

3. *Conversie van gestandaardiseerd vragenlijst- onderzoek naar internet: kritische succes- factoren*

E.C. DEUTSKENS, P.M. WILLEMS, S.S.M. VAN DEN BERG, P. OOSTERVELD,
J.C. DE RUYTER en M.G.M. WETZELS

SAMENVATTING

In de online marktonderzoeksliteratuur zijn de algemeen geaccepteerde richtlijnen voor het inrichten van online vragenlijstonderzoek nog volop in ontwikkeling. De best practices schieten tekort op het moment dat omvangrijke, gestandaardiseerde schriftelijke vragenlijsten worden omgezet naar internet. In een experiment is onderzocht wat de invloed is van beloningsvorm, vragenlijstlengte, presentatievorm en timing van reminders op de deelnamebereidheid van respondenten aan een uitgebreide online vragenlijst. Uit de resultaten bleek dat lange vragenlijsten weliswaar een lagere respons hebben, maar dat er toch een bevredigend responspercentage mee behaald kan worden. Daarnaast was loterij de meest effectieve beloningsvorm, aangezien met een loterij min of meer hetzelfde responspercentage werd verkregen als met waardebonnen, maar een loterij kostefficiënter is. Hoewel een vragenlijst met visuele productpresentaties een wat lagere respons had dan de tekstuele versie, was de responskwaliteit wel hoger.

1. INTRODUCTIE

Voor de ontwikkeling van internetmarktonderzoek tot volwaardig alternatief is het van belang dat mogelijkheden en grenzen van dit veldwerkkanaal expliciet worden gemaakt. Op dit moment bestaan er voor vragenlijstonderzoek via internet slechts richtlijnen of 'best practices'. De bekende regels zoals dat een vragenlijst via internet niet langer dan 10 minuten invultijd mag vergen, zijn te weinig gebaseerd op empirische bevindingen en betekenen in de praktijk een beperking voor de inzet van internet in het brede scala van onderzoeksproducten.

Er is een toenemende behoefte om ook uitgebreidere studies via internet uit te voeren. Onderzoekers van zowel bureauezijde als vanuit de internationale producenten en dienstverleners staan te trappelen om deze volgende stap te zetten. Gestandaardiseerde, omvangrijke meetinstrumenten die nu meestal nog via schriftelijke vragenlijsten worden afgenomen, roepen om een conversie naar internet. De omvangrijke handelingskosten, lange doorlooptijd, slechtere bereikbaarheid van respondenten en de afnemende respons op schriftelijke vragenlijsten doen een beroep op onderzoeksmethodologen om het instrument geschikt te maken voor internet als responsmedium.

Er zijn twee fundamentele problemen bij de conversie van uitgebreide, gestandaardi-

seerde vragenlijsten naar internet. Voor de eerste, trendbreuk, is inmiddels niemand meer zo bang. Trendbreuk is het eenmalige gevolg van de omzetting van het veldwerk naar een ander kanaal en betekent dat gemiddelde scores op vragenlijstitems niet langer rechtstreeks met benchmarks vergeleken kunnen worden. De oplossing voor trendbreuk is calibratie. In een experimenteel design waarin op aselechte wijze respondenten aan verschillende veldwerkvarianten van de vragenlijst worden toegewezen, kunnen correctiecoëfficiënten worden bepaald waarmee toekomstige resultaten alsnog met benchmarks vergeleken kunnen worden (Willems et al., 2002).

De tweede reden waarom onderzoekers op dit moment nog huiverig zijn voor het omzetten van standaardvragenlijsten naar internet bestaat uit het feit dat men onvoldoende zeker is van de kwaliteit en kwantiteit van respons via internet op dit soort uitgebreide vragenlijsten. Bestaat een ad-hoonderzoek via internet doorgaans uit vijftienving vragen, gestandaardiseerde paper-and-pencil vragenlijsten kennen vaak meerdere pagina's met stellingenblokken en tientallen achtergrondvragen. Het is vervolgens onduidelijk of respondenten bereid zijn om dit soort lange vragenlijsten via internet in te vullen, welke beloningsvorm moet worden gekozen als compensatie voor deelname en hoe de vormgeving van de vragenlijst moet worden aangepast voor presentatie op internet.

Experimenteel onderzoek kan ook ten aanzien van dit tweede probleem praktische richtlijnen aanreiken om conversie naar internet van gestandaardiseerde vragenlijsten te faciliteren. In dit artikel worden resultaten gepresenteerd van een studie waarin deze conversie experimenteel is onderzocht. Deze studie is het resultaat van een samenwerking tussen Centrum voor Marketing Analyses en de Universiteit Maastricht in het kader van een onderzoeksprogramma betreffende de equivalentie tussen online en traditionele marktonderzoeksmethoden. Meerdere varianten van een gestandaardiseerd, in de FMCG branche alom bekend, multicient instrument zijn via internet aangeboden aan een uitgebreide steekproef van respondenten. De lengte van deze oorspronkelijk paper-and-pencil vragenlijst vroeg om een creatieve omzetting naar internet. Daarvoor is eerst een wetenschappelijke inventarisatie gemaakt van de belangrijkste succesfactoren voor een online vragenlijst.

2. KRITISCHE SUCCESFACTOREN VOOR EEN ONLINE VRAGENLIJST

Voor een aantal factoren bij het ontwikkelen van een online vragenlijst bestaan reeds eenduidige richtlijnen. Voorbeelden hiervan zijn onder andere het tonen van voortgangsinformatie, het bieden van zekerheid wat betreft anonimiteit en geheimhouding en het aanbieden van online help.

Feedback over de voortgang van het invullen van de vragenlijst vermindert het aantal respondenten dat voortijdig afhaakt (Couper et al., 2001). Voortgangsinformatie vormt een belangrijke motivator. De visuele presentatiemogelijkheden van een online vragenlijst maken het bij uitstek mogelijk respondenten te informeren over de voortgang van het invulproces.

Het benadrukken van anonimiteit en het opnemen van een verklaring tot geheimhouding op de vragenlijstpagina('s) vormt eveneens een onbetwiste voorwaarde voor een goede respons. Studies van onder meer Sheehan en Hoy (1999) tonen aan dat respondenten zich bewust zijn van de vertrouwelijkheid van de door hen gegeven antwoorden

en dat ze actief op zoek gaan naar informatie in de vragenlijst die wijst op een zorgvuldige behandeling van de gegeven antwoorden.

Online help wordt bij internetvragenlijsten vaak aangeboden in de vorm van een zogenaamde FAQ (Frequently Asked Questions list). Het aanbieden van dit soort informatie vergroot de slagingskans van een online onderzoek substantieel (Bauman et al., 2000). Belangrijke elementen van FAQ zijn het introduceren van de onderzoeksorganisatie, antwoorden op vragen of problemen die kunnen optreden, het verzekeren van vertrouwelijkheid van antwoorden en het geven van een contactnaam of e-mailadres voor verdere vragen. Bij voorkeur zou een link naar de FAQ op elke pagina van de vragenlijst aangeboden moeten worden.

Voor een aantal andere vragenlijstkenmerken geldt dat richtlijnen minder eenduidig zijn. Op het moment dat een vragenlijst bijzondere eisen stelt aan het invulproces, bijvoorbeeld door de lengte van de vragenlijst, lijken bevindingen uit de literatuur niet langer duidelijkheid te verschaffen over de inrichting van de vragenlijst. Vier van deze kenmerken worden hieronder belicht. Ze vormen de onafhankelijke variabelen in de experimentele studie.

2.1. Beloningsvorm

Het gebruik van financiële beloningen of incentives is effectief in zowel offline als online onderzoek (Yammarino et al., 1991; Bergmann et al., 1990; The Future Place Consultancy, 2001). Een incentive kan respondenten een reden geven om een vragenlijst in te vullen en fungeert tevens als een teken van dank namens de klant. Incentives kunnen geclassificeerd worden naar de manier waarop beloond wordt: individuele incentives belonen iedereen die deelneemt aan het onderzoek, zoals bijvoorbeeld in het geval van waardebonnen of spaarsystemen. In andere gevallen worden slechts enkele van de deelnemende respondenten beloond, bijvoorbeeld in het geval van loterijen. Daarnaast bestaat er de beloning in de vorm van donaties aan goede doelen. Hierbij ontvangt de respondent zelf niet rechtstreeks een beloning.

Resultaten van een aantal eerdere studies geven aanleiding te kiezen voor individuele, financiële incentives (Warriner et al., 1996; James en Bolstein, 1990; James en Bolstein, 1992), zelfs in situaties waarin slechts symbolische bedragen worden gegeven. Echter, Frick et al. (1999) vonden dat loterijen eveneens resulteren in een relatief lage uitval van respondenten. Daarbij werd een additioneel argument ten gunste van een loterijbeloning genoemd, namelijk dat de kosten voor de respondentbeloning onafhankelijk kunnen zijn van het aantal deelnemers in het onderzoek. Het namens respondenten doneren van geld aan goede doelen, tenslotte, doet een appèl op sociaalnormatieve waarden. Het ingaan op het verzoek deel te nemen aan een vragenlijstonderzoek heeft een gunstig effect op de beleving van eigenwaarde van de respondent (Warriner et al., 1996). Studies tonen echter geen helder verband aan tussen het gebruik van de donatiebeloningsvorm en respons op de vragenlijst.

Een vergelijking van de effecten van individuele, financiële beloningen met loterijen en donaties in schriftelijke vragenlijsten geven dan ook geen eenduidig beeld (Warriner et al., 1996). Omdat beloningen een substantiële component vormen van het onderzoeksbudget, dient de invulling ervan nauwkeurig te worden bepaald. Om die reden vormt beloning een belangrijke factor in het experiment in dit artikel.

2.2. Lengte van de vragenlijst

Resultaten van onderzoek naar de relatie tussen het aantal vragen in een vragenlijst en het responspercentage zijn allerm minst eenduidig. Terwijl sommige studies suggereren dat de lengte van de vragenlijst geen invloed heeft op de respons (Brown, 1965; Bruvold en Comer, 1988; Mason et al., 1961; Fox et al., 1988; Linsky, 1975), wijzen andere onderzoeken erop dat er wel degelijk een relatie bestaat tussen vragenlijstlengte en het responspercentage, zowel in off- als online marktonderzoek (Dillman, 1991; Heberlein en Baumgartner, 1978; Steele et al., 1992; Yammarino et al., 1991; MacElroy, 2000; Sheehan, 2001).

Uit onderzoek door Marketfacts (2001) bleek dat respondenten van een online vragenlijst veelal vinden dat vragenlijsten te lang zijn. Omdat respondenten met één klik kunnen afhaken, lijkt het erop dat uitsluitend korte vragenlijsten tot een hogere respons leiden. Het is echter niet duidelijk wat precies onder kort moet worden verstaan. Volgens Rosenblum (2001) zou een online vragenlijst uit ongeveer twintig vragen moeten bestaan. Dit lijkt echter onrealistisch klein voor de meeste uitgebreide onderzoeksvragen.

Wat we kunnen verwachten bij de conversie naar internet van een vragenlijst met vragen over 190 producten is daarom niet precies te zeggen. De afname van respons als functie van vragenlijstlengte vormt dan ook een belangrijk aandachtspunt in het hieronder beschreven experiment.

2.3. Presentatie van de vragenlijst

In internetmarktonderzoek kan de vraagtekst worden aangevuld met visuele elementen, zoals kleur, afbeeldingen en interactieve elementen (bijvoorbeeld de 3D-presentatie van producten). Dit zou het aantrekkelijker kunnen maken voor respondenten om de online vragenlijsten in te vullen (Redline en Dillman, 1999). Bovendien kunnen productafbeeldingen ook effectief zijn bij het verhogen van de kans op productherkenning, aangezien respondenten meer associatiemogelijkheden krijgen aangeboden. Een probleem van bij het gebruik van afbeelding is de laadtijd (downloadtijd). Relatief eenvoudige internetvragenlijsten die weinig downloadtijd vragen, leiden dan ook in sommige gevallen tot een hoger responspercentage dan vragenlijsten met een overdaad aan grafische elementen (Dillman et al., 1998).

Het lijkt erop dat de voordelen van aantrekkelijkheid en hogere productherkenning dienen te worden afgewogen tegen nadelen als langere downloadtijd voor respondenten en de kosten die gepaard gaan met programmeren. De cruciale vraag is of grafische elementen resulteren in betere resultaten die tijd en andere investeringen rechtvaardigen. Op basis van de resultaten van eerder onderzoek verwachten we dat afbeeldingen in een online vragenlijst zowel tot hogere productherkenning als tot hogere responskwaliteit zullen leiden, maar dat het responspercentage lager zal zijn vanwege de langere downloadtijd.

2.4. Reminders

Reminders, of follow-ups, beïnvloeden het responspercentage positief bij schriftelijke vragenlijsten die per post worden verzonden (Sheehan, 2001; Solomon, 2001). Van een briefkaartreminder wordt verwacht dat die het responspercentage verhoogt om twee

redenen. Ten eerste wordt verondersteld dat de extra aandacht die blijkt uit het sturen van een reminder de medewerking stimuleert. In de tweede plaats vormt de reminder ook letterlijk een geheugensteuntje omdat respondenten nogal eens vergeten de vragenlijst in te sturen (Fox et al., 1988). Reminders in schriftelijke studies kunnen het responspercentage 8 tot 48% verhogen (Sheehan, 2001). Bij internet ligt het voor de hand vooral het tweede effect van reminders te veronderstellen, namelijk de herinneringswerking ervan. In een studie naar onderzoek via e-mail (Solomon, 2001) werd gevonden dat het mogelijk is om met reminders het responspercentage te verdubbelen. Hoewel het verzenden van meerdere reminders in online onderzoeken vrijwel kosteloos is, moet zorgvuldig met dit middel omgesprongen worden. Het valt te verwachten dat het overdadig sturen van reminders gevoelens van irritatie bij respondenten wekt en op langere termijn kan leiden tot een algemene vermindering in de bereidheid van respondenten om via internet aan onderzoek mee te doen.

Het is voorts niet duidelijk wat het beste moment is om een reminder te versturen. Het moment van versturen van reminders wordt in de onderhavige studie experimenteel gemanipuleerd. We verwachten dat wanneer een reminder vrij snel na de uitnodiging wordt verstuurd, dit een hoger responspercentage oplevert dan wanneer de reminder pas na twee weken wordt verstuurd.

3. METHODE

3.1. De vragenlijst

De vragenlijst die in dit experiment is geconverteerd naar internet, betreft een onderzoeksinstrument voor het meten van het gebruik van consumentenproducten binnen huishoudens en de bekendheid van merken. De vragenlijst wordt oorspronkelijk in schriftelijke vorm afgenomen. Resultaten worden in de vorm van multiclient participatie afgenomen door verschillende producenten. Voorbeelden van productgroepen in de vragenlijst zijn bier, shampoo, kaas, olijfolie en toast. Iedere productgroep bevat een aantal merken waarvan respondenten moeten aangeven of ze deze kennen, hoe vaak de producten gebruikt worden en hoe de prijs-kwaliteitverhouding wordt beoordeeld. De vragenlijst is een lange, uitgebreide vragenlijst die van respondenten al snel gemiddeld 30 minuten invultijd vraagt.

3.2. Steekproef

De deelnemers aan het onderzoek werden uit een database gerekruteerd. Deze bestaat uit namen en e-mailadressen van mensen in Nederland die verzameld zijn door middel van eerder telefonisch contact. De steekproef werd verdeeld naar de Minicensus, op basis van de volgende demografische gegevens:

- leeftijd (13-74) en sociale klasse
- geslacht
- regio (5).

Onze steekproef bestond uit 5.413 e-mailadressen, verkregen van 1999 tot en met 2002. We besloten om niet alle uitnodigingen in één keer uit te mailen, maar in plaats daar-

van in batches die we één voor één uitmailden. In batch 1 verdeelden we alle respondenten gelijkmatig over de verschillende cellen van ons design (zie paragraaf 3.3). De verdeling in de latere batches, batch 2, 3 en 4, werd aangepast aan de tot dan toe behaalde respons in iedere cel, om ervoor te zorgen dat de absolute respons in iedere cel aan het einde van het onderzoek ongeveer gelijk zou zijn.

3.3. De onafhankelijke variabelen

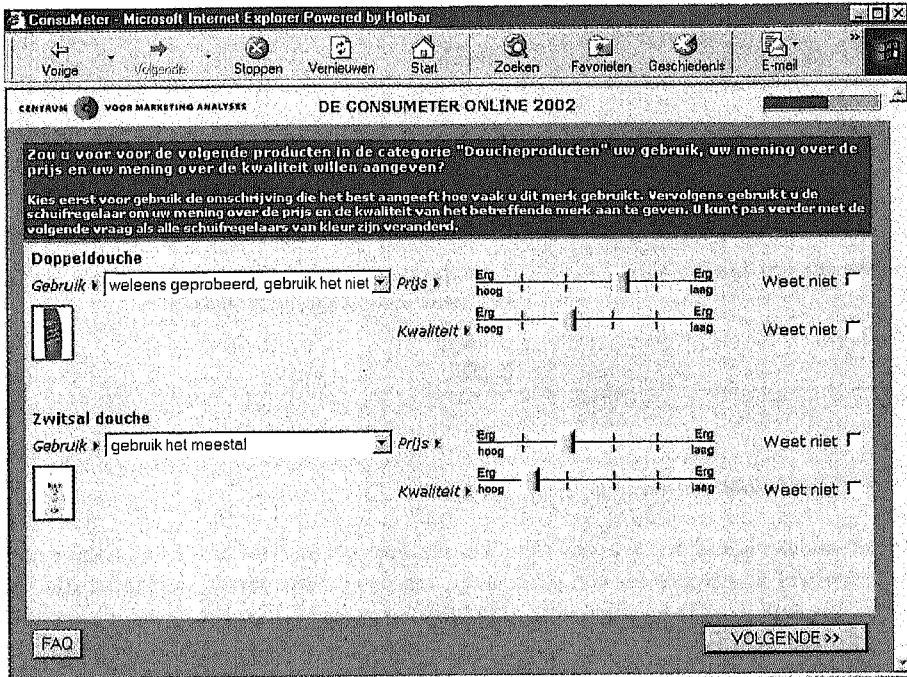
Experimentele factoren in het onderzoek waren het soort incentive, de lengte van de vragenlijst, de presentatie van de vragenlijst en de timing van de reminder (zie Tabel 1).

Beloningsvorm. Drie soorten beloningen zijn in de studie meegenomen: een waardebon, donatie aan een goed doel en een loterij. Participanten in de waardebonconditie kregen een waardebon voor een online boeken- en platenwinkel na het completeren van de vragenlijst. We kozen voor waardebonnen in plaats van contant geld aangezien het meer tijd en moeite kost om de uitbetaling van contant geld af te handelen. De hoogte van het bedrag was afhankelijk van de lengte van de vragenlijst: € 2,50 voor de korte versie, € 5,00 voor de lange versie. De donatieconditie hield in dat namens de respondenten een bepaald bedrag aan een goed doel werd gedoneerd als dank voor hun deelname. Daarbij kon men kiezen uit drie goede doelen: Wereld Natuur Fonds, Amnesty International of de Kankerbestrijding. De loterijconditie hield in dat respondenten een kans hadden een waardebon van een relatief hoog bedrag te winnen. In de korte vragenlijst maakten respondenten kans op één van de vijf waardebonnen van € 25,00, in de lange vragenlijst op één van de vijf waardebonnen van € 50,00.

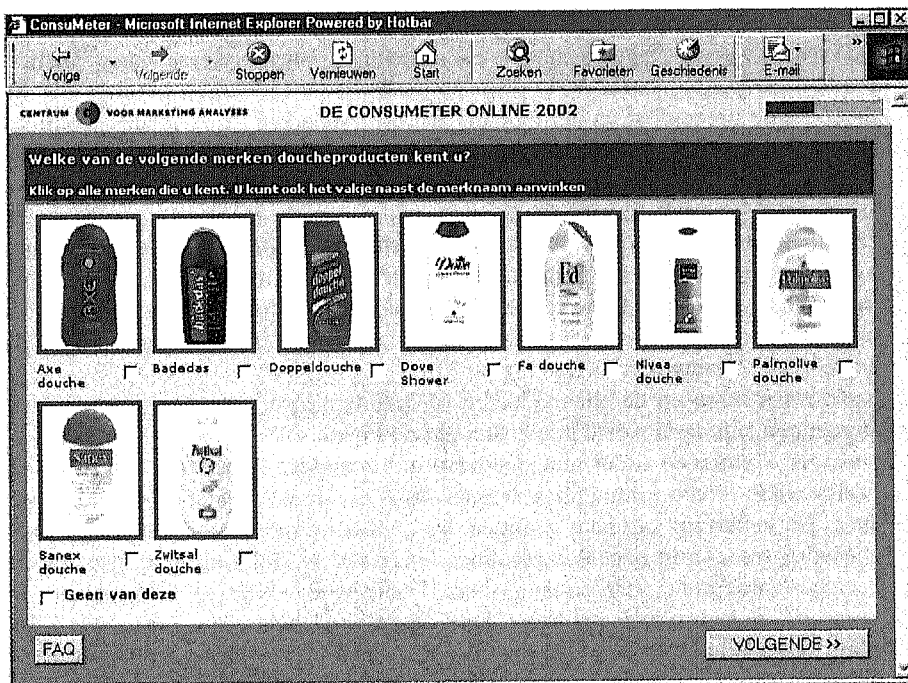
Lengte van de vragenlijst. In het onderzoek zijn een lange en een korte versie van de vragenlijst gebruikt. De lange versie bevatte evenveel vragen en productcategorieën (19) als de schriftelijke vragenlijst, terwijl de korte versie de helft van het aantal vragen en productcategorieën (9) van de schriftelijke vragenlijst bevatte. De lange versie duurde ongeveer 30 tot 45 minuten om in te vullen, de korte 15 tot 30 minuten.

Presentatie van de vragenlijst. Met de vragenlijst wordt voor een groot aantal producten en merken (190 voor de lange vragenlijst, 91 voor de korte) vastgesteld of respondenten deze kennen en gebruiken en hoe men deze beoordeelt wat betreft prijs en kwaliteit. In de schriftelijke vragenlijst moeten respondenten deze vragen beantwoorden op basis van de naam van de producten en merken alleen. In de presentatie van de online vragenlijst onderscheiden we twee varianten: tekstuele presentatie van producten, waarbij respondenten alleen de naam van het product te zien krijgen en visuele presentatie van producten, waarbij respondenten naast de naam ook een afbeelding van het product zien (zie Figuur 1 en 2).

Timing van de reminder. De vierde factor die we in deze studie varieerden, was de timing van de reminder en wel op twee niveaus: vroeg en laat. Respondenten in de vroege reminderconditie ontvingen een reminder vijf werkdagen nadat ze de e-mail met uitnodiging hadden ontvangen. Respondenten in de late reminderconditie ontvingen de reminder tien werkdagen na ontvangst van de e-mail en uitnodiging. Deze reminder werd alleen naar diegenen gestuurd die de vragenlijst nog niet hadden ingevuld.



Figuur 1. Productherkenning in de visuele vragenlijst.



Figuur 2. Productherkenning in de tekstuele vragenlijst.

Tabel 1. Overzicht factoren en niveaus.

Factor	Niveau
1. Beloningsvorm	a. waardebon b. donatie (aan goed doel) c. loterij
2. Lengte van de vragenlijst	a. lang (19 productcategorieën) b. kort (9 productcategorieën)
3. Presentatie van de vragenlijst	a. tekstuele presentatie van producten b. visuele presentatie van producten
4. Reminder	a. vroeg b. laat

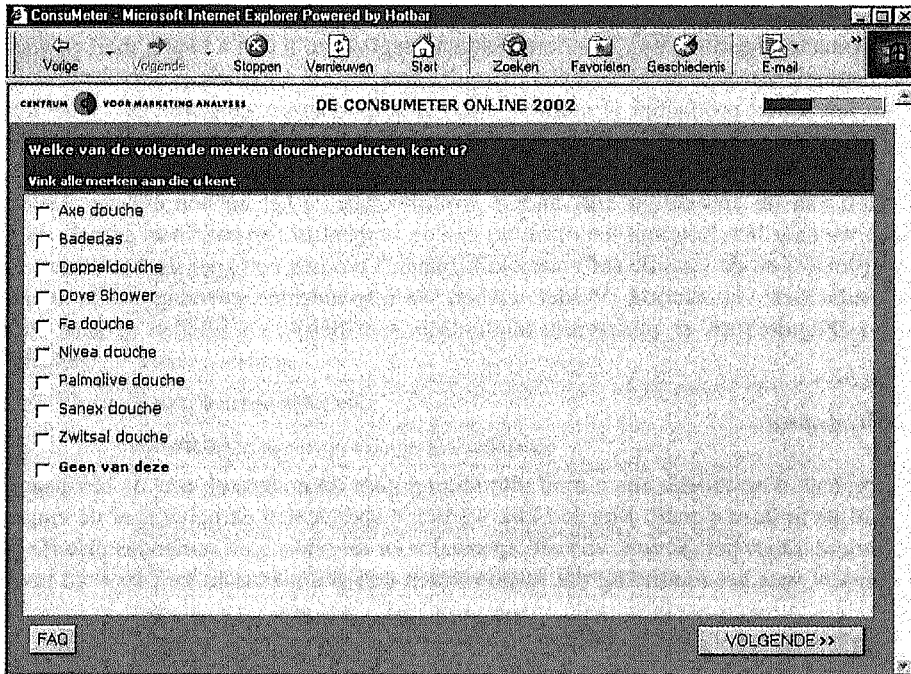
3.4. De afhankelijke variabelen

In deze studie hanteerden we vijf afhankelijke variabelen. Ten eerste het responspercentage, ofwel de respons als een percentage van de gecontacteerde steekproef (dat wil zeggen, het aantal volledig ingevulde vragenlijsten gedeeld door het aantal uitgemailde vragenlijsten). We kunnen onderscheid maken tussen bruto en netto responspercentages. Daarbij wordt het bruto responspercentage berekend op basis van het totaal aantal uitgemailde vragenlijsten. Echter, omdat een groot aantal e-mailadressen fout is of niet meer bestaat, komt het nogal eens voor dat vragenlijsten niet aankomen bij potentiële respondenten. Daarom hanteren we in deze studie het netto responspercentage, ofwel de respons op basis van het aantal werkelijk aangekomen e-mails is. E-mails die niet bezorgd konden worden, zijn bij dit responsepercentage buiten beschouwing gelaten.

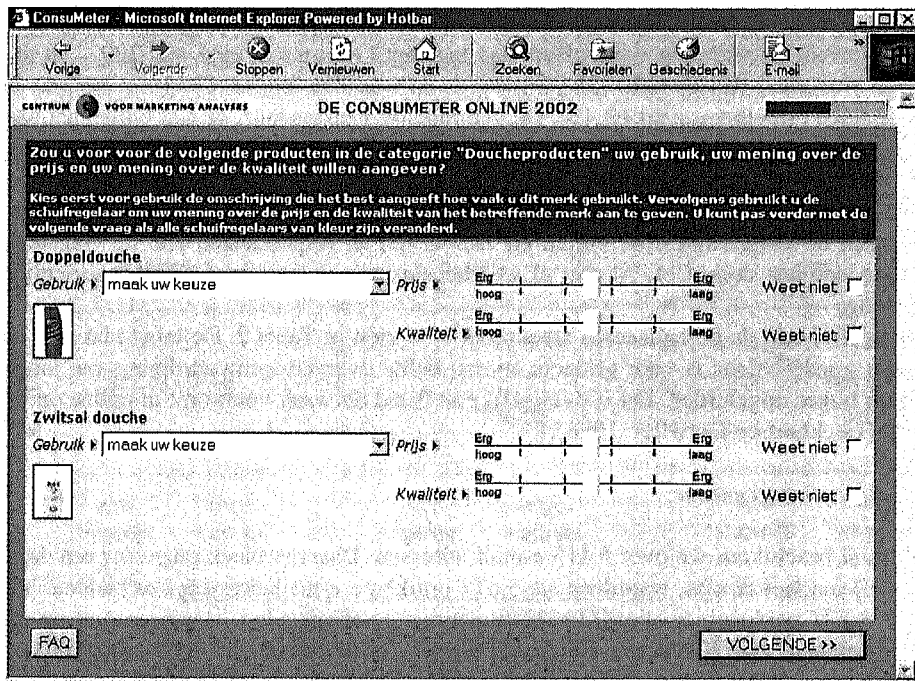
De tweede afhankelijke variabele is responstijd. Dit is de tijd tussen het moment van uitmailen van de vragenlijst en het moment dat de respondent de vragenlijst compleet heeft. De responstijd werd gemeten in dagen. De derde afhankelijke variabele, responskwaliteit, is op twee manieren onderzocht. Ten eerste door te kijken naar de gemiddelden en varianties van de antwoorden op de prijs- en kwaliteitvraag. Aangezien er 190 merken/producten in de vragenlijst voorkomen, zijn er 190 prijsvragen en 190 kwaliteitvragen. Dit zijn de vragen waarin respondenten antwoord moeten geven met behulp van een zogenaamde slider. Deze slider staat standaard in het midden (op waarde 'noch hoognoch laag' of 'noch goednoch slecht') en is dan grijs van kleur, wat betekent dat de vraag nog niet beantwoord is (zie Figuur 3).

Respondenten moeten op de slider klikken of deze bewegen, waardoor die van kleur verandert er een antwoord wordt geregistreerd (zie Figuur 4).

Respondenten konden de slider niet standaard in het midden laten staan zonder iets te doen, dan volgde er een foutmelding waarin stond dat de vraag niet beantwoord was. Daarom is het belangrijk om naar gemiddelden en met name varianties te kijken om te achterhalen of mensen in bepaalde condities heel vaak de slider uit gemakszucht niet eens bewegen maar in het midden laten staan. Ten tweede hebben we de responskwaliteit getest door te kijken naar het aantal keer dat respondenten het 'weet ik niet' vakje aanvinken per productcategorie. Over het algemeen wordt aangenomen dat wanneer respondenten veelvuldig 'weet ik niet' aanklikken dit de responskwaliteit verslechtert (Singer et al., 2000; Goetz et al., 1984).



Figuur 3. Sliders in onbeantwoorde situatie.



Figuur 4. Sliders in beantwoorde situatie.

De vierde afhankelijke variabele is productherkenning. De vragenlijst bevatte negentien productcategorieën met in iedere productcategorie zo'n tien à twaalf merken waarvan respondenten aan moeten geven of ze die kennen. De variabele productherkenning is dan het aantal producten of merken waarvan respondenten aangeven dat ze die kennen per productcategorie.

Tot slot hebben we gekeken naar de beleving van de respondenten ten aanzien van het invullen van de vragenlijst, met andere woorden hoe ze het hebben ervaren. Daarbij keken we naar hun houding ten opzichte van de vragenlijst (op basis van gemak, moeite en plezier) en de intentie (of waarschijnlijkheid) om nog een keer deel te nemen aan een onderzoek via internet. Verder hebben we respondenten gevraagd welke manier van marktonderzoek ze prefereren: telefonisch, schriftelijk, via internet of geen voorkeur.

3.5. Procedure

Respondenten ontvingen een e-mail uitnodiging voor dit onderzoek met de link naar de vragenlijst in deze e-mail. Met één klik werden respondenten derhalve naar de vragenlijst geleid. Door het gebruik van een speciaal voor dit onderzoek samengestelde ID per respondent was het niet nodig dat respondenten een gebruikersnaam en password invulden.

4. RESULTATEN

4.1. Gegevensverzameling

De gegevensverzameling vond plaats van half april tot en met eind mei 2002. Na drie dagen was meer dan de helft (52,9%) van de antwoorden binnen. Na één week was dit percentage opgelopen naar 63,4%. De gemiddelde responsduur, de tijd tussen het aanklikken van de link en completering van de lijst, was ongeveer 45 minuten. De belangrijkste reden voor deze relatief lange tijd is waarschijnlijk dat sommige respondenten het invullen onderbreken om hem later verder in te vullen. 16,5% deed er een kwartier over om de vragenlijst in te vullen en 78,9% van de respondenten had de vragenlijst binnen een uur compleet. Al met al was het dus een vragenlijst van een aanzienlijke omvang.

Het profiel van de gerealiseerde steekproef is te zien in Tabel 2. De tabel laat zien dat er een goede balans is voor geslacht en een lichte oververtegenwoordiging van jongeren en hoger opgeleiden. Dit is overigens een beeld dat vaak voorkomt in online onderzoek (Lockhart en Detwiler, 1998).

4.2. Responspercentage

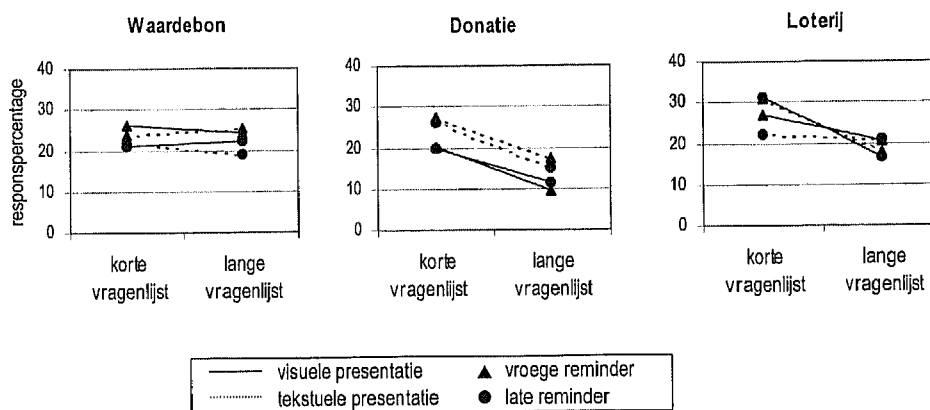
In totaal beschikten we over 5.413 e-mail adressen. Daarvan bleek ongeveer een derde (1.836) inactief te zijn, waardoor we 3.577 bruikbare e-mailadressen overhielden. Van deze 3.577 personen hebben 730 de vragenlijst gecompleteerd, wat neerkomt op een netto respons van 20,4%. Er zijn significante verschillen ($\chi^2_{23} = 74,8$, $p < 0,001$) in responspercentage tussen de cellen van het design waarbij de korte, visuele versie van de vragenlijst met loterij als incentive en een late reminder een respons van 31,4% ken-

Tabel 2. Profiel van de gerealiseerde steekproef.

Geslacht	Man	51,5%
	Vrouw	48,5%
Leeftijd	15-24 jaar	12,5%
	25-34 jaar	23,7%
	35-49 jaar	40,0%
	50-64 jaar	22,1%
	65-74 jaar	1,8%
Opleiding	Lager onderwijs	1,0%
	Lager beroepsonderwijs	4,0%
	Middelbaar algemeen voortgezet onderwijs	11,2%
	Middelbaar beroepsonderwijs	31,4%
	Hoger algemeen en voorbereidend wetenschappelijk onderwijs	10,0%
	Hoger beroepsonderwijs en wetenschappelijk onderwijs	30,0%
	Wetenschappelijk onderwijs - doctoraal	10,4%
Wil niet zeggen	2,1%	

de, terwijl de lange visuele versie met donatie als incentive en een vroege reminder bleef steken bij een respons van slechts 9,4%. In Figuur 5 staan de responspercentages van de combinaties van de factoren weergegeven.

De responspercentages voor de hoofdeffecten van de factoren staan in Tabel 3. Voor 'type incentive' vonden we een significant verband tussen incentive en responspercentage. De waardebon en de loterij hadden vrijwel gelijke responspercentages van respectievelijk 22,9% en 23,0%, terwijl de donatiegroep een responspercentage van 16,6%



Figuur 5. Interactie-effecten van beloningsvorm, vragenlijstlengte, presentatievorm en timing van de reminder.

Tabel 3. Hoofdeffecten responspercentages.

Factor		Lengte van de vragenlijst		Presentatie van de vragenlijst		Reminder	
waardebon	22,9%	lang	17,1%	tekstueel	21,9%	vroeg	21,2%
donatie	16,6%	kort	24,6%	visueel	19,1%	laat	19,6%
loterij	23,0%						

Noot: $\chi^2_2 = 20,9$; $p < 0,001$ (beloningsvorm); $\chi^2_1 = 30,3$; $p < 0,001$ (lengte); $\chi^2_1 = 4,4$; $p = 0,04$ (presentatie); $\chi^2_1 = 1,4$; $p = 0,23$ (reminder).

had (overigens begunstigde het merendeel, 61%, de Kankerbestrijding, 25% het Wereld Natuur Fond en 15% Amnesty International). Zoals verwacht, bleek er een significant hoger responspercentage voor de korte versie van de vragenlijst dan voor de lange versie te zijn: 17,1% voor de lange, 24,6% voor de korte versie. Ook de verwachting omtrent het lagere responspercentage voor de visuele versie van de vragenlijst in vergelijking met de tekstuele versie is uitgekomen. Het responspercentage voor de tekstuele versie was 21,9% en voor de visuele versie 19,1%. De nulhypothese omtrent de timing van de reminder kon niet worden verworpen.

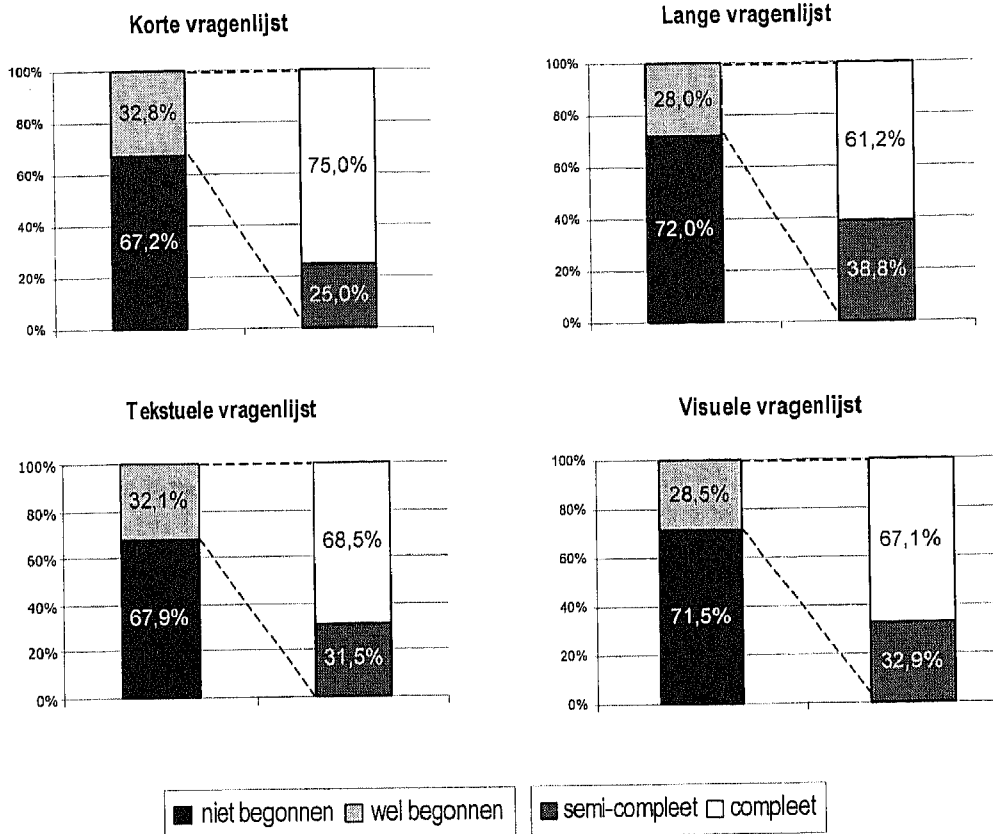
Waarom is de respons voor de lange vragenlijst zoveel lager dan voor de korte vragenlijst? Is dat omdat respondenten worden afgeschrikt door de aankondiging in de e-mail uitnodiging dat het invullen van de vragenlijst ongeveer 30 minuten zal duren? Of is dat omdat respondenten tijdens het invullen van de vragenlijst ervaren dat de vragenlijst lang en eentonig is en stoppen ze er dan mee?

Om nu te achterhalen of de uitnodiging of ervaring een belangrijkere factor is, hebben we gekeken naar hoe groot het percentage semi-completes is ten opzichte van het totaal aantal respondenten dat wel aan de vragenlijst is begonnen. Op basis van de uitnodiging begint 32,8% van de respondenten in de conditie korte vragenlijst met het invullen, in de conditie lange vragenlijst is dat marginaal lager (28,0%). Het percentage uitval laat wel een groot verschil zien, namelijk 25,0% versus 38,8% (zie Figuur 6). Er is dus een effect van ervaring.

Een soortgelijke vraag kunnen we ook stellen met betrekking tot de presentatie van de vragenlijst. Is het responspercentage voor de visuele versie van de vragenlijst lager vanwege de frustratie van de langere download tijd die gepaard gaat met de afbeeldingen of heeft het toch een andere oorzaak? Wanneer we hier kijken naar hoe groot het percentage semi-completes is ten opzichte van het totaal aantal respondenten dat wel aan de vragenlijst is begonnen, dan zien we dat ook deze percentages heel dicht bij elkaar liggen: 31,5% voor de tekstuele versie, 32,9% voor de visuele (zie Figuur 6). Het lagere responspercentage voor de visuele versie van de vragenlijst wordt kennelijk niet veroorzaakt door de langere downloadtijd van de vragenlijst vanwege de afbeeldingen.

4.3. Responstijd

De verwachting was dat de respondenten in de loterij- en in de korte vragenlijstconditie sneller de vragenlijst zouden completeren. De respondenten hadden na gemiddeld 6,6 dagen de vragenlijst ingevuld. Alleen voor type incentive vonden we een significant



Figuur 6. Effect van uitnodiging vs. ervaring op basis van non-response, semi-compleet en compleet voor vragenlijstlengte en presentatievorm.

verschil in gemiddelde responstijd ($F_{2, 727} = 3.8$, $p < 0.05$). Voor de waardebonconditie was dit 7,6 dagen, voor de donatieconditie 7,3 dagen en voor loterijconditie 5,9 dagen.

4.4. Responskwaliteit

Zoals hierboven vermeld, kijken we bij responskwaliteit enerzijds naar de gemiddelden en varianties van de vragen over prijs en kwaliteit en anderzijds naar het aantal weet niet responsen op deze vragen. Dit houdt in dat we een groot aantal toetsen moeten uitvoeren, wat de kans op het vinden van toevallige significanties vergroot. Daarom zullen we bij het vinden van significante resultaten nagaan of het op basis van het significantieniveau te verwachten aantal ligt. Bij een significantieniveau van 5% is immers de verwachting dat 5% van de nulhypothesen onterecht zal worden verworpen.

Bij de eerste benadering vonden we bij geen van de factoren die we gemanipuleerd hebben vaker significantie verschillen in gemiddelden en standaarddeviaties dan op basis van toeval kon worden verwacht. We gaan er daarom van uit dat er voor dit

aspect van responskwaliteit geen effecten zijn van lengte, presentatie en incentive. Voor het aantal maal dat respondenten 'weet ik niet' invullen, hadden we verwacht dat dit aantal groter zou zijn voor de lange versie. Er bestaat echter wel een verschil voor dit aspect van responskwaliteit voor de presentatie van de vragenlijst. Bij vijf van de productcategorieën rapporteren respondenten met de tekstuele versie van de vragenlijst significant vaker 'weet ik niet'.

4.5. Productherkenning

We hadden verwacht hier een verschil te vinden bij de presentatie van de vragenlijst, waarbij productherkenning hoger is voor de visuele versie van de vragenlijst. Ook hier volgden we een zelfde toetsingprocedure als bij responskwaliteit: de toetsen worden per productcategorie uitgevoerd, de nulhypothese wordt verworpen als het aantal significante resultaten groter is dan op basis van kans kan worden verwacht. Bij vier van de negentien productcategorieën is er sprake van een significant verschil. In tegenstelling tot de verwachting is echter de productherkenning hoger voor de tekstuele versie.

4.6. Houding

In Tabel 4 zijn de gemiddelden voor houding en intentie weergegeven. De respondenten vonden het doorgaans makkelijk om de vragenlijst in te vullen en het kostte ze ook niet veel moeite. Wat betreft het plezier waren ze neutraal. De respondenten waren gematigd positief over de waarschijnlijkheid dat ze nog een keer deel zullen nemen aan een dergelijk onderzoek, waarbij men duidelijk de voorkeur had voor deelname via internet.

Wanneer we kijken of er verschillen in houding en intentie zijn voor de verschillende versies van de vragenlijst, dan blijken er alleen voor de factor lengte significante verschillen op te treden. Respondenten met de korte versie zijn positiever over gemak (26,4 vs. 30,8), moeite (33,1 vs. 45,6), plezier (43,4 vs. 53,5) en intentie (2,1 vs. 2,4). In termen van effectgrootte (het verschil ten opzichte van de standaarddeviatie) moet men hier spreken van kleine tot middelgrote effecten.

5. DISCUSSIE

In dit artikel hebben we getracht inzicht te verschaffen in de manier waarop online marktonderzoek het best kan worden ingericht en uitgevoerd wanneer het gaat om de inzet van een oorspronkelijk zeer uitgebreide schriftelijke vragenlijst.

De behaalde respons in onze studie mag een opvallend resultaat genoemd worden, aangezien de respondenten zijn geworven uit een database met personen die eenmalig telefonisch zijn benaderd. Deze personen hebben wel aangegeven dat ze opnieuw benaderd mogen worden, maar vaak was dit al zeer lang geleden. Voor ongeveer een derde van de respondenten die we voor dit onderzoek hebben uitgenodigd was dit telefonisch contact zelfs in 1999. Als je ziet dat 80 à 90% van de online onderzoeken momenteel met behulp van panels wordt uitgevoerd, dan is het resultaat van ons onderzoek waarbij we met respondenten met minder commitment dan een panel toch nog een respons van ruim 31% (max.) kunnen behalen, positief voor de toekomst van online marktonderzoek. Het behaalde responspercentage is vooral opvallend wanneer we in oenschouw nemen

Tabel 4. Gemiddelde houding van respondenten ten opzichte van de vragenlijst en preferentie voor marktonderzoek.

Houding	Gemiddelde (sd)
Gemak (makkelijk 0 – moeilijk 100)	28,5 (16,6)
Moeite (weinig moeite 0 – veel moeite 100)	39,0 (23,3)
Plezier (plezierig 0 – onplezierig 100)	48,2 (21,6)
Intentie	Gemiddelde (sd)
Intentie (zeer waarschijnlijk 1 – zeer onwaarschijnlijk 5)	2,3 (1,1)
Welke manier van marktonderzoek prefereert u	Percentage (%)
Telefonisch	5,1
Schriftelijk	13,2
Via Internet	66,6
Geen voorkeur	16,2

dat het hier om een voor online onderzoek ongebruikelijke lange vragenlijst gaat. Uiteraard vonden we een duidelijk effect voor lengte: de korte vragenlijst had een hoger responspercentage dan de lange vragenlijst. Daar dienen twee kanttekeningen bij te worden gemaakt. In de eerste plaats was zelfs de korte versie van de vragenlijst nog relatief lang voor internetonderzoek. Met het behaalde responspercentage van bijna 25% lijkt het toch mogelijk vragenlijsten van substantiële lengte via internet af te nemen. Bovendien doet de lengte van de vragenlijst niet af aan de kwaliteit van de antwoorden en zijn respondenten na het completeren van de vragenlijst nog steeds gematigd positief en achten ze het waarschijnlijk vaker aan online onderzoek mee te zullen werken. Daarnaast blijkt de aankondiging (e-mail uitnodiging) van de lengte van de vragenlijst een belangrijke oorzaak te zijn dat mensen de vragenlijst niet invullen naast de ervaring van het invullen zelf. Een groot gedeelte van de respondenten die beginnen aan de vragenlijst maakt deze, ondanks de lengte, namelijk ook af. Een observatie die we nog niet veel hebben gezien in de bestaande literatuur. We concluderen dat het grootste gedeelte van de mensen a-priori, op basis van de e-mail uitnodiging, besluiten om niet aan de vragenlijst te beginnen.

De resultaten laten vervolgens zien dat de vragenlijst met grafische elementen een lager responspercentage vertoont dan de vragenlijst met alleen tekst. Dit is echter maar een klein effect, dat afwezig is bij respondenten die wel beginnen aan de vragenlijst maar deze niet afmaken. De aantrekkelijkheid van een vragenlijst met afbeeldingen blijkt geen motiverende factor te zijn, maar werkt tegelijkertijd ook niet demotiverend.

Belangrijker is het resultaat dat respondenten minder producten herkennen in de visuele vragenlijst dan in de tekstuele. Aanvankelijk veronderstelden we dat de visuele versie een hogere productherkenning zou opleveren, omdat de afbeeldingen een extra ondersteuning voor herkenning vormen. In de visuele variant hebben de respondenten twee aanwijzingen voor productherkenning. Gezien de lagere productherkenning moeten kennelijk zowel de naam als de afbeelding bekend voorkomen. Met andere woorden in de visuele conditie hanteren de respondenten een strenger criterium voor herkenning. Deze veronderstelling wordt gesteund door de bevindingen dat in de visuele vragenlijst minder vaak een 'weet ik niet' wordt gescoord. Wanneer respondenten eenmaal een product hebben herkend door middel van een afbeelding, lijken ze een actievere asso-

ciatie van relevante kenmerken te hebben. Dit kan van betekenis zijn voor onderzoek naar herkenning van producten en bijvoorbeeld reclame-uitingen. Door namelijk gebruik te maken van deze kennis kunnen afbeeldingen een strengere selectie-eis teweegbrengen en een hogere associatie. Natuurlijk moet hiermee rekening worden gehouden in de interpretatie, maar het kan tot een betere kwaliteit van antwoorden en onderzoek leiden.

Deze studie laat duidelijk zien dat respondenten de beloning duidelijk zien als iets dat aan hen zelf ten goede zou moeten komen. De respons bij de donatie bleef immers achter bij de andere beloningsvormen. In dit opzicht lijkt de loterij de meest geschikte incentive, aangezien een loterij kosteffectiever is dan een waardebon. Bovendien zagen we dat de responstijd voor loterij ruim anderhalve dag sneller is dan voor de waardebon. Wellicht dat mensen sneller reageren in een loterijconditie omdat ze denken dan meer kans te maken op een prijs. Dit resultaat mag opvallend worden genoemd, aangezien contant geld meestal als optimale stimulus wordt gezien en verondersteld wordt dat contant geld het hoogste responspercentage genereert. Hierbij moeten we ons natuurlijk wel afvragen in hoeverre een waardebon voor een online boeken- en platenwinkel als contant geld kan worden beschouwd.

Wat betreft de optimale timing voor het uitmailen van de reminder geeft onze studie geen eenduidig beeld. Er is een lichte voorkeur om de reminder na vijf werkdagen te sturen in plaats van na tien werkdagen. Niet alleen vanwege het iets hogere responspercentage voor de vroege reminder, maar ook als we ons realiseren dat online onderzoek zoveel sneller is dan schriftelijk onderzoek en dat onze resultaten laten zien dat een groot gedeelte van de antwoorden na drie dagen binnen is.

Concluderend kunnen we stellen dat de resultaten van ons onderzoek als een signaal mogen worden opgevat dat de grenzen van online onderzoek minder beperkt zijn, dan toe nu toe werd aangenomen. De bevinding dat deelnemers in ons onderzoek een sterke voorkeur hebben voor marktonderzoek via internet in vergelijking met telefonisch of schriftelijk onderzoek, versterkt dit alleen maar. Het lijkt mogelijk om relatief lange vragenlijsten op succesvolle wijze via internet af te nemen, ook als dat niet met behulp van een panel plaatsvindt. Bovendien moet kritisch worden omgegaan met afbeeldingen in productherkenning. Tot slot is creativiteit vereist bij de communicatie met respondenten, met name bij de e-mail uitnodiging. Daar ligt de uitdaging voor onderzoekers om creatief te zijn en respondenten te interesseren voor deelname. Doen respondenten eenmaal mee aan een onderzoek, dan vullen de meesten de vragenlijst volledig in. Met het experiment hebben we aangetoond dat ook voor relatief complex en grootschalig onderzoek het internet een alternatief kan zijn, mits met een aantal belangrijke factoren rekening wordt gehouden.

LITERATUUR

- Bauman, S., N. Jobity, J. Airey and H. Atak (2000), "Invites, Intros and Incentives: Lessons from a Web Survey", retrieved January 7th, 2002 from <http://www.wintellitech.com>.
- Bergmann, T.J., W.J. Hannaford and J. Wenner (1990), "Amount, Timing, and Value of Financial Incentives in Mail Surveys: Does it Make a Difference?", *Marketing Research: A Magazine of Management and Applications*, 2, pp. 30-36.
- Brown, M. (1965), "Use of a Postcard Query in Mail Surveys", *Public Opinion Quarterly*, Winter, pp. 635-637.
- Bruvold, N.T. and J.M. Comer (1988), "A model for Estimating the Response Rate to a Mailed

Survey", *Journal of Business Research*, 16 (2), pp. 101-116.

Couper, M.P., M.W. Traugott and M.J. Lamias (2001), "Web Survey Design and Administration", *Public Opinion Quarterly*, 65, pp. 230-253.

Dillman, D.A. (1991), "The Design and Administration of Mail Surveys", *Annual Review of Sociology*, 17, pp. 225-249.

Dillman, D.A., R.D. Tortora, J. Conradt and D. Bowker (1998), "Influence of Plain vs. Fancy Design on Response Rates for Web Surveys", retrieved January 11th, 2002 from <http://survey.sesrc.wsu.edu/dillman/papers.htm>.

Fox, R.J., M.R. Crask and J. Kim (1988), "Mail Survey Response Rate: A Meta-Analysis of Selected Techniques for Inducing Response", *Public Opinion Quarterly*, 52, pp. 467-491.

Frick, A., M.T. Bächtiger and U.D. Reips (1999), "Financial Incentives, Personal Information and Drop-out Rate in Online Studies", *Current Internet Science. Trends, Techniques, Results*, retrieved March 13th, 2002 from http://www.dgof.de/tband99/pdfs/a_h/frick.pdf.

Goetz, E.G., T.R. Tyler and F. Lomax Cook (1984), "Promised Incentives in Media Research: A look at Data Quality, Sample Representativeness, and Response Rate", *Journal of Marketing Research*, 21, pp. 148-154.

Heberlein, T.A and R.A. Baumgartner (1978), "Factors Affecting Response Rates to Mailed Questionnaires: A Quantitative Analysis of the Published Literature," *American Sociological Review*, 43 (August), pp. 447-462.

James, J.M. and R. Bolstein (1990), "The Effect of Monetary Incentives and Follow-up Mailings on the Response Rate and Response Quality in Mail Surveys", *Public Opinion, Quarterly*, 54, pp. 346-361.

James, J.M. and R. Bolstein (1992), "Large Monetary Incentives and Their Effect on Mail Survey Response Rate", *Public Opinion, Quarterly*, 56, pp. 442-453.

Linsky, A. (1975), "Stimulating Responses to Mailed Questionnaires: A Review", *Public Opinion Quarterly*, 39, pp. 82-101.

Lockhart, D.C. and F.R.J. Detwiler (1998), "Internet Surveys: a Description of the Demographics", *Quirk's Marketing Review*, retrieved October 1st, 2001 from http://www.quirks.com/articles/article_print.asp?arg_articleid=353.

MacElroy, B. (2000), "Measuring Response Rates in Online Surveys", *Quirk's Marketing Research Review*, retrieved January 7th, 2002 from http://www.quirks.com/articles/article_print.asp?articleid=583.

Marketfacts (2001), "Online Consumers: Beyond Fiction to Fact Distilling Reality from the Hype", retrieved March 13th, 2002 from <http://www.marketfacts.com/publications/pdf/r62.pdf>.

Mason, W., R. Dressel and R. Bain (1961), "An experimental study of factors affecting response to a mail survey of beginning teachers", *Public Opinion Quarterly*, 25, pp. 296-299.

Redline, C.D. and D.A. Dillman, (1999), "The Influence of Auxiliary, Symbolic, Numeric, and Verbal Languages on Navigational Compliance in Self-administered Questionnaires", retrieved June 18th, 2002 from <http://survey.sesrc.wsu.edu/dillman/papers/Auxiliary,Symbolic,Numeric%20paper—with%20Cleopdf>.

Rosenblum, J. (2001), "Give and Take: Use Online Surveys to Increase the Web's Communication Potential", *Quirk's Marketing Research Review*, retrieved June 14th, 2002 from http://www.quirks.com/articles/article_print.asp?arg_articleid=705.

Singer, E., J. van Hoewyk, and M.P. Maher (2000), "Experiments with Incentives in Telephone Surveys", *Public Opinion Quarterly*, 64, pp. 171-188.

Sheehan, K. (2001), "E-mail Survey Response Rates: A Review", *JCMC*, 6(2), retrieved October 8th, 2001 from <http://www.ascusc.org/jcmc/vol6/issue2/sheehan.html>.

Sheehan, K.B. and M.G. Hoy (1999), "Using E-mail to Survey Internet Users in the United States: Methodology and Assessment", *JCMC*, 4(3), retrieved January 11th, 2002 from <http://www.ascusc.org/jcmc/vol4/issue3/sheehan.html>.

Solomon, D.J. (2001), "Conducting Web-Based Surveys", *Practical Assessment, Research & Evaluation*, 7(19), retrieved January 9th, 2002, from <http://ericae.net/getvn.asp?v=7&n=19>.

Steele, T. J., W.L. Schwendig and J.A. Kilpatrick (1992), "Duplicate Responses to Multiple Survey Mailings: A Problem?", *Journal of Advertising Research*, 37 (March/April), pp. 26-34.

The Future Place Consultancy (2001), "A Guide to Best Practices in Online Quantitative Research", retrieved May 21st, 2002 from http://www.thefutureplace.com/best_practice.htm.

Warriner, K., J. Goyder, H. Gjertsen, P. Hohner and K. McSpurren (1996), "Charities, No; Lotteries, No; Cash, Yes: Main Effects and Interactions in a Canadian Incentives Experiment", *Public Opinion Quarterly*, 60, pp. 542-562.

Willems, P.M., P. Oosterveld, J. Lakerveld en T. Wildenberg (2002), "Modaliteits- en populatie-effecten onder de loep genomen", *Onderzoek*, 2, pp. 14-16.

Yammarino, F.J., S.J. Skinner and T.L. Childers (1991), "Understanding Mail Survey Response Behavior: A Meta-Analysis", *Public Opinion Quarterly*, 55, pp. 613-639.