

Tracking the eye of the beholder

De invloed van lay-out op biased response behavior in webvragenlijstonderzoek

Jelmer de Groot

1^e begeleider: Dr. J. W. M. Verhoeven

2^e begeleider: Drs. J. F. Gosselt

Communicatiewetenschap

Faculteit Gedragswetenschappen

Universiteit Twente

Voorwoord

Blijven gaan

Het onderste deel is het steilst. Dat is bekend. Hoe vaak je kapot gaat onderweg, ligt niet vast (ergens tussen de 6 en 12 keer). De mentale kwestie die Alpe d'Huez heet, is gewoon trappen op het lichtste tandje en dat 14,2 kilometer lang. Na bocht 3 zie je het dorp al liggen en ruik je de finishlijn. Nog maar 1,5 kilometer en je bent boven! Dan alleen nog door het dorp richting finish. In ieder geval niet zo steil meer. Het wonen in de Limburgse heuvels heeft mij aangemoedigd de fiets te pakken en ooit die verschrikkelijke puist in Zuid-Frankrijk te bedwingen. Het lukte.

Het afstuderen kan vergeleken worden met een dergelijke fietstocht. Verblind door mijn optimisme dacht ik het varkentje wel even te wassen. Dikwijls werd ik met beide benen op de grond gezet; toen ik niet in één week 60 proefpersonen wierf of toen de resultaten anders waren dan verwacht. Het motto bleek: blijven gaan. En het lukte.

Deze scriptie vormt het slotwerk van mijn studie. Een belangrijke fase in mijn leven die op het moment van afronden staat. Voor mijn afstudeerscriptie heb ik met veel plezier een half jaar begeleiding gehad vanuit het Centraal Bureau voor de Statistiek. Vanuit het CBS was er interesse naar de waarde van eye tracking als onderzoeksinstrument binnen webvragenlijsten. Met veel plezier kijk ik terug op de fijne periode die ik heb gehad en de vriendelijke en geïnteresseerde collega's.

Deze scriptie heeft mij veel inzicht verschaft in het fascinerende onderzoeksveld over webvragenlijsten. Met grote interesse heb ik veel gelezen en veel ervaring opgedaan binnen het vragenlijstontwerpveld.

Middels dit voorwoord wil ik de mensen bedanken die hebben meegeholpen dit product tot een succes te brengen. Het vragenlab van het CBS en in het bijzonder mijn dagelijkse begeleider Dirkjan Beukenhorst; voor je onuitputtelijke commentaar en onze overlegsessies met in ieder geval genoeg humor. Joost Verhoeven en Jordy Gosselt van de Universiteit Twente; voor jullie kritische blikken op mijn producten en het zeer vruchtbare overleg dat we vaak voerden. Zonder de proefpersonen zou dit onderzoek nooit voltooid worden; daarom is een dankwoord voor jullie op zijn plaats. Tot slot wil ik Sanne, mijn familie en mijn vrienden bedanken voor de steun tijdens mijn afstuderen en niet op de laatste plaats voor het fantastische studentenleven dat ik heb mogen meemaken!

Jelmer de Groot
Augustus 2010

Samenvatting

Het gebruik van matrixvragen in webvragenlijsten is niet ongebruikelijk. Matrixvragen worden gebruikt om pagina's te besparen, aangezien er veel vragen op een pagina passen. Respondenten worden in het geven van antwoorden op attitudevragenlijsten al snel gestuurd door de context van de vragen. Het groeperen van vragen tot één matrix, is een dergelijke sturende, contextuele factor. Door het groeperen van vragen ziet de respondent (onbewust) een verband tussen de diverse vragen. Door dit verband worden vragen niet meer als op zichzelf staand gezien, maar als bij elkaar horend. Respondenten zijn door dit optische verband eerder geneigd zichzelf de cognitieve last te besparen. De antwoorden die vervolgens worden gegeven, zijn minder gedifferentieerd en weerspiegelen niet altijd de werkelijke opinies. Over cognitieve verwerking van de vragenlijst is zeer weinig bekend. Dit onderzoek zal inzicht verschaffen in de cognitieve verwerking van matrixvragen en de ervaring ervan door de respondent.

Een experiment werd uitgevoerd volgens een 2x2 design. De helft van de respondenten kreeg een vragenlijst voorgelegd waarbij vragen van elkaar werden gescheiden door horizontale banen, terwijl de andere versie juist door middel van verticale banen in een matrix de antwoordopties van elkaar scheidde. Binnen deze versies werd er gevarieerd in betrokkenheid: respondenten vulden één vragenlijst in waarbij de betrokkenheid hoog was en één met lage betrokkenheid. Eye tracking werd gebruikt om na te gaan hoe respondenten de vragenlijsten verwerkten. Daarnaast rapporteerden de respondenten zelf de inspanning die ze leverden.

De eye tracker liet geen verschillen zien tussen de ontwerpen in overwegingen bij het geven van een antwoord. Deze verschillen waren wel aanwezig bij het lezen van de vragen. De duur van het kijken (fixatie) was bij het verticale ontwerp langer dan bij de respondenten die het horizontale ontwerp voorgelegd kregen. Daarnaast toonden de eye tracker data aan dat respondenten zorgvuldiger te werk gingen naarmate zij meer betrokken waren bij het ontwerp van de vragenlijst. Dit bleek ook uit de *straightlinescores* (het gemiddelde van de absolute verschillen tussen vraag x en vraag x-1 voor een matrix). Respondenten rapporteerden voor het horizontale ontwerp ook zelf een verschil in inspanning tussen de verschillende onderwerpen, maar respondenten die het verticale ontwerp voor zich kregen ervoeren dit verschil niet.

Wordt alleen gebruik gemaakt van de objectieve methode, eye tracking, dan is het gebruik van horizontale lijnen aan te raden. Respondenten hadden bij het verticale ontwerp meer fixaties en een langere duur van deze fixaties nodig voor het lezen van vragen. Als alleen naar de uitkomsten van de zelf-gerapporteerde inspanning gekeken zou worden, zou het verticale ontwerp beter gebruikt kunnen worden. Respondenten rapporteerden bij het verticale ontwerp namelijk zelf minder inspanning geleverd te hebben. Eye tracking zou een goed instrument kunnen zijn deze mate van cognitieve inspanning van de vragenlijst in kaart te brengen. Uit de data is echter te concluderen dat meer inspanning niet per se meer zorgvuldigheid hoeft te betekenen en dit maakt duidelijk dat data triangulatie in webvragenlijstonderzoek onmisbaar is.

Inhoudsopgave

1. Theoretische aanleiding.....	5
1.1 Attituderapportage	5
1.2 Matrixvragen	6
1.3 Satisficing.....	7
1.4 Cognitieve verklaring voor satisficing	9
1.5 Eye tracking	11
2. Methode.....	13
2.1 Design.....	13
2.2 Onafhankelijke variabelen.....	13
2.2.1 Ontwerp van de vragenlijst	13
2.2.2 Betrokkenheid ten aanzien van het onderwerp van de vragenlijst	14
2.3 Afhankelijke variabelen.....	15
2.3.1 Zelf-gerapporteerde inspanning	15
2.4 Mate van satisficing	15
2.4.1 Data kwaliteit	15
2.4.2 Waargenomen cognitieve verwerking	16
2.5 Onderzoeksmateriaal.....	17
2.6 Proefpersonen	17
2.7 Procedure	18
2.7.1 Het experiment	18
2.7.2 Kwaliteit fixaties.....	19
2.7.3 Debriefing	19
3. Resultaten	20
3.1 Manipulatiecheck	20
3.2 Kwaliteit van de antwoorden	20
3.3 Eye tracker data	23
3.4 Eye tracking als objectief onderzoeksinstrument	25
3.5 Zelf-gerapporteerde inspanning en betrokkenheid als subjectief instrument.....	27
4. Discussie	28
4.1 De invloed van lay-out op satisficing	28
4.2 Objectief gemeten zorgvuldigheid versus subjectieve inspanning	30
4.3 Eye tracking ter aanduiding van satisficing	31
4.4 Eye tracking als onderzoeksmethode	32
4.5 Toekomstig onderzoek.....	33
4.6 Tot slot	33
5. Referenties	34
Bijlage I: Vragenlijsten	37
Hoge betrokkenheid	37
Lage betrokkenheid	39
Bijlage II: Technische uitleg installatie eye tracking en databestanden.....	42
Bijlage III: semantische differentiaalschalen betrokkenheid en inspanning	43
Betrokkenheid	43
Inspanning	43

1. Theoretische aanleiding

“Geef aan in hoeverre u het eens bent met de volgende stellingen” is een veelvoorkomende tekst bij webvragenlijsten. Vervolgens verschijnt er op het scherm een matrix met een behoorlijk aantal vragen. De mate waarin de visuele presentatie van een vragenlijst de antwoorden op de vragen beïnvloedt, is zeer sterk (Couper, 2008). Dit hoofdstuk geeft een overzicht van de literatuur op het gebied van attitudematrices in vragenlijstontwerp. Vervolgens wordt een aantal onderzoeksvragen geformuleerd die in dit onderzoek beantwoord zullen worden.

1.1 Attituderapportage

Webvragenlijsten worden steeds vaker gebruikt (Dillman, 2007). Vergeleken met telefonische of *face-to-face* vragenlijsten wordt, door de afwezigheid van de interviewer, de betekenis van onduidelijke vragen dikwijls onbewust afgeleid uit de verbale en visuele elementen van de interface (Ware, 2000). In veel gevallen worden de door de respondenten gegeven antwoorden beïnvloed door de wijze waarop de vragen visueel worden gepresenteerd. Dit wordt ook wel aangeduid met de term contexteffect. Er dient bij het ontwerp van webvragenlijsten daarom veel rekening gehouden te worden met diverse ontwerpprincipes. Deze principes kunnen bepaalde invulgedragingen sturen of juist tegenhouden (Redline & Dillman, 1999). Deze invulgedragingen worden aangeduid als ‘de cognitieve aspecten van vragenlijstmethodologie’ (Schwarz, 1999).

De contexteffecten kunnen zorgen voor de meest voorkomende meetfout: biased response behavior (Tourangeau, Rips & Rasinski, 2000). Sudman, Bradburn en Schwarz (1996) bespreken de belangrijkste contexteffecten: (1) het aantal en de inhoud van voorafgaande vragen, (2) de context van één specifieke vraag, (3) de indeling van verschillende vragen op één pagina, (4) (het ontbreken van) introducties en (5) de grafische weergave van vragenlijsten. Veel ander onderzoek is bovendien verricht naar de invloed van woorden en lay-out op de antwoorden die worden gegeven (Couper, Traugott & Lamias, 2001; Lozar Manfreda, Batagelj, & Vehovar 2002; Smyth, Dillman, Christian & Stern, 2006). Eerdere onderzoeken (Toepoel, Das & van Soest, 2005 en Couper et al, 2001) naar het aantal vragen op één pagina toonden aan dat hier grote verschillen zijn in antwoordgedrag en daarom wordt in dit onderzoek aandacht besteed aan de zogenaamde matrixvragen en het specifieke invulgedrag dat hiermee gepaard gaat. Onderzoek dat tot op heden is uitgevoerd, hield zich vooral bezig met de effecten van de visuele presentatie op de kwaliteit van de data. De cognitieve processen die ten grondslag liggen aan het invulgedrag, bleven onbelicht. Dit onderzoek zal inzicht verschaffen in de achterliggende processen die een rol spelen bij het invullen van vragenlijsten en de manier waarop deze processen worden gestuurd door de visuele presentatie.

Deze contextafhankelijkheid is niet altijd even sterk. Respondenten hebben over sommige onderwerpen zeer sterke attitudes, die door de context niet snel te beïnvloeden zijn. De sterkste invloed die context zou hebben op attitudes, is als er ter plekke een attitude gevormd moet worden over een bepaald onderwerp. De attitude wordt gebaseerd op twee verschillende typen informatie: de altijd toegankelijke en de tijdelijk toegankelijke informatie. De altijd toegankelijke informatie reflecteert karaktereigenschappen van de respondent en onderwerpen waar de respondent veel van af weet. Dit in tegenstelling tot de tijdelijk

toegankelijke informatie, die afhankelijk is van de context van de vraag. De belangrijkste vorm van tijdelijk toegankelijke informatie bij het invullen van vragenlijsten, is het antwoord dat op de vorige vraag is gegeven. Dit antwoord zit nog vers in het geheugen, dus het is zeer waarschijnlijk dat deze informatie gebruikt wordt om tot een volgend antwoord te komen (Sudman, Bradburn & Schwarz, 1996). Het groeperen van vragen moedigt de respondent aan om de antwoorden op elkaar te baseren. Zeer weinig is bekend over de cognitieve processen die een rol spelen bij het invullen van matrixvragen en de invloed van lay-out hierop.

1.2 Matrixvragen

Het aanbieden van diverse vragen op één pagina gebeurt dikwijls in de vorm van een matrix (zie figuur 1). Het gebruik van matrixvragen, of *grid questions*, is niet ongebruikelijk bij het ontwerp van webvragenlijsten (Couper, 2008). In de literatuur is men het er niet over eens of het gebruik van matrixvragen ten goede komt aan de gebruiksvriendelijkheid van de vragenlijst. Poynter (2001) noemt dat *grids* zorgen voor hogere *drop out rates*. Wojtowicz (2001) noemt zelfs dat het gebruik van matrixvragen zo veel mogelijk vermeden moet worden.

Figuur 1: Voorbeeldsjabloon van een matrixvraag

	Optie 1	Optie 2	Optie 3	Optie 4
Vraag 1	O	O	O	O
Vraag 2	O	O	O	O
Vraag 3	O	O	O	O
Vraag 4	O	O	O	O
Vraag 5	O	O	O	O
Vraag 6	O	O	O	O

Het effect van het groeperen van vragen is zeer groot te noemen, zo blijkt uit de literatuur. Het groeperen van vragen beïnvloedt de antwoorden die respondenten geven. Vragen die gegroepeerd werden aangeboden, correleerden sterker dan wanneer vragen duidelijk van elkaar gescheiden werden (Schwarz & Hippler, 1995). Respondenten zijn in staat de invultijd te verkorten als gebruik wordt gemaakt van een matrix (Couper et al, 2001). Toepoel, Das en van Soes (2005) rapporteren dat de kwaliteit van de antwoorden hierbij wel achteruit gaat, omdat respondenten de vragen eerder afbreken en vaker *weet niet* invullen. In hetzelfde onderzoek wordt gerapporteerd dat meerdere vragen op één pagina hogere interne correlaties (Cronbach's α) opleveren in vergelijking met één vraag per pagina.

Deze hogere α -score roept echter de nodige vraagtekens op. Een hoge α zou betekenen dat het meetinstrument dat is gebruikt, zeer betrouwbaar is. Het is echter de vraag of men tevreden moet zijn met een hoge α op matrixvragen, omdat dit ook minder differentiatie tussen de antwoorden impliceert. De antwoorden worden eerder gebaseerd op elkaar dan dat ze als onafhankelijke vragen worden gezien. Het gevolg hiervan is dan ook dat met één enkele vraag per pagina Cronbach's α afneemt (zie o.a. Couper, 2008 en Toepoel et al, 2005), terwijl het meetinstrument alleen verschillend is qua indeling.

Een andere 'zorg' is dat respondenten bij het beantwoorden van matrixvragen minder differentiatie in de antwoorden laten zien dan wanneer vragen stuk voor stuk worden voorge-

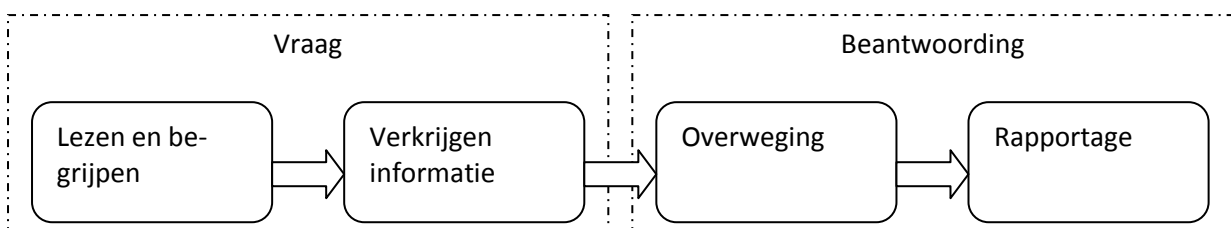
legd. Bovendien wordt een negatief geformuleerde vraag al snel over het hoofd gezien. Door in plaats van 'hoe gelukkig vindt u zichzelf' de vraag 'hoe ongelukkig vindt u zichzelf' te stellen, kost het de respondent meer cognitief vermogen de vraag te beantwoorden. Als een dergelijke negatieve vraagstelling niet wordt opgemerkt door de respondent en wordt beantwoord als een 'gewone' positieve vraagstelling, leidt dit tot minder valide data (Krosnick, 1991).

Samengevat: het gebruik van matrixvragenlijsten heeft zowel voor- als nadelen. Respondenten hebben kortere *responsetimes* als gebruik wordt gemaakt van matrices in webvragenlijsten (Couper et al, 2001). Daarbij is de interne consistentie hoger (Toepoel et al, 2005; Couper, 2008). Nadelen zijn echter dat matrixvragen zorgen voor meer drop-outs, en dat de antwoorden op de vragen op elkaar gebaseerd worden. De volgende stap is het evalueren van bestaande theorieën met betrekking tot dit invulgedrag. *Satisficing* van Jon Krosnick (1991) is een dergelijke theorie. In het volgende deel zal deze theorie worden besproken.

1.3 Satisficing

Het baseren van een antwoord op de vragen die ervoor werden gesteld, kan worden verklaard aan de hand van de *satisficing theorie* van Jon Krosnick (1991). Door het groeperen van vragen tot een matrix wordt *satisficing* in de hand gewerkt. *Satisficing* (een combinatie van de woorden *satisfy* en *suffice*) is de mate waarin respondenten zichzelf de cognitieve last besparen door vragen zo snel mogelijk in te vullen. Deze term is gebaseerd op het feit dat het zeer zorgvuldig invullen van de vragen in een vragenlijst veel cognitief werk vereist (Krosnick, 1991). Het gehele invulproces van één vraag in een vragenlijst bestaat uit vier stappen; het begrijpen van de vraag, het verkrijgen van alle relevante informatie uit het geheugen, het gebruiken van deze informatie om tot een oordeel te komen en tot slot het correct rapporteren van dit oordeel. Aan de hand van de door Krosnick (1991) genoemde cognitieve stappen kan het volgende fasemodel geconstrueerd worden:

Figuur 2: Het antwoordproces van Krosnick (1991) in beeld gebracht



Tijdens de eerste fase, het lezen en begrijpen van de vraag, wordt de informatie door de respondent gelezen en geïnterpreteerd. Gedurende de fase die daarop volgt worden de herenen geraadpleegd om relevante informatie te vinden. Deze gevonden informatie wordt tijdens de derde fase gebruikt om tot de juiste attitude te komen ten aanzien van de stelling. Tot slot bevat de laatste fase de rapportage; de respondent ziet de antwoorden en overweegt welk antwoord het best de uit fase drie voortgekomen attitude reflecteert.

Krosnick (1991) noemt drie modererende variabelen die een rol spelen in de mate van zorgvuldigheid waarmee het model wordt doorlopen. Dit zijn:

1. De vaardigheden van de respondent; respondenten die minder in staat zijn om de cognitieve stappen uit te voeren, zijn eerder geneigd om *short-cuts* te nemen en zichzelf de cognitieve last te besparen.
2. De motivatie van de respondent; minder gemotiveerde respondenten zijn eerder geneigd *satisficing* gedrag te vertonen dan gemotiveerde respondenten. Dit wordt ondersteund door Ajzen, Brown & Rosenthal (1996) en Jemmott, Ditto & Croyle (1986), die allen aantoonde dat naarmate respondenten minder betrokken waren, motivatie lager bleek te zijn en dit van invloed was op de antwoorden die werden gegeven.
3. De taak zelf; moeilijkere vragen zijn cognitief vermoeiender voor de respondenten en dit zou eerder leiden tot *satisficing* gedrag. Dit *satisficing* gedrag uit zich in het minder zorgvuldig doorlopen van de vier fasen.

Krosnick (1991) onderscheidt twee typen invullers van vragenlijsten: *optimizers* en *satisficers*. *Optimizers* proberen zo accuraat mogelijk antwoord te geven op elke vraag. Hierbij zijn ze bereid de inspanning te leveren die nodig is om zo zorgvuldig mogelijk de vier vereiste stappen te doorlopen. *Satisficers* vinden het voornamelijk belangrijk om de cognitieve last zo laag mogelijk te houden. Bij het doorlopen van de vier cognitieve stappen wordt de vraag zo snel mogelijk gelezen, het geheugen minder goed doorzocht, de opgeroepen informatie al snel als 'waar' beschouwd en een antwoordoptie wordt met weinig of geen overweging gekozen en antwoorden zijn hierdoor sneller gebaseerd op de instemming die de respondent eerder rapporteerde.

Satisficers hebben de neiging om zichzelf de cognitieve last te besparen. Dit gebeurt door de antwoorden op vragen niet te overwegen, maar door zo snel mogelijk tot een antwoord te komen. Krosnick (1991) veronderstelt dat veel vragenlijsten standaard geformuleerde stellingen bevatten waarbij het vaak voor zich spreekt dat men het eens is met de stelling. De *satisficer* laat zich graag sturen door deze vraagstelling. Dit mondt uit in positievere attitudes (*acquiescence bias*). Andere kenmerken van respondenten die zich de cognitieve last willen besparen zijn: het willekeurig kiezen van een antwoord, het kiezen van een neutraal antwoord, het invullen van *weet niet* en antwoorden die op elkaar gebaseerd zijn in plaats van vragen op zichzelf te beantwoorden (Krosnick, 1991). Door de vragen en antwoorden met elkaar in verband te zien in plaats van vragen op zichzelf te beoordelen, vindt *nondifferentiation* plaats en *straightlining*. *Nondifferentiation* houdt in dat de gerapporteerde attitudes weinig differentiatie vertonen en dus weinig van elkaar verschillen. *Straightlining* lijkt hier veel op; *straightliners* vertonen bij het geven van antwoorden een duidelijke verticale lijn; alle vragen worden met dezelfde mate van instemming beantwoord, ongeacht de inhoud van de vragen. Vanuit de *preference for consistency theory* van Cialdini kost het de respondent meer cognitieve inspanning na een aantal positieve attitudes te wisselen naar een negatieve (Cialdini, 1993).

Doordat *satisficing* zich kan uiten in een groot aantal zaken, vormt het meten van *satisficing* een zeer gecompliceerde taak. Onderzoekers en ontwerpers weten dat *satisficing* voorkomt, maar een eenduidige meetmethode ontbreekt. Dit wordt ten eerste veroorzaakt doordat bij attitudes een *true value* ontbreekt; er is altijd sprake van meetfouten. Ten tweede speelt het feit dat de daadwerkelijke cognitieve inzichten in invulgedrag ontbreken een rol ter aanduiding van *satisficing*. Ten derde kan *satisficing* zich in veel verschillende zaken uiten en er

is dus lastig één maat aan te geven. Tot slot hoeft niet elk symptoom van *satisficing* te duiden op *satisficing*. Een respondent kan namelijk *weet niet* invullen omdat hij het antwoord echt niet weet. Daarnaast zou het kunnen voorkomen dat een respondent drie keer hetzelfde antwoord geeft, omdat dit zo precies mogelijk zijn attitude weergeeft.

1.4 Cognitieve verklaring voor *satisficing*

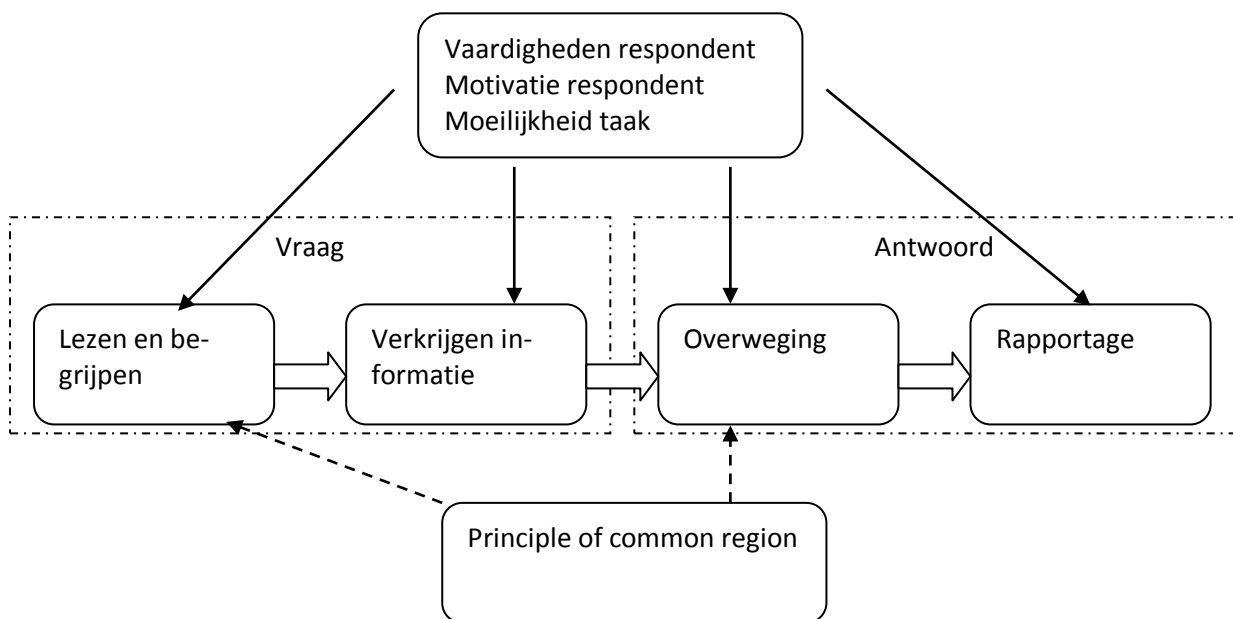
Er is weinig inzicht in de daadwerkelijke verwerking van de vragen en de antwoorden in de hersenen. De precieze invloed van de vaardigheid en motivatie van de respondent en de moeilijkheid van de taak op het invulgedrag is onbekend. Het is tot nu toe onduidelijk hoe een variatie hierin zich kan vertalen naar het fasemodel van Krosnick (1991). Tourangeau, Couper & Conrad (2004) maakten wel duidelijk dat respondenten verschillende visuele heuristieken gebruiken om een vraag te kunnen interpreteren. Een voorbeeld van een heuristiek wordt gevormd door de *Gestaltprincipes*. Deze principes uit de cognitieve psychologie verklaren visuele perceptie. Uitgangspunt is dat de waarneming van objecten niet wordt verklaard door de optelsom van alle delen, maar door de totale configuratie of de organisatie ervan. De hersenen groeperen de verschillende objecten op basis van deze *Gestaltprincipes*. De voornaamste drie principes die invloed hebben op het vraag-antwoord proces zijn: *principle of proximity* (dat wat bij elkaar staat, hoort bij elkaar), *principle of similarity* (dat wat op elkaar lijkt, hoort bij elkaar) en het *principle of common region* (dat wat omsloten wordt door dezelfde lijnen, hoort bij elkaar). Aan de hand van deze principes kan worden verklaard dat de vragen binnen één matrix als samenhangend worden gezien en dit heeft duidelijke invloed op de antwoorden die worden gegeven (Dillman, 2007).

Op matrixniveau is er niet te variëren in het *similarity* principe, dat uitgaat van hetgeen dat op elkaar lijkt, bij elkaar hoort. Hier is alleen in te variëren door elke vraag in een aparte context te presenteren en daardoor staan de vragen niet meer in een matrix. Het *proximity* principe vindt alleen plaats bij matrixvragen, omdat daarbij vragen bij elkaar staan en dus bij elkaar horen. In deze twee principes is dus niet te variëren binnen een matrixvraag. Wel kan er op matrixniveau verschil aangebracht worden in de manier waarop het *principle of common region* wordt toegepast door de respondenten. Horizontale en verticale lijnen zorgen namelijk voor *grouping* in website design. Als vragen gescheiden worden door middel van horizontale of verticale lijnen, ziet de proefpersoon de objecten binnen de lijnen als samenhangend (Palmer, 1992).

Door de vragen van elkaar te scheiden door middel van horizontale lijnen ontstaat er een horizontale 'region' waarbij de vragen van elkaar gescheiden worden en er een verband bestaat tussen de vraag en de antwoordopties. Het gevolg hiervan is dat de vragen minder in verband met elkaar zouden worden gebracht en zo vindt *satisficing* in mindere mate plaats. Anderzijds ontstaat er een verticale 'region' door juist de antwoordopties van elkaar te scheiden (Palmer, 1992). Zo ontstaat er een verticaal verband tussen de antwoorden die op de verschillende vragen zijn gegeven. Hierdoor zou een sterke samenhang tussen de verschillende antwoorden geïmpliceerd worden en dit versterkt (het eventueel aanwezige) *satisficing* gedrag in zekere mate. Het verband dat er al wordt geconstateerd tussen de verschillende vragen door ze in één matrix te plaatsen, zou door deze verticale lijnen versterkt worden. De lijn vormt een cognitieve barrière om andere opties te kiezen. Het houdt de respondent tegen om over de lijn heen te durven. De aanname is daarom dat antwoordopties die gescheiden zijn door middel van verticale lijnen ervoor zorgen dat *satisficing* wordt aange-

moedigd. Deze satisfice gedragingen zullen in het bijzonder plaatsvinden tijdens fase 1 en fase 3 van het antwoordproces. Respondenten lezen de vragen minder zorgvuldig en het overwegen van de antwoordopties gebeurt sneller. Antwoorden worden namelijk niet meer overwogen op basis van de verkregen informatie uit het geheugen, maar op basis van de antwoorden die men eerder had gegeven.

Figuur 3: Fasemodel Krosnick (1991) met principle of common region toegevoegd



Veel respondenten die vrijwillig meedoen aan een vragenlijstonderzoek zijn bereid om het nodige cognitieve werk te verrichten om de vragenlijst zorgvuldig in te vullen. Maar veel andere respondenten die aangeven een vragenlijst in te willen vullen, worden naarmate de tijd verstrijkt minder gemotiveerd en vermoeid. Het gevolg hiervan is dat de cognitieve stappen minder zorgvuldig worden doorlopen en zorgen voor antwoorden waarbij de attitudes minder nauwkeurig gereflecteerd worden. Zo kan de vraag sneller gelezen worden of minder goed geïnterpreteerd, is het moeilijker voor de respondent om informatie te verkrijgen, gebeurt het overwegen van het antwoord sneller en minder zorgvuldig en wordt er naar aanleiding van deze korte overweging meteen een attitude gerapporteerd.

Eén van de modererende variabelen die Krosnick (1991) noemt, is betrokkenheid. In dit onderzoek zal nagegaan worden in hoeverre de betrokkenheid van de respondent een rol speelt bij de effecten van lay-out op satisfice gedrag. Te verwachten is dat door het verschil in motivatie, respondenten eerder geneigd zijn zich te laten leiden door de lay-out en dus meer kenmerken van *satisficing* vertonen. Omdat *satisficing* moeilijk in één maat is te meten, worden de hypothesen opgesplitst naar drie vormen van *satisficing*, te weten *straightlining*, *weet niet scores* en *acquiescence* van de gerapporteerde attitudes

H1: Respondenten die het verticale ontwerp voorgelegd krijgen laten lagere *straightlinescores* zien dan respondenten die het horizontale ontwerp voorgelegd krijgen.

H2: De waargenomen verschillen in *straightlinescores* tussen de ontwerpen worden versterkt door een lage betrokkenheid.

H3: Respondenten die het verticale ontwerp voorgelegd krijgen laten een hoger aandeel *weet niet* zien dan de respondenten die het horizontale ontwerp voorgelegd krijgen.

H4: De waargenomen verschillen in *weet niet* scores tussen de ontwerpen worden versterkt door een lage betrokkenheid.

H5: Respondenten rapporteren bij het verticale ontwerp positievere attitudes dan respondenten die het horizontale ontwerp voorgelegd krijgen.

1.5 Eye tracking

Een goed instrument om na te gaan hoe respondenten bezig zijn met de cognitieve verwerking van vragen, is eye tracking. Eye tracking is binnen het human-computer-interactionveld een betrekkelijk nieuwe onderzoeksmethode. Bij eye tracking worden, zoals het woord al doet vermoeden, de ogen gevolgd tijdens het presenteren van één enkele stimulus of meerdere stimuli op het beeldscherm van een computer. Eye tracker data bestaan uit saccades en fixaties.

Saccades zijn de oogbewegingen van het ene naar het andere punt. Tijdens de saccade wordt er geen informatie verwerkt door de hersenen (zie o.a. Jacob & Karn, 2003 en Cooke, 2006). Aan saccades werd in dit onderzoek geen aandacht besteed.

Fixaties duiden op de verwerking van informatie in de hersenen. Tijdens fixaties staan de ogen even stil op een bepaald punt om informatie te verwerken. Een groter aantal fixaties op een bepaald object duidt erop dat de respondent het desbetreffende object zorgvuldig wil verwerken of juist dat het object meer verwerking van de proefpersoon vereist om verwerkt te worden. De duur van deze fixaties vertelt iets over de zorgvuldigheid die respondenten hadden met het verwerken van de informatie. Hoe zorgvuldiger een proefpersoon te werk gaat, des te meer cognitieve inspanning de proefpersoon besteedt. Daarnaast zou er vanuit de respondent behoefte kunnen zijn zorgvuldiger een bepaald deel van de informatie te verwerken (Poole & Ball, 2009). De gevolgde oogbewegingen kunnen geanalyseerd worden en op deze manier kunnen duur en lengte van fixaties in kaart gebracht worden. Daarnaast legt eye tracking zogenaamde 'scan paths' bloot (Cooke, 2004). Scan paths tonen de zoekstrategieën van respondenten. Voorheen werd eye tracking gebruikt om inzicht te verschaffen in leestechieken en zoekstrategieën binnen het menselijk brein. De laatste jaren wordt eye tracking voornamelijk gebruikt om gebruiksvriendelijkheid te testen van websites of software. Bij deze tests gaat het er dan vooral om waar de proefpersonen naar kijken. Er wordt onderzocht waar respondenten als eerste naar kijken en welke zoekstrategieën respondenten hanteren bij het zoeken naar bepaalde informatie. Eye tracking wordt dan voornamelijk gebruikt als exploratieve methode, om te kijken hoe iemand reageert zodra er een webpagina wordt getoond.

Usability testing met behulp van eye tracking gaat uit van de eye-mind hypothesis (Just & Carpenter, 1976), namelijk dat de informatie die wordt bekeken, verwerkt wordt in de hersenen. Op deze aanname valt wel af te dingen dat het onduidelijk is wat er precies gebeurt met de bekeken informatie. De data laten wel zien waar iemand naar kijkt, maar dan hoeft het niet zo te zijn dat de bekeken informatie in zijn geheel wordt waargenomen en dat naar aanleiding van de bekeken informatie actie ondernomen wordt (Jacob & Karn, 2003). Daar-

naast wordt wel gemeten waarnaar men kijkt, maar waarom is vaak onduidelijk (Cooke, 2006). Op deze manier is het verschil tussen kijken en zien moeilijk waar te nemen. Sommige informatie kan wel bekeken worden, maar of het ook daadwerkelijk gezien wordt, is een heel andere vraag. Poole & Ball (2009) adviseren om deze reden alleen fixaties te analyseren die langer duren dan 100 milliseconden, om er zeker van te zijn dat de bekeken informatie in de hersenen verwerkt wordt.

In dit onderzoek wordt eye tracking gebruikt om de zorgvuldigheid waarmee het fasemodel van Krosnick (1991) wordt doorlopen te onderzoeken. Het aantal fixaties en de duur van deze fixaties is gebruikt om de cognitieve last die respondenten ondervonden te onderzoeken. Op basis hiervan worden de volgende vier hypothesen geformuleerd:

H6: Respondenten die het verticale ontwerp voorgelegd krijgen vertonen minder fixaties dan respondenten die het horizontale ontwerp voorgelegd krijgen.

H7: Respondenten die het verticale ontwerp voorgelegd krijgen vertonen een kortere duur van de fixaties dan respondenten die het horizontale ontwerp voorgelegd krijgen.

H8: De waargenomen verschillen in aantal fixaties worden versterkt door een lage betrokkenheid.

H9: De waargenomen verschillen in de duur van de fixaties worden versterkt door een lage betrokkenheid.

Waar veel onderzoeksmethoden die vragenlijsten evalueren van subjectieve aard zijn, is eye tracking juist van objectieve aard. De rapportage in subjectieve onderzoeksmethoden is gebaseerd op de belangen van de respondent zelf, wat de respondent belangrijk vindt. De zaken die de respondent noemt zijn contextafhankelijk en ook hierbij spelen processen als *satisficing* en sociale wenselijkheid een rol. Dit in tegenstelling tot objectieve onderzoeksmethoden, waarbij zaken waargenomen worden die wellicht niet altijd door de respondent worden genoemd (bewust of onbewust). Dit onderzoek zal het gebruik eye tracking evalueren ten aanzien van subjectieve onderzoeksmethoden die *satisficing* aanduiden. Hypothese 10 luidt daarom:

H10: Er is geen verschil tussen de door de eye tracker waargenomen cognitieve inspanning in vergelijking tot de door de respondenten gerapporteerde cognitieve inspanning.

2. Methode

Eye tracking werd in dit onderzoek gebruikt om inzicht te krijgen in de cognitieve verwerking die de respondenten hadden bij het invullen van de vragenlijst. Eerst zal het design van het experiment besproken worden. Vervolgens zullen de onafhankelijke en afhankelijke variabelen gepresenteerd worden en tot slot volgt er uitleg over het aantal proefpersonen, het onderzoeksmateriaal en de procedure tijdens het experiment.

2.1 Design

Het experiment bestond uit een 2 (horizontale banen versus verticale banen) x 2 (hoge betrokkenheid versus lage betrokkenheid) design. Aan elke proefpersoon werden twee vragenlijsten voorgelegd. De helft van de proefpersonen kreeg de vragenlijsten met horizontale lijnen voorgelegd, bij de andere helft bestonden de vragenlijsten uit verticale lijnen. Per ontwerp (horizontaal versus verticaal) werden de twee verschillende onderwerpen counterbalanced voorgelegd aan de proefpersonen. De respondenten verschilden in betrokkenheid ten aanzien van de twee onderwerpen. Op deze manier werd gecontroleerd voor de factor motivatie in het Krosnick (1991) model. Het ene onderwerp ging over de afschaffing van de studiefinanciering, een onderwerp dat voor studenten zeer relevant was en dus voor hogere betrokkenheid zorgde. Het andere onderwerp betrof lage betrokkenheid, namelijk overwegingen die voor een persoon belangrijk kunnen zijn bij het kiezen voor een pensioenfonds.

2.2 Onafhankelijke variabelen

2.2.1 Ontwerp van de vragenlijst

Om antwoord te krijgen op de vraag of lay-out een rol speelt bij het invullen van een vragenlijst en de cognitieve verwerking ervan, werd gebruik gemaakt van twee verschillende designs. Deze twee designs varieerden van elkaar in de manier waarop de vragen van elkaar werden gescheiden. In de ene versie werd gebruik gemaakt van verticale banen om de antwoordopties van elkaar te scheiden. De andere versie maakte juist gebruik van horizontale banen om de vragen van elkaar te scheiden. Ter illustratie de screenshots van de verschillende ontwerpen.

Figuur 4: Verticaal ontwerp

Bij het kiezen van een verzekeringssmaatschappij vind ik het belangrijk dat...

Geef aan in hoeverre je het oneens/eens bent met de volgende stellingen:

	Zeer oneens	Mee oneens	Neutraal	Mee eens	Zeer eens	Weet niet
De transacties bij een pensioenfonds moeten veilig zijn.	☐	☐	☐	☐	☐	☐
De premie bij een pensioenfonds moet laag zijn.	☐	☐	☐	☐	☐	☐
Een pensioenfonds moet een breed assortiment hebben.	☐	☐	☐	☐	☐	☐
Een pensioenfonds moet binnen een bepaalde productcategorie veel keuze hebben.	☐	☐	☐	☐	☐	☐
Ik vind het belangrijk dat ik het uitkeringsbedrag bij mijn pensioenfonds zelf kan bepalen.	☐	☐	☐	☐	☐	☐
De prijs/kwaliteit verhouding van een pensioenfonds moet slecht zijn.	☐	☐	☐	☐	☐	☐
Een pensioenfonds hoeft niet goed bereikbaar te zijn.	☐	☐	☐	☐	☐	☐
De locatie van de vestigingen van het pensioenfonds moeten dichtbij zijn.	☐	☐	☐	☐	☐	☐
Ik verwacht dat een pensioenfonds veel gebruik maakt van e-mail.	☐	☐	☐	☐	☐	☐

[Verder...](#)

Figuur 5: Horizontaal ontwerp

Bij het kiezen van een verzekeringssmaatschappij vind ik het belangrijk dat...

Geef aan in hoeverre je het oneens/eens bent met de volgende stellingen:

	Zeer mee oneens	Mee oneens	Neutraal	Mee eens	Zeer mee eens	Weet niet
De transacties bij een pensioenfonds moeten veilig zijn.	☐	☐	☐	☐	☐	☐
De premie bij een pensioenfonds moet laag zijn.	☐	☐	☐	☐	☐	☐
Een pensioenfonds moet een breed assortiment hebben.	☐	☐	☐	☐	☐	☐
Een pensioenfonds moet binnen een bepaalde productcategorie veel keuze hebben.	☐	☐	☐	☐	☐	☐
Ik vind het belangrijk dat ik het uitkeringsbedrag bij mijn pensioenfonds zelf kan bepalen.	☐	☐	☐	☐	☐	☐
De prijs/kwaliteit verhouding van een pensioenfonds moet slecht zijn.	☐	☐	☐	☐	☐	☐
Een pensioenfonds hoeft niet goed bereikbaar te zijn.	☐	☐	☐	☐	☐	☐
De locatie van de vestigingen van het pensioenfonds moeten dichtbij zijn.	☐	☐	☐	☐	☐	☐
Ik verwacht dat een pensioenfonds veel gebruik maakt van e-mail.	☐	☐	☐	☐	☐	☐

[Verder...](#)

2.2.2 Betrokkenheid ten aanzien van het onderwerp van de vragenlijst

Per lay-out versie werd een tweetal vragenlijsten opgesteld. De ene vragenlijst ging over studiefinanciering, een onderwerp waar studenten zeer betrokken bij zijn. Dit in tegenstelling tot het andere onderwerp, namelijk pensioenfonds. Dit staat wat verder van de studenten af en zorgt voor lagere betrokkenheid. Om tijdens dit onderzoek te kunnen controleren of de persoonlijke betrokkenheid bij de ene vragenlijst groter was dan bij de andere, werd bij beide vragenlijsten de persoonlijke betrokkenheidsschaal van Mittal (1995) voorgelegd aan de proefpersonen. Dit is een 5 items tellende semantische differentiaal schaal. Voorbeelden van de uitersten van deze schaal waren: belangrijk-onbelangrijk en boeit mij-boeit mij niet. Cronbach's α voor deze schaal was voor de hoge betrokkenheidsvragenlijst 0,87 en voor de lage betrokkenheidsvragenlijst 0,71.

2.3 Afhankelijke variabelen

2.3.1 Zelf-gerapporteerde inspanning

Na elke vragenlijst werd de respondent een vragenlijst voorgelegd die het zelf-gerapporteerde *satisficing* gedrag in kaart brengt. Omdat invulgedrag nauw samenhangt met de mentale inspanning die wordt geleverd, werd een vragenlijst voorgelegd waarbij 'mental effort' wordt gemeten. Deze korte vragenlijst was een combinatie van de inspanningsvragenlijst die Krosnick et al (2002) eerder gebruikten en een aantal items uit de NASA-TLX (Hart & Steveland, 1988), waarbij mentale inspanning gemeten wordt.

De vragenlijst van Krosnick et al (2002) bestaat uit een mental effort schaal om cognitieve inspanning te meten bij het invullen van een vragenlijst. De volgende drie vragen met 7-puntsschaal kwamen aan bod om mentale inspanning te meten:

1. *Hoe zorgvuldig hebt u deze vragenlijst ingevuld?*
2. *Hoe grondig hebt u nagedacht alvorens een antwoord te geven?*
3. *Hoeveel inspanning hebt u geleverd bij het invullen van de vragenlijst?*

Daarnaast werd een deel van de gevalideerde vragenlijst van NASA gebruikt, de NASA-TLX (Hart & Staveland, 1988). Deze vragenlijst bestaat uit 6 vragen, die elk meten in hoeverre een proefpersoon inspanning leverde bij het volbrengen van een taak. Aangezien deze vragenlijst van toepassing is op de sector vliegtuigbouw, zijn de vragen gereduceerd tot drie. Omdat één vraag overeenkwam met de schaal die Krosnick et al (2002) gebruikten, bleven er van de NASA-TLX schaal twee vragen over. Dit zijn:

1. *Hoeveel mentale inspanning kostte het invullen van de vragenlijst?*
2. *Hoe verveeld was u bij het invullen van deze vragenlijst?*

De respondenten dienden hierbij te antwoorden door de best passende mate van instemming op een zevenpunts Likertschaal aan te geven.

Door het samenvoegen van deze twee vragen met de vragenlijst van Krosnick et al (2002) ontstond een construct om zelfgerapporteerde inspanning na te gaan (zie bijlage IV). Voor het persoonlijk betrokken onderwerp was Cronbach's $\alpha=0,75$ in eerste instantie. Cronbach's α voor het minder betrokken onderwerp bedroeg in eerste instantie 0,60. Na het verwijderen van het vijfde item bedroegen de α -scores respectievelijk 0,81 en 0,65. Het item dat werd verwijderd was 'hoe verveeld was u tijdens het invullen van de vragenlijst?'

2.4 Mate van *satisficing*

2.4.1 Data kwaliteit

Om na te gaan in hoeverre *satisficing* plaatsvond, wordt aan de hand van de ingevulde antwoorden op de vragenlijst het volgende berekend:

1. Het aandeel *weet niet*; het aantal keren dat een proefpersoon *weet niet* invulde, wordt geteld en gedeeld door het aantal vragen. Hieruit komt een percentage. Hoe hoger dit percentage, hoe sterker de persoon bezig is met het toepassen van *satisficing*, omdat de respondent zichzelf de cognitieve last bespaart door geen echte attitude aan te kruisen.
2. *Straightlinescore* van de gerapporteerde attitudes; de *straightlinescore* geeft aan in hoeverre het antwoord op vraag x afwijkt van het antwoord op vraag x-

1. De *straightlinescore* is gebaseerd op de mate waarin het antwoord op een bepaalde vraag afwijkt van het antwoord dat men op de vraag ervoor gaf. Deze absolute score wordt per vraag berekend en zo kan er een gemiddelde score per matrix worden opgesteld. Proefpersonen die door de verticale banen meer geleid worden om hetzelfde antwoord te geven, zouden een lagere *straightlinescore* laten zien dan de respondenten die de versie kregen met de horizontale banen, omdat de afwijking tussen de antwoorden kleiner is.

Een aantal maten waarbij geprobeerd werd *satisficing* aan te tonen, is de Cronbach's α (Couper, 2008 en Toepoel et al, 2005), variantie (Chang & Krosnick, 2009) of het tellen van het aantal identieke antwoorden in een matrix (Holbrook et al, 2003; Genassali, 2008). Zowel variantie als interne correlatie gaan na in welke mate de antwoorden binnen één matrix overeenkomen. De variantie wordt gemeten aan de hand van de gemiddelde afstand van elk antwoord tot de gemiddelde attitude. Als een respondent 5 keer achter elkaar een 1 antwoordt en daarna 5 keer een 5, heeft deze respondent dezelfde variantie als een respondent die ombeurten een 1 en een 5 rapporteert, terwijl hier duidelijk een verschil is waar te nemen in differentiatie tussen twee verschillende antwoorden. Het nadeel bij zowel de variantie als Cronbach's α , is dat de volgorde in de matrix niet in acht wordt genomen. De mate waarin de attitude gebaseerd is op de attitude die ervoor werd gerapporteerd, is niet aan te duiden aan de hand van α - en variantiescores. Om deze reden werd er in dit onderzoek gekozen voor een maat waarbij nagegaan werd in hoeverre het antwoord op vraag x gebaseerd was op vraag x-1. Dit gebeurde door de som van de absolute verschillen tussen vraag x en vraag x-1 per matrix te delen door het aantal vragen.

2.4.2 Waargenomen cognitieve verwerking

De eye tracker werd aan de hand van verschillende lookzones ingesteld. Dit werd zo gedaan, dat zones werden gemaakt voor het vraagdeel van elke vraag en voor het antwoorddeel van elke vraag. De zone voor het vraagdeel bevatte alleen de tekst waarin de stelling werd voorgelegd aan de proefpersoon. Het antwoorddeel van de vraag bevatte de antwoordopties.

Figuur 5: Indeling vraag naar lookzones

Een pensioenfonds moet een aantrekkelijk bedrijfspand hebben.	☐	☐	☐	☐	☐	☐
Ik vind het belangrijk dat een pensioenfonds modern uitziende faciliteiten heeft.	☐	☐	☐	☐	☐	☐

De ruimte binnen het gemarkeerde gebied wordt gedefinieerd als een 'look zone'. Zo kan eenvoudig bekeken worden welke fixaties en hoeveel fixaties door de proefpersoon werden besteed aan het lezen van de vraag. De duur van deze fixaties zal tevens aangegeven worden. Hetzelfde geldt voor het antwoorddeel van elke stelling.

Satisficing werd in kaart gebracht aan de hand van de hoeveelheid en de duur van de fixaties die plaatsvonden. Het aantal fixaties en de duur van de fixaties per vraagdeel geeft aan in hoeverre de respondenten zorgvuldig te werk zijn gegaan bij het lezen van de vraag. Hoe groter het aantal fixaties en hoe langer de duur ervan, des te zorgvuldiger de vraag wordt gelezen (Poole & Ball, 2009). Bij deze fixaties werden alleen fixaties meegenomen met een

duur langer dan 100 milliseconden, om er zeker van te zijn dat de informatie waar fixaties plaatsvinden, daadwerkelijk wordt verwerkt in de hersenen.

Dit gaat ook op voor de hoeveelheid en de duur van de fixaties die zich bevinden in het antwoorddeel van de vraag; meer fixaties zou een langere overweging betekenen en dus sterkere mate van *optimizing* (en mindere mate van *satisficing*). Andersom zal een kleiner aantal fixaties en een gemiddeld kortere duur minder overweging betekenen (Poole & Ball, 2009). Hieruit zou te concluderen zijn dat de proefpersonen zichzelf de cognitieve lasten willen besparen en dus een sterkere mate van *satisficing* laten zien.

2.5 Onderzoeksmateriaal

Het onderzoeksmateriaal bestond uit twee webvragenlijsten. De ene vragenlijst bevatte het persoonlijk betrokken onderwerp, studiefinanciering. De andere ging over de overwegingen die een consument maakt bij het kiezen van een pensioenfonds. Dit onderwerp was voor de respondenten minder betrokken. De vragenlijst over pensioenfondsen bestond uit 68 vragen, verdeeld over zeven pagina's met per pagina een matrix. De vragenlijst over de studiefinanciering bestond uit 56 vragen, verdeeld over zes pagina's en ook per pagina een vragenmatrix. Hierdoor had de webvragenlijst met lage betrokkenheid zeven matrices en die met hoge betrokkenheid zes. Per onderwerp werden in de eerste matrix geen negatieve vragen geformuleerd, dit begon bij matrix 2. Hiervan werd gebruik gemaakt omdat respondenten die de vragenlijst snel willen invullen, minder aandacht zouden besteden aan de inhoud van de stellingen. Op negatief geformuleerde stellingen zou dus ook een positief antwoord volgen. Dit werkt de *straightlinescores* aanzienlijk in de hand, omdat de afwijking van vraag x ten opzichte van vraag $x-1$ kleiner is.

Elke matrix bestond naast de stellingen uit een set met zes antwoordopties. Een Likert-schaal met vijf antwoordopties die de attitudeschaal besloegen en een optie *weet niet*. De vijf antwoordopties reikten van 'geheel mee oneens' tot 'geheel mee eens'. Omdat *satisficing* zich ook uit in de mate waarin de respondent *weet niet* invult, werd er een optie *weet niet* in de vragenlijst opgenomen om *satisficing* gedrag goed te kunnen aantonen. De vragen werden per matrix ingedeeld naar bepaalde thema's. In de bijlage staan de vragenlijsten en staan de verschillende thema's bij de vragen genoemd.

2.6 Proefpersonen

Aan het onderzoek deden 40 proefpersonen mee. Per matrixontwerp zijn er 20 proefpersonen toegewezen. De werving van de proefpersonen vond onder andere plaats via de proefpersoonpool vanuit de Universiteit Twente. Proefpersonen deden mee in ruil voor 'credits' die studenten in de eerste twee jaar moeten halen door mee te doen aan onderzoek. Daarnaast stelde het Centraal Bureau voor de Statistiek per respondent vijf euro beschikbaar. Deze beloning vond plaats in de vorm van irischeques. Daarbij werden proefpersonen geworven buiten het creditsysteem om. Door middel van het uitdelen van flyers op diverse plekken en het sturen van een mail aan een grote groep studenten is uiteindelijk het onderzoek uitgevoerd bij een groep van 40 studenten. Van vier deelnemers bleken de data onbruikbaar, vanwege technische redenen. Deze vier zijn daarom verwijderd uit de resultaten. De verhouding man-vrouw was met 19 om 17 bijna gelijk. De gemiddelde leeftijd was 21,90 jaar ($SD=2,21$).

Omdat één van de twee vragenlijsten ging over studiefinanciering, was het belangrijk er zeker van te zijn dat de proefpersonen niet minder betrokken waren bij het onderwerp als er geen studiefinanciering werd ontvangen. De proefpersonen werd daarom gevraagd of ze op zichzelf woonden of bij hun ouders en welk aandeel studiefinanciering had in het totale inkomen. Op deze manier kon vastgesteld worden dat variaties in de resultaten eventueel toe te wijzen waren aan de studiefinanciering die de studenten ontvingen. Naar de woonsituatie werd gevraagd omdat respondenten die in het ouderlijk huis wonen aanzienlijk minder afhankelijk zijn van de studiefinanciering dan uitwonende studenten en dit van invloed kon zijn op de betrokkenheid ten aanzien van het onderwerp. Van de 36 proefpersonen woonden er 32 buiten het ouderlijk huis, vier proefpersonen lieten weten bij de ouders te wonen. De respondenten die bij de ouders in huis woonden vertoonden geen significant lagere betrokkenheid dan de respondenten die buitenshuis woonden ($t=-1,25$, $p=0,22$, n.s.). Bij respondenten die geen studiefinanciering ontvingen bleek geen verschil te zijn in de betrokkenheid ten aanzien van de studiefinanciering ($t=-1,34$, $p=0,19$, n.s.). Er werd bij verdere analyses geen onderscheid gemaakt in de mate waarin proefpersonen studiefinanciering ontvingen.

2.7 Procedure

De respondenten hadden in totaal tussen de 20 en 30 minuten nodig voor de deelname aan het onderzoek. 10 minuten voor de briefing, het installeren van de proefpersoon, het instellen van de eye tracker en achteraf het debriefen. Vervolgens werd ongeveer 10 minuten besteed aan vragenlijst 1 en nog eens 10 minuten aan het invullen van vragenlijst 2. De duur van deze installatie verschilde per proefpersoon. Bij een aantal proefpersonen ging de installatie erg vlot en kon na vijf minuten al begonnen worden aan het onderzoek, terwijl het bij anderen wel op kon lopen tot een minuut of 12.

De gehele procedure begon met het welkomstwoord en de vraag of de respondent de vragenlijst wilde invullen over de demografische kenmerken. Na het invullen van deze vragenlijst volgde uitleg over het onderzoek. De proefpersoon werd verteld wat hem/haar tijdens het onderzoek te wachten stond en hoe het onderzoek in zijn werk ging. Hierbij werd kort uitgelegd dat de eye tracker werkt door middel van weerkaatsing van infrarood licht op de ogen en dat deze weerkaatsing wordt opgevangen door de twee camera's die onder het beeldscherm opgesteld zijn.

2.7.1 Het experiment

Na de installatie van de eye tracker volgde de uitleg over het vervolg van het onderzoek. De proefpersoon werd verteld dat zijn mening gevraagd zou worden over twee verschillende onderwerpen en dat het geheel vanzelf duidelijk zou worden. Als de proefpersoon geen vragen meer had, werd gestart met het onderzoek. Vervolgens verscheen de startpagina op het beeldscherm met daarop uitleg over het onderzoek. In deze uitleg stond dat het onderzoek over twee webvragenlijsten ging en dat begonnen zou worden met vragenlijst 1. Na de startpagina volgde uitleg over vragenlijst 1. Deze uitleg werd zo geformuleerd, dat de proefpersoon onmogelijk kon raden dat het onderzoek ging over de manier waarop de vragenlijst werd ingevuld en in mindere mate over de attitudes die de proefpersoon rapporteerde. Na het invullen van vragenlijst 1 verscheen op het scherm de korte vragenlijst over de betrokkenheid van de respondent bij het onderwerp en over de inspanning die de proefpersoon leverde tijdens het invullen van de vragenlijst.

Daarna werd overgegaan tot de uitleg over vragenlijst 2. Deze uitleg was wederom zo geformuleerd dat de proefpersoon het doel van het onderzoek niet duidelijk werd. Na het invullen van de tweede vragenlijst volgde de korte vragenlijst over de inspanning en betrokkenheid. Op het moment dat de proefpersoon hiermee klaar was, verscheen de slotpagina, waarop de proefpersoon werd bedankt en het daadwerkelijke doel van het onderzoek werd aangegeven.

2.7.2 Kwaliteit fixaties

Vooraf werd de proefpersoon verteld dat het belangrijk is het hoofd zo goed als stil te houden, kleine bewegingen zouden geen probleem zijn. De respondent werd verteld dat de proefleider aanwijzingen zou geven als de eye tracker software de ogen niet meer goed kon volgen. Als de proefleider bijvoorbeeld opmerkte dat de proefpersoon iets te ver naar links zat, kon hij vragen of laatstgenoemde iets naar rechts zou willen opschuiven. Tijdens het onderzoek werd op deze manier de kwaliteit van de data bewaakt. FaceLab, de eye tracker software, gaf continu scores aan de fixaties van de proefpersoon. Deze score, op een schaal van 1 tot 3, werd door de proefleider in de gaten gehouden. Een score onder de 2,5 betekende vaak een knippering van de ogen of een verkeerde positie van de proefpersoon. Afhankelijk van de oorzaak gaf de proefleider aanwijzingen om de score van de fixaties zo hoog mogelijk te houden. Bij drie proefpersonen moest een enkele keer een aanwijzing gegeven worden.

2.7.3 Debriefing

Na het invullen van de vragenlijst werd de proefpersoon in de gelegenheid gesteld vragen te stellen over het onderzoek. Dit gebeurde nadat de proefpersoon de slotpagina had gezien. Om er zeker van te zijn dat de proefpersoon het uiteindelijke doel van het onderzoek had begrepen, werd gevraagd of het doel van het onderzoek duidelijk was voor de respondent. Als dit het geval bleek te zijn, kon de respondent de ruimte verlaten.

3. Resultaten

In de volgende paragrafen wordt gekeken in hoeverre de hypothesen worden ondersteund. Ten eerste wordt gekeken naar de verschillende kenmerken van *satisficing* die ontlokt worden door het verschil in lay-out. Per verschijnsel van *satisficing* (*straightlining*, *weet niet* en *acquiescence*) werden de twee ontwerpen met elkaar vergeleken. Daarnaast werd per verschijnsel gekeken in hoeverre het al dan niet aanwezige effect versterkt werd door de variatie in betrokkenheid. Vervolgens worden de eye tracker data behandeld; het aantal fixaties en de duur van de fixaties zijn geanalyseerd om na te gaan in hoeverre respondenten zorgvuldiger of juist minder zorgvuldig te werk gingen bij een bepaald ontwerp. Daarnaast werd gekeken naar de betrokkenheid en inspanning die de respondenten zelf rapporteerden. De inspanning die de respondenten zelf rapporteerden wordt in dit hoofdstuk vergeleken met de cognitieve verwerking die volgens de eye tracker plaatsvond. Hieruit kan afgeleid worden wat de waarde is van eye tracking als onderzoeksinstrument. Tot slot is gekeken naar de kwaliteit van de antwoorden die de proefpersonen gaven op de stellingen. Er werd gekeken naar de zelfgerapporteerde attitudes en de verschillen hierin tussen het horizontale en het verticale ontwerp.

3.1 Manipulatiecheck

De respondenten werd een vragenlijst voorgelegd om de zelf gerapporteerde inspanning en betrokkenheid ten aanzien van de beide onderwerpen te meten. Tabel 3.1 geeft aan dat tussen de verschillende vragenlijsten een duidelijk verschil waar te nemen was in betrokkenheid en dat sprake was van een succesvolle manipulatie ($M=3,20$ vs. $5,67$, $t=-10,33$, $p=0,000$).

Tabel 3.1: Gemiddelde betrokkenheid (s.d.) per vragenlijst voor alle respondenten

	Lage betrokkenheid	Hoge betrokkenheid	t-waarde (df)	p-waarde
Betrokkenheid (1-7)	3,20 (1,03)	5,67 (1,07)	-10,33 (17)	0,000***

*** $p<0,01$, ** $p<0,05$, * $p<0,10$, n.s. = niet significant

3.2 Kwaliteit van de antwoorden

Als respondenten de vragenlijst zo snel mogelijk willen invullen, gaat dit vaak ten koste van de kwaliteit van de gerapporteerde antwoorden. Dit houdt in dat respondenten niet lang nadenken en het antwoord op de volgende vraag baseren op het gegeven antwoord op de vorige vraag. Om na te gaan in hoeverre de antwoorden die door de respondent zijn gegeven gebaseerd zijn op het voorgaande antwoord, wordt het absolute verschil tussen vraag x en vraag x-1 berekend. Op basis hiervan wordt de afstand bepaald tussen de verschillende gerapporteerde attitudes. De som van deze absolute waarden vormt de *straightlinescore*. Proefpersonen die door de verticale banen beïnvloed worden, zouden hun antwoorden minder overwegen en het antwoord eerder baseren op de antwoorden die op voorgaande vragen werden gegeven. Om na te gaan in hoeverre dit het geval is, laat tabel 3.2 de gemiddelde *straightlinescores* zien. Hoe lager de score, des te kleiner de afstand tussen vraag x en vraag x-1. Een hogere score zou dus impliceren dat de respondent zeer zorgvuldig de vragen aan het invullen was.

Tabel 3.2: Gemiddelde *straightlinescores* (s.d.) per vragenlijst

	Horizontaal	Verticaal	t-waarde (df)	p-waarde
Vragenlijst				
Lage betrokkenheid	0,95 (,26)	1,01 (,22)	-,736 (34)	n.s.
Hoge betrokkenheid	1,27 (,24)	1,25 (,19)	,227 (34)	n.s.
t-waarde (df)	-6,37 (17)	-5,423 (17)		
P-waarde	,000***	,000***		

*** $p < 0,01$, ** $p < 0,05$, * $p < 0,10$, n.s. = niet significant

Uit tabel 3.2 komt naar voren dat op vragenlijstniveau hypothese 1 niet wordt ondersteund. Tussen de verschillende ontwerpen werd geen verschil geconstateerd in gemiddelde *straightlinescores*. Het ontwerp zou dus niet van invloed zijn op de gerapporteerde attitudes. Een interessant verschil dat bij tabel 3.2 wel naar voren komt, is dat respondenten bij de vragenlijst met hoge betrokkenheid minder geneigd waren het antwoord op de volgende vraag te baseren op die eraan voorafging volgde. Dit ging op voor beide ontwerpen. Dat wil zeggen; de antwoorden zijn gevarieerder en in mindere mate gebaseerd op elkaar.

Tabel 3.3: Betrokkenheid, lay-out en lay-out x betrokkenheid voor *straightlinescores*

	F (df)	p-waarde
Betrokkenheid	69,59 (1, 34)	,000***
Lay-out	0,19 (1, 34)	n.s.
Betrokkenheid x lay-out	0,52 (1, 34)	n.s.

*** $p < 0,01$, ** $p < 0,05$, * $p < 0,10$, n.s. = niet significant

Tabel 3.3 toont aan dat er geen sprake is van een interactie-effect voor de gemiddelde *straightlinescores*. Hypothese 2 wordt dus verworpen. De verwachte verschillen in *straightlinescores* tussen de hoge en lage betrokkenheid bleken significant ($F=69,59$, $p=0,000$). Dit significante verschil werd niet versterkt door het verschil in ontwerp. Voor zowel het horizontale als het verticale ontwerp waren de verschillen tussen hoge en lage betrokkenheid even groot ($F=0,52$, $p=0,477$).

Een andere maat ter aanduiding van *satisficing* is het aantal keer dat de proefpersoon *weet niet* heeft ingevuld ten opzichte van de hele vragenlijst. Dit geeft informatie over het aantal keer dat de respondenten zichzelf de last van het nadenken bespaarden en geen attitude rapporteerden op de vragen. Tabel 3.4 geeft het gemiddelde aantal keren weer dat een respondent gebruik maakte van de optie *weet niet*.

Tabel 3.4: Percentage *weet niet* ingevuld per vragenlijst

	Horizontaal	Verticaal	t-waarde (df)	p-waarde
Lage betrokkenheid	1,88 (3,63)	3,76 (6,06)	-1,13 (34)	n.s.
Hoge betrokkenheid	2,58 (3,14)	1,98 (2,29)	0,56 (34)	n.s.
t-waarde (df)	0,91 (17)	-1,35 (17)		
p-waarde	n.s.	n.s.		

*** $p < 0,01$, ** $p < 0,05$, * $p < 0,10$, n.s. = niet significant

Tabel 3.4 toont aan dat het percentage *weet niet* wel verschilde per ontwerp, maar dat deze verschillen niet significant zijn. Hypothese 3 wordt niet ondersteund. Opvallend is dat het verschil tussen het aantal keer dat de respondent geen duidelijke attitude liet merken bij het verticale ontwerp groter is per vragenlijst dan bij het horizontale ontwerp. Dit verschil bleek echter niet significant. Volgens deze maat zou er dus geen verschil zijn in *satisficing* tussen het horizontale en het verticale ontwerp.

Tabel 3.5: Betrokkenheid, lay-out en lay-out x betrokkenheid voor percentage *weet niet*

	F (df)	p-waarde
Betrokkenheid	1,59 (1, 34)	n.s.
Lay-out	0,40 (1, 34)	n.s.
Betrokkenheid x lay-out	2,52 (1, 34)	n.s.

*** $p < 0,01$, ** $p < 0,05$, * $p < 0,10$, n.s. = niet significant

Tabel 3.5 laat zien dat er geen interactie-effect optreedt voor de percentages *weet niet* scores. Dit was te verwachten, aangezien de *weet niet* scores voor beide ontwerpen niet significant verschilden tussen hoge en lage betrokkenheid ($F=1,59$, $p=0,22$, n.s.). Hypothese 4 wordt dus verworpen.

Naast de variatie in de antwoorden en het aantal *weet niet* werden de gerapporteerde attitudescores geanalyseerd. Hieruit moest blijken in welke mate de lay-out van invloed is op de attitudes die respondenten rapporteren. Aangezien de vragenlijsten alleen verschilden in lay-out, kan gekeken worden naar de invloed die deze lay-out had op de attitudes die respondenten rapporteerden bij de verschillende stellingen. De optie *weet niet* is voor het opstellen van deze tabel omgezet naar de gemiddelde attitude die respondenten gaven per matrix.

Tabel 3.6: Gemiddelde attitudescores (s.d.) per matrix per ontwerp

	Horizontaal	Verticaal	t-waarde (df)	p-waarde
Lage betrokkenheid				
Matrix 1 (1-5)	3,56 (,40)	3,76 (,32)	-1,644 (34)	n.s.
Matrix 2 (1-5)	3,04 (,33)	3,16 (,32)	-1,030 (34)	n.s.
Matrix 3 (1-5)	2,52 (,18)	2,81 (,35)	-3,193 (34)	,003**
Matrix 4 (1-5)	3,61 (,28)	3,66 (,25)	-,666 (34)	n.s.
Matrix 5 (1-5)	3,13 (,26)	3,25 (,29)	-1,231 (34)	n.s.
Matrix 6 (1-5)	3,00 (,19)	3,04 (,24)	-,571 (34)	n.s.
Matrix 7 (1-5)	3,11 (,26)	3,15 (,34)	-,390 (34)	n.s.
Vragenlijst (1-5)	3,14 (,17)	3,27 (,15)	-2,296 (34)	,028**
Hoge betrokkenheid				
Matrix 1 (1-5)	2,99 (,42)	3,14 (,40)	-1,086 (34)	n.s.
Matrix 2 (1-5)	2,68 (,35)	2,78 (,38)	-,859 (34)	n.s.
Matrix 3 (1-5)	2,94 (,31)	2,91 (,25)	,340 (34)	n.s.
Matrix 4 (1-5)	2,98 (,41)	3,20 (,39)	-1,630 (34)	n.s.
Matrix 5 (1-5)	2,58 (,33)	2,80 (,37)	-1,812 (34)	,079*
Matrix 6 (1-5)	2,76 (,36)	2,96 (,43)	-1,557 (34)	n.s.
Vragenlijst (1-5)	2,82 (,18)	2,97 (,19)	-2,354 (34)	,025**

*** $p < 0,01$, ** $p < 0,05$, * $p < 0,10$, n.s. = niet significant

In tabel 3.6 is af te lezen dat er op matrixniveau slechts één significant verschil per vragenlijst was waar te nemen tussen de beide ontwerpen op de antwoorden die werden gegeven. Als gekeken wordt naar de gemiddelde scores voor de gehele vragenlijst, verschillen beide vragenlijsten in de antwoorden die de respondenten gaven. Hypothese 5 wordt dus ondersteund. De respondenten die de vragenlijst met de verticale lijnen invulden, waren over het algemeen meer geneigd positieve antwoorden te geven dan de respondenten die de vragenlijsten met horizontale lijnen voorgelegd kregen. Dit zou verklaard kunnen worden door de neiging die satisficers hebben positievere attitudes te rapporteren.

3.3 Eye tracker data

De volgende paragrafen beschrijven de verwerkingstijd van de informatie voor de hersenen. Voor de analyse van de data zijn de verschillende fixaties ingedeeld in een zone voor het vraagdeel en een zone voor het antwoorddeel. Als een verschil tussen de twee ontwerpen in fixaties en langere fixaties gevonden zou worden, zou dit enerzijds kunnen betekenen dat de proefpersoon meer tijd nodig had om de informatie te kunnen verwerken in de hersenen. Anderzijds zou een verschil kunnen duiden op de zorgvuldigheid waarmee te werk is gegaan tijdens het invullen. De eerste twee tabellen in deze paragraaf geven de uitkomsten voor de twee ontwerpen apart van elkaar om verschillen te constateren tussen de diverse ontwerpen van de vragenlijsten (hoge versus lage betrokkenheid). Vervolgens wordt een aantal tabellen getoond waarbij de verschillen per onderwerp tussen de diverse ontwerpen getoetst worden en in hoeverre deze verschillen de geconstateerde verschillen tussen de ontwerpen versterken.

Tabel 3.7: Gemiddeld aantal fixaties (s.d.) duur fixaties in seconden

	Horizontaal	Verticaal	t-waarde (df)	p-waarde
Aantal fixaties totaal				
Lage betrokkenheid	1011,00 (308,19)	1197,46(298,54)	-1,84 (34)	,074*
Hoge betrokkenheid	1172,39 (357,83)	1406,89 (389,17)	-1,88 (34)	,068*
Aantal fixaties vraagdeel				
Lage betrokkenheid	621,67 (268,37)	797,95 (283,95)	-1,91 (34)	,064*
Hoge betrokkenheid	778,72 (325,22)	978,67 (349,26)	-1,78 (34)	,084*
Aantal fixaties antwoorddeel				
Lage betrokkenheid	389,33 (66,03)	399,51 (67,55)	-0,46 (34)	n.s.
Hoge betrokkenheid	393,67 (74,75)	428,22 (103,08)	-1,15 (34)	n.s.
Duur fixaties totaal				
Lage betrokkenheid	191,45 (59,43)	237,42 (69,64)	-2,13 (34)	,040**
Hoge betrokkenheid	221,03 (77,91)	274,40 (87,73)	-1,93 (34)	,062*
Duur fixaties vraagdeel				
Lage betrokkenheid	101,61 (43,73)	142,44 (59,65)	-2,34 (34)	,025**
Hoge betrokkenheid	134,38 (65,83)	176,78 (75,55)	-1,80 (34)	,081*
Duur fixaties antwoorddeel				
Lage betrokkenheid	89,84 (20,15)	94,98 (21,62)	-0,74 (34)	n.s.
Hoge betrokkenheid	86,66 (20,35)	97,62 (25,97)	-1,41 (34)	n.s.

*** $p < 0,01$, ** $p < 0,05$, * $p < 0,10$ n.s. = niet significant

Tabel 3.7 geeft de verschillen in fixaties weer tussen de beide ontwerpen. Het is duidelijk te zien dat de proefpersonen bij het verticale ontwerp meer fixaties nodig hadden voor het lezen van de vragen. Hypothese 6 wordt dus niet ondersteund. De lay-out zorgde ervoor dat respondenten bij het verticale ontwerp gemiddeld een kleine 798 fixaties nodig hadden om alle vragen in de lage betrokkenheidsvragenlijst te lezen. Dit tegenover de 622 fixaties die gemiddeld bij het horizontale ontwerp nodig waren om alle vragen te lezen. Dit verschil bleek statistisch significant ($t = -1,91$, $p = 0,064$). Ditzelfde verschil is ook significant voor de vragenlijst waarbij de betrokkenheid hoog is: gemiddeld 779 fixaties in de horizontale groep tegenover 978 bij de verticale groep ($t = -1,78$, $p = 0,084$).

Als we echter kijken naar de verschillen tussen de twee ontwerpen bij het antwoorddeel zijn er geen significante verschillen op te merken. Het verschil in lay-out is dus alleen van invloed op de verwerking van de vragen en niet op de verwerking van de antwoorden of het overwegen van de verschillende antwoordopties.

Bij de som van de duur van de fixaties voor zowel het vraagdeel als het antwoorddeel, zijn er ook significante verschillen op te merken. Hypothese 7 wordt dus ook niet ondersteund. Het gehele invulproces lijkt beïnvloed te worden door de lay-out. Van het lezen van de vraag tot het aankruisen van een antwoord; het geheel kostte meer tijd en meer fixaties voor de res-

pondenten die het verticale ontwerp invulden. Voor het onderwerp waarbij de betrokkenheid laag was, was de som van de fixaties voor de gehele vragenlijst gemiddeld 191,45 seconden voor het horizontale ontwerp, tegenover 237,42 in de versie met verticale banen ($t=2,13$, $p=0,040$). Voor het onderwerp waarbij de betrokkenheid hoog was, bedroeg de gemiddelde som van de fixaties 221,03 seconden voor het horizontale ontwerp, waar dit cijfer 274,40 was bij het verticale ontwerp ($t=-1,93$, $p=0,062$).

3.4 Eye tracking als objectief onderzoeksinstrument

De volgende stap is te kijken in hoeverre eye tracking als nuttig instrument kan dienen om de gebruiksvriendelijkheid van vragenlijsten te testen. Om dit na te gaan, werden allereerst de twee vragenlijsten met elkaar vergeleken, binnen elk ontwerp. Op die manier kan worden nagegaan welke vragenlijst objectief gemeten zorgvuldiger door de respondenten werd verwerkt. Door de zelfgerapporteerde inspanning met de duur van de fixaties en het totale aantal fixaties met elkaar te vergelijken, kan blijken of de eye tracker de zelfgerapporteerde inspanning van de respondenten in kaart kan brengen.

Tabel 3.8 laat de verschillen tussen de hoge en lage betrokkenheid zien. De lage betrokkenheidsvragenlijst bevatte 68 vragen. De vragenlijst waarbij het onderwerp hoog betrokken was, bevatte 56 vragen. Om deze reden zijn de fixaties en de duur van de fixaties die op het antwoorddeel plaatsvonden in de vragenlijst met de hoge betrokkenheid gedeeld door 56 en vervolgens vermenigvuldigd met 68. Zo kon het aantal fixaties en de duur van de fixaties binnen de twee ontwerpen op beide vragenlijsten goed vergeleken worden. Bij de vraagdelen van de vragenlijsten werd het aantal fixaties en de duur ervan gewogen naar het aantal woorden dat elke vragenlijst had. De vraagdelen van elke vraag werden gewogen naar het aantal woorden dat elke vragenlijst bevatte. Zo had de vragenlijst met lage betrokkenheid 836 woorden, de hoge betrokkenheid 813. Het aantal fixaties en de duur ervan is voor de vragenlijst met de hoge betrokkenheid gedeeld door 813 en vermenigvuldigd met 836. De onderstaande tabel laat het gewogen aantal fixaties zien en de gewogen duur van deze fixaties.

Tabel 3.8: Gemiddeld totaal aantal fixaties en duur fixaties per vragenlijst (s.d.), gewogen naar aantal woorden en aantal vragen

	Lage betrokkenheid	Hoge betrokkenheid	t-waarde (df)	p-waarde
Aantal fixaties totaal				
Horizontaal	1011,00 (308,19)	1278,77 (375,38)	-6,18 (17)	,000***
Verticaal	1197,46 (434,51)	1526,34 (410,09)	-7,00 (17)	,000***
Aantal fixaties vraagdeel				
Horizontaal	621,67 (268,37)	800,75 (334,42)	-4,99 (17)	,000***
Verticaal	797,95 (283,95)	1006,35 (359,14)	-5,81 (17)	,000***
Aantal fixaties antwoorddeel				
Horizontaal	389,33 (66,03)	478,03 (90,76)	-5,46 (17)	,000***
Verticaal	399,51 (67,55)	519,98 (125,17)	-5,47 (17)	,000***
Duur fixaties totaal				
Horizontaal	191,45 (59,43)	243,40 (82,72)	-5,43 (17)	,000***

Verticaal	237,42 (69,64)	300,32 (93,12)	-6,70 (17)	,000***
Duur fixaties vraagdeel				
Horizontaal	101,61 (43,73)	138,17 (67,69)	-4,68 (17)	,000***
Verticaal	142,44 (59,65)	181,78 (77,69)	-5,28 (17)	,000***
Duur fixaties antwoorddeel				
Horizontaal	89,84 (20,15)	105,23(24,71)	-3,94 (17)	,001***
Verticaal	94,98 (21,62)	118,53 (31,54)	-5,28 (17)	,000***

*** $p < 0,01$, ** $p < 0,05$, * $p < 0,10$, n.s. = niet significant

Uit tabel 3.8 blijkt dat het invulgedrag van de proefpersonen verschilde tussen de hoge en lage betrokkenheid. Alle respondenten gingen significant zorgvuldiger te werk in de vragenlijst met de hoge betrokkenheid dan in de vragenlijst met lage betrokkenheid. Er vonden meer fixaties plaats bij de vragenlijst met het persoonlijk betrokken onderwerp dan in de vragenlijst waarbij het onderwerp minder persoonlijk betrokken bleek. Dit ging op voor zowel het vraagdeel als het antwoorddeel. Respondenten lezen de vraag zorgvuldiger als het onderwerp persoonlijk betrokken is. Bij een persoonlijk minder betrokken onderwerp gingen de proefpersonen sneller door de vragen heen en werd sneller tot een antwoord gekomen.

Tussen de twee versies (horizontaal versus verticaal) bleek ook sprake te zijn van significante verschillen. Met name als er gebruik werd gemaakt van het verticale design, hadden de respondenten meer tijd en meer fixaties nodig voor het persoonlijk minder betrokken onderwerp. De onderstaande interactie-effecten (tabel 3.9) laten zien in hoeverre dit verschil versterkt werd door het verschil in ontwerp.

Tabel 3.9: Hoofdeffecten voor betrokkenheid, lay-out en interactie-effect voor lay-out x betrokkenheid

Totaal					
Aantal fixaties	F (df)	p-waarde	Duur fixaties	F (df)	p-waarde
Betrokkenheid	87,032 (1, 34)	,000***	Betrokkenheid	73,43 (1, 34)	,000***
Lay-out	3,72 (1, 34)	,062*	Lay-out	4,28 (1, 34)	,046*
Betrokkenheid x lay-out	0,91 (1, 34)	n.s.	Betrokkenheid x lay-out	0,67 (1, 34)	n.s.
Vraagdeel					
Aantal fixaties	F (df)	p-waarde	Duur fixaties	F (df)	p-waarde
Betrokkenheid	58,33 (1, 34)	,000***	Betrokkenheid	51,17 (1, 34)	,000***
Lay-out	3,55 (1, 34)	,068*	Lay-out	4,26 (1, 34)	,047*
Betrokkenheid x lay-out	0,33 (1, 34)	n.s.	Betrokkenheid x lay-out	0,80(1, 34)	n.s.
Antwoorddeel					
Aantal fixaties	F (df)	p-waarde	Duur fixaties	F (df)	p-waarde
Betrokkenheid	58,35 (1, 34)	,000***	Betrokkenheid	43,14 (1, 34)	,000***
Lay-out	0,94 (1, 34)	n.s.	Lay-out	1,41 (1, 34)	n.s.
Betrokkenheid x lay-out	1,35 (1, 34)	n.s.	Betrokkenheid x lay-out	1,90 (1, 34)	n.s.

*** $p < 0,01$, ** $p < 0,05$, * $p < 0,10$, n.s. = niet significant

Tabel 3.9 toont aan dat er geen interactie-effect bleek te zijn tussen de lay-out en de betrokkenheid. Binnen elk ontwerp tussen de hoge en lage betrokkenheid waren wel duidelijk significante verschillen te ontdekken, maar de betrokkenheid en de lay-out hadden gezamenlijk

geen invloed op de invultijd. Er kan dus gezegd worden dat de som van de fixaties en het aantal fixaties verschilden per ontwerp en per onderwerp (hoge versus lage betrokkenheid) maar dat deze effecten elkaar niet versterkten. Hypothesen 8 en 9 worden niet ondersteund.

3.5 Zelf-gerapporteerde inspanning en betrokkenheid als subjectief instrument

Tabel 3.10 toont aan in hoeverre door de proefpersonen verschil in betrokkenheid en inspanning werd gerapporteerd tussen de beide vragenlijsten en tussen de verschillende ontwerpen.

Tabel 3.10: Gemiddelde zelf-gerapporteerde inspanning (s.d.) per onderwerp naar ontwerp

	Lage betrokkenheid (1-7)	Hoge betrokkenheid (1-7)	t-waarde (df)	p-waarde
Horizontaal (n=18)	4,25 (0,72)	5,04 (0,81)	-4,69 (17)	,000***
Verticaal (n=18)	4,35 (1,06)	4,56 (1,10)	-,793 (17)	n.s.
t-waarde	-,32 (34)	1,49 (34)		
p-waarde	n.s.	n.s.		

*** $p < 0,01$, ** $p < 0,05$, * $p < 0,10$, n.s. = niet significant

Uit de eye tracker data bleek dat de respondenten die de verticale versie voorgelegd kregen in het gehele antwoordproces meer fixaties hadden, die bovendien gezamenlijk langer duurden dan de fixaties van de respondenten die te maken kregen met het horizontale ontwerp. Er kan dus verwacht worden dat deze respondenten ook subjectief meer inspanning zouden leveren voor het invullen van de vragenlijsten. De zelfrapportage bewijst echter dat dit niet zo is. Respondenten lieten zien zelf geen verschil te hebben ervaren tussen de inspanning die geleverd moest worden bij het beantwoorden van de verticale vragenlijst versus de inspanning die geleverd moest worden bij het beantwoorden van de horizontale vragenlijst. Wat bovendien interessant is aan tabel 3.10, is dat de proefpersonen bij het verticale ontwerp geen verschil rapporteerden tussen de inspanning die geleverd werd tijdens het invullen van de vragenlijst met hoge betrokkenheid en de inspanning die geleverd werd bij de vragenlijst met lage betrokkenheid.

Tabel 3.11: Betrokkenheid, lay-out en lay-out x betrokkenheid voor zelfgerapporteerde inspanning

	F (df)	p-waarde
Betrokkenheid	11,95 (1, 34)	,002***
Lay-out	0,56 (1, 34)	n.s.
Betrokkenheid x lay-out	4,80 (1, 34)	,036**

*** $p < 0,01$, ** $p < 0,05$, * $p < 0,10$, n.s. = niet significant

Tabel 3.11 toont de interactie-effecten tussen de lay-out en betrokkenheid. De verwachte verschillen in inspanning tussen hoge en lage betrokkenheid (zie tabel 3.5) worden versterkt door het ontwerp. De respondenten die het horizontale ontwerp voorgelegd kregen, rapporteerden wél een verschil in inspanning tussen hoge en lage betrokkenheid, terwijl de respondenten die het verticale ontwerp voorgelegd kregen, dit verschil niet zelf rapporteerden. Dit verschil in variatie bleek significant ($F=4,80$, $p=.036$). Hier biedt de eye tracker zeer waardevolle gegevens, want dit interactie-effect trad niet op voor de door de eye tracker gemeten zorgvuldigheid van verwerking. Er kan geconcludeerd worden dat hypothese 10 niet wordt ondersteund.

4. Discussie

4.1 De invloed van lay-out op *satisficing*

Van alle 10 hypotheses is alleen 5 overeind gebleven: de respondenten die de vragenlijst met de verticale lijnen invulden, waren over het algemeen meer geneigd positieve antwoorden te geven dan de respondenten die de vragenlijsten met horizontale lijnen voorgelegd kregen. De resultaten lieten zien dat het aantal fixaties dat de gemiddelde proefpersoon nodig had voor het invullen van de vragenlijsten tussen het horizontale ontwerp en het verticale ontwerp alleen verschilde voor het vraagdeel en niet voor het antwoorddeel. Voornamelijk het verschil in antwoorddeel was wel te verwachten, omdat werd aangenomen dat de respondenten door de lay-out dermate gestuurd werden, dat het erg verleidelijk werd zichzelf de cognitieve last te besparen. Respondenten waren echter niet gevoelig voor de lay-out als het ging om het overwegen van een antwoord. Bij beide ontwerpen werden hier even veel en even lange fixaties gemeten. De waargenomen verschillen in fixaties tussen het horizontale en het verticale ontwerp vonden alleen plaats in het vraagdeel van de vraag. Dit ging op voor zowel de hoge betrokkenheid als de lage betrokkenheid. Er mag dus geconcludeerd worden dat volgens objectieve meting van de zorgvuldigheid deze verticale lijnen geen invloed hadden voor het overwegen van de antwoordopties en het rapporteren van de best passende attitude.

Zoals Cooke (2006) aangaf, is het niet eenvoudig te achterhalen waarom proefpersonen op bepaalde delen veel fixaties en een langere duur van deze fixaties hadden. Een tweetal verklaringen is aan de hand van het theoretisch kader te formuleren:

1. Bekeken vanuit een ontwerp-perspectief in het *human-computer-interaction* veld, zouden de langere fixaties en de langere duur ervan kunnen betekenen dat het ontwerp dermate onvriendelijk was dat respondenten meer moeite moesten doen om met deze onvriendelijkheid om te kunnen gaan.
2. Bekeken vanuit cognitief perspectief op het gebied van invulgedrag in webvragenlijsten, kan gesteld worden dat meer fixaties juist langere verwerking van de informatie in de hersenen betekenen en dat er zorgvuldiger te werk wordt gegaan door de respondenten. Krosnick (1991) noemt dit *optimizing*.

Als er daadwerkelijk meer *optimizing* plaatsvond tijdens het invullen van de vragenlijst, zou dit betekenen dat er een verschil zou zijn in de kwaliteit van de antwoorden die respondenten gaven. Dit verschil in kwaliteit van data ontbrak tussen respondenten in de verticale, vergeleken met respondenten in het horizontale ontwerp. Gezegd kan worden dat meer fixaties en een langere duur van de som van de fixaties duiden op een onvriendelijk ontwerp waarbij de grotere cognitieve zorgvuldigheid niet ten goede komt aan de kwaliteit van de beantwoording, maar wordt besteed aan het overwinnen van de barrières opgeworpen door de slechte lay-out.

De variatie in ontwerp had alleen invloed op de manier waarop het vraagdeel van de vraag werd gelezen en niet op de manier waarop respondenten de antwoorden overwogen. Respondenten waren zorgvuldiger in de cognitieve verwerking voor het vraagdeel bij het verticale ontwerp, vergeleken met de respondenten die het horizontale ontwerp invulden. Bekeken vanuit het *principle of common region* van Palmer (1992), is hieruit een tweetal zaken te concluderen op het gebied van de cognitieve verwerking van de lay-out:

1. Respondenten zagen de verticale baan aan het einde van de vraag als barrière en stopten daar, alvorens tot een antwoord te komen. Het aantal fixaties in het antwoorddeel zou dus ook gelijk kunnen zijn omdat de respondenten bij het verticale ontwerp al nadenken over het antwoord, vlak nadat ze de vraag lezen, maar de bijbehorende fixaties vinden niet plaats in het antwoorddeel. Deze aanname lijkt erg onwaarschijnlijk, aangezien alleen de fixaties en de duur ervan in het vraagdeel groter waren dan in het antwoorddeel. Als de respondenten bij het verticale ontwerp daadwerkelijk aan het *optimizen* waren, zou op beide delen een significant verschil verwacht worden ten opzichte van het horizontale ontwerp.
2. De andere aanname is dat door het ontbreken van de horizontale banen de respondenten geen referentiekader hadden voor het lezen van de vraag. De ogen bewogen onbewust omhoog of naar beneden, omdat er geen horizontale baan was om ze in de goede richting vast te houden. De respondent moest bij het verticale ontwerp dus de aandacht er goed bij houden om de vragen te kunnen lezen en dit zorgde voor meer fixaties en ook voor langere fixaties bij het verticale ontwerp. Aan de hand van de data is het moeilijk na te gaan in hoeverre de ogen afweken van een bepaalde lijn, omdat een reeks fixaties vaak niet op eenzelfde lijn plaatsvond, maar dat er kleine variaties waren ten opzichte van de ideale lijn. Wellicht dat deze afwijking groter is bij één van de twee ontwerpen, maar dan zou de positie van elke fixatie ten opzichte van de vorige fixatie bekeken moeten worden. Omwille van de tijd is hier in dit onderzoek geen aandacht aan besteed.

Uitgaande van aanname 1 wordt aangeraden om vanwege het onvriendelijke design geen gebruik te maken van verticale banen. Respondenten zouden de vragen zorgvuldiger lezen, maar deze toename in fixaties werd alleen veroorzaakt door het gebruiksonvriendelijke ontwerp van de vragenlijst en heeft dus weinig te maken met de zorgvuldigheid van de verwerking. Dit wordt bovendien ondersteund doordat er alleen een verschil bleek voor het lezen van de vragen en niet voor het komen tot een zo passend mogelijk antwoord.

Uit de literatuur kwam naar voren dat het meten van *satisficing* een zeer gecompliceerde taak vormt. Dit onderzoek kan dat ondersteunen. De beoogde indicatoren van *satisficing* gebruikt in dit onderzoek waren het aandeel *weet niet* en de *straightlinescores*. Door het gebruik van verticale banen werd er echter een ander aspect van *satisficing* in de hand gewerkt, namelijk het rapporteren van positievere attitudes. Op deze manier wordt de meetfout 'biased response behavior' alleen maar groter en komen de gerapporteerde attitudes verder af te staan van de attitudes die de respondenten daadwerkelijk hebben. Aan de hand van de resultaten is niet eenvoudig te zeggen of *satisficing* in sterkere mate plaatsvond bij een bepaald ontwerp. In ieder geval vertoonden de respondenten bij het verticale ontwerp

niet meer fixaties bij het overwegen van de antwoorden, waren de data niet minder gedifferentieerd en was er geen verschil in het aandeel *weet niet*. Toch blijkt dat de attitudes positiever waren. Gesproken zou kunnen worden over een verschil in satisfice gedrag, maar dit verschil is wel erg klein. Op basis van deze bevindingen zou er geen voorkeur voor een bepaald design bestaan. Bij het verticale ontwerp was het niet zo dat *satisficing* in veel sterkere mate plaatsvond, daar slechts één van de drie *satisfice* verschijnselen verschilde.

Het is aan te nemen dat dit verschil ook aanwezig is bij papieren vragenlijsten. Vanuit de *gestaltprincipes* blijft de verwerking van de visuele elementen hetzelfde, ook al verandert het medium. Een belangrijk verschil is wel dat bij papieren vragenlijsten respondenten terug kunnen bladeren en vooruit, om te zien hoeveel vragen er nog beantwoord moeten worden. Dit kan van invloed zijn op het eerder afbreken van de vragenlijst of het zo snel mogelijk afmaken ervan. In dit onderzoek was er voor de respondenten niet de mogelijkheid terug te 'bladeren' in de vragenlijst. Door de ontwikkelingen binnen de nieuwe media, lijkt het onvermijdelijk dat vragenlijsten binnenkort op de mobiele telefoon of op de iPad ingevuld kunnen worden. De presentatie van matrixvragen zal op deze platforms andere vormen aannemen. Het is aannemelijk dat veel ontwerpprincipes, zoals het groeperen van vragen, de nodige effecten behouden vanwege de aard van de cognitieve verwerking. Een kleiner formaat kan ertoe leiden dat er slechts een aantal vragen op één scherm past, waardoor vragen minder snel met elkaar in verband worden gezien en satisfice gedragingen dus minder plaatsvinden. Anderzijds is er op een klein scherm minder mogelijk qua gebruiksvriendelijkheid. Hierdoor vergt het invullen van een vragenlijst meer cognitieve inspanning en de vraag is of respondenten bereid zijn die inspanning te leveren. Met als resultaat: juist meer satisfice gedrag. Mocht het invullen van vragenlijsten veelvuldig gaan gebeuren op de kleinere schermen, dan is onderzoek naar dit specifieke invulgedrag zeer nuttig.

4.2 Objectief gemeten zorgvuldigheid versus subjectieve inspanning

De eye tracker data gaven inzicht in de cognitieve verwerking van de proefpersonen tijdens het invullen van de vragenlijst. Hoe meer fixaties en hoe langer de duur van de fixaties, des te groter de zorgvuldigheid die er werd gemeten door de eye tracker. Hieruit bleek dat het aantal fixaties groter en de duur ervan langer waren voor de vragenlijst met hoge betrokkenheid, vergeleken met de vragenlijst met lage betrokkenheid. Dit verschil in zorgvuldigheid bleek voor beide ontwerpen significant.

Daarnaast werd proefpersonen gevraagd om zelf aan te geven hoeveel inspanning er was geleverd bij het invullen van de beide onderwerpen. Hierbij is het opvallendste resultaat dat de zelfgerapporteerde inspanning tussen hoge en lage betrokkenheid wél verschilde voor het horizontale ontwerp, maar dat dit verschil niet significant bleek voor het verticale ontwerp. Respondenten die te maken kregen met het verticale ontwerp gaven aan dat er voor beide vragenlijsten evenveel mentale inspanning werd geleverd. Dit in tegenstelling tot respondenten die het horizontale ontwerp te zien kregen, waar wel verschil in zelfgerapporteerde inspanning tussen beide onderwerpen gerapporteerd werd. Hier bleek sprake te zijn van een interactie-effect; de verschillen in inspanning tussen hoge en lage betrokkenheid werden versterkt door het verticale ontwerp. De rol die de lay-out hierbij speelt is opvallend te noemen. De respondenten hadden bij het horizontale ontwerp zelf door dat meer inspan-

ning geleverd werd naarmate de betrokkenheid varieerde, terwijl de respondenten dit niet zo ervoeren bij het verticale ontwerp.

Dit kan betekenen dat de variatie in lay-out onbewust invloed heeft op het invulgedrag, aangezien de inspanning die men zegt te leveren op beide vragenlijsten constant bleef bij het verticale ontwerp. De respondenten beseften zelf niet dat er meer inspanning geleverd diende te worden. Dit biedt interessante conclusies voor het gebruik van eye tracking als onderzoeksinstrument. Eye tracking kan op objectieve wijze de zorgvuldigheid van de cognitieve verwerking in kaart brengen. Deze verwerking kan sneller gebeuren door gebruik te maken van een gebruiksvriendelijk design. Echter, de respondenten hadden zelf niet door dat deze variatie zorgde voor verschil in inspanning. Vragenlijstbouwers kunnen de inspanning wel zo klein mogelijk maken door de lay-out gebruiksvriendelijk te maken, maar de mindere inspanning die de respondent levert, gaat ongemerkt.

Omdat deze twee onderzoeksmethoden elkaar lijken tegen te spreken, rijst de vraag welke van de twee het best gebruikt kan worden. Als eye tracking in dit onderzoek niet gebruikt zou worden, zouden er geen verschillen in inspanning waargenomen zijn tussen de verschillende typen lay-out. Zowel de horizontale als de verticale vragenlijst zou zonder problemen gelanceerd worden, aangezien het de respondenten niet meer moeite kostte de vragenlijsten in te vullen. Eye tracking laat echter zien dat de respondenten bij het verticale ontwerp beduidend meer moeite besteedden aan het gehele invulproces.

Dit maakt eye tracking tot een nuttig onderzoeksinstrument voor gebruiksvriendelijkheids-tests. Hierbij moet wel rekening gehouden worden met het verschil tussen subjectieve en objectieve onderzoeksmethoden. De voorkeur voor een bepaalde onderzoeksmethode kan verschillen per onderzoeker en elk type methode heeft zowel voor- als nadelen. Objectieve onderzoeksmethoden kunnen de tijd die de respondent te besteden heeft zo goed mogelijk in kaart brengen en aanbevelingen doen over het zo efficiënt mogelijk indelen van deze tijd. Dit terwijl aanhangers van subjectieve methoden aanvoeren dat het gaat om de beleving van de vragenlijst. Als er onbewust maar een bepaald aantal seconden of minuten besteed wordt aan een vragenlijst, kunnen de objectieve methoden deze goed in kaart brengen en zorgdragen voor efficiënte tijdsindeling. De lay-out zorgde er in dit onderzoek bijvoorbeeld voor dat respondenten voor het lezen van de vraag extra tijd kwijt waren. Deze tijd kan, bij een beperkt aantal seconden of minuten, niet besteed worden aan het geven van een antwoord, omdat er al relatief te veel inspanning is geleverd voor het lezen van de vraag.

4.3 Eye tracking ter aanduiding van *satisficing*

De resultaten aan de hand van de eye tracker data tonen aan dat respondenten bij het verticale ontwerp niet bezig waren met *satisficing*, aangezien er meer fixaties werden besteed aan het gehele antwoordproces. Op het eerste gezicht zou eerder sprake kunnen zijn van *optimizing*, omdat de vragen zorgvuldiger worden gelezen bij het verticale ontwerp. De *straightlinescores* die berekend werden, lieten echter geen verschillen zien tussen het horizontale en het verticale ontwerp. Het mag dus duidelijk zijn dat respondenten niet per se de vragen zorgvuldiger lazen, maar dat men door het onvriendelijke design meer moeite had de vragen te lezen. Hierdoor ontstond variatie in het aantal fixaties en waren de fixaties langer voor de proefpersonen bij het verticale ontwerp.

Een ander punt dat deze aanname ondersteunt is dat verwacht werd dat men bij het verticale ontwerp minder fixaties zou besteden en dat daarmee het percentage *weet niet* hoger zou zijn bij dit ontwerp.

Hoe meer fixaties de respondenten nodig hadden, des te hoger het aandeel *weet niet* bleek. Dit ging wel op voor de variatie in lay-out, maar niet voor de variatie in betrokkenheid. Het verschil in lay-out had wel invloed op het aantal fixaties en de duur ervan, maar niet op het aantal keren dat de respondent geen duidelijke attitude rapporteerde. In lijn met de *satisficing* verwachtingen is echter wel dat de respondenten bij lage betrokkenheid minder fixaties en kortere fixaties hadden. Bovendien hadden de respondenten bij de lage betrokkenheid een hogere *straightlinescore*. Een andere aanduider van *satisficing*, het percentage *weet niet* bleek echter niet significant.

Als het aantal fixaties dus toeneemt, hoeft dit niet per se te betekenen dat dit ook leidt tot beter overwogen antwoorden. Dit gold wel voor de variatie in betrokkenheid, maar niet voor de variatie in lay-out. Bij dat laatste hadden respondenten meer fixaties en ook langere fixaties, terwijl het aandeel *weet niet* gelijk bleek tussen de verschillende ontwerpen.

4.4 Eye tracking als onderzoeksmethode

De eerste instrumenten waarmee oogbewegingen gevolgd werden, waren complexe constructies die de respondent op het hoofd moest zetten om de oogbewegingen goed te kunnen volgen. Door de jaren heen is geprobeerd de eye tracker minder aanwezig te laten zijn in de onderzoekopstelling. Door de duidelijke aanwezigheid krijgt de proefpersoon het gevoel dat hij/zij in de gaten wordt gehouden en dit leidt tot ander invulgedrag. Zo gingen respondenten uit zichzelf al zorgvuldig te werk (zelfs op de vragenlijst met lage betrokkenheid), omdat wellicht het gevoel heerste dat ze in de gaten werden gehouden. Dit is vaak het probleem in laboratoriumonderzoek; respondenten reageren door de laboratoriumsetting anders op het onderzoeksmateriaal dan wanneer ze ermee geconfronteerd worden in de eigen omgeving. Een ander nadeel van *eye tracking* is dat het veel tijd in beslag neemt. Het verwerken van de data kostte ongeveer 2 uur per proefpersoon. Het uitvoeren van het onderzoek nam ongeveer een half uur in beslag. Hierdoor is het aantal proefpersonen lager dan bij ander vragenlijstonderzoek. Verschillen die werden gevonden tussen beide ontwerpen waren hierdoor niet altijd significant, terwijl ze dat wel konden zijn als dezelfde variatie gevonden werd bij een groter aantal proefpersonen.

Eye tracking biedt wel inzicht in de manier waarop respondenten te werk gaan bij het lezen van de vraag en de rol die variatie in lay-out speelt bij het invullen van een vragenlijst. Bovendien werd aangetoond dat eye tracking zaken boven water kan halen die met zelfrapportage niet gevonden kunnen worden. Eye tracking is een beloftevolle methode voor het testen van de gebruiksvriendelijkheid. Het komt vaak voor dat gebruiksvriendelijkheid getest wordt aan de hand van zelfrapportage. Respondenten dienen dan zelf aan te geven hoe gebruiksvriendelijk een bepaalde vragenlijst was. Hierbij worden de resultaten altijd gestoord door wat de respondent belangrijk vindt aan een vragenlijst en hoe zorgvuldig de proefpersoon zelf rapporteerde dat hij/zij te werk was gegaan. Met eye tracking kan juist de onbewuste zorgvuldigheid in kaart gebracht worden. Daarbij kan, zo is aangetoond in dit onderzoek, aangegeven worden in welke mate de lay-out van invloed was op de mentale inspanning die geleverd werd bij het lezen van de vragen en het geven van een antwoord.

Eye tracking wordt gebruikt om na te gaan waar de respondenten naar kijken en hoe lang of hoeveel fixaties er plaatsvinden. Deze laatste cijfers geven inzicht in hoe lang de respondenten ergens naar kijken. Duidelijk wordt echter niet of dit komt door de onvriendelijkheid van het design of door de mate waarin proefpersonen nadenken over de vragen. Door deze onduidelijkheid is het gebruik van data-triangulatie noodzakelijk als gebruik wordt gemaakt van eye tracking.

4.5 Toekomstig onderzoek

De verschillen in *weet niet* scores bleken niet significant. Met een grotere groep respondenten zou dit verschil wel significant kunnen zijn. Hier is vervolgonderzoek aan te bevelen. Daarnaast zou met een grotere groep personen een vragenlijststudie kunnen worden ingericht om na te gaan in hoeverre de *straightlinescore* die hier is gebruikt een goede maat is voor het aanduiden van *satisficing*. *Straightlining* is hier aangeduid met een maat die is gebaseerd op gezond verstand. De score bleek goed gebruikt te kunnen worden aangezien de gemiddelde *straightlinescore* hoger was voor de vragenlijst met het hoge betrokkenheidsonderwerp. Een hogere score betekende dat er sprake was van een grote afstand tussen vraag x en vraag x-1. Dit ondersteunt de *satisficing* theorie dat respondenten zich minder inspannen bij een minder betrokken onderwerp. Dit zou leiden tot antwoorden die in sterkere mate gebaseerd zijn op het antwoord dat gegeven is op de voorgaande vragen. Dit wordt aangeduid door de *straightlinescores*, die significant verschilden tussen de onderwerpen van de vragenlijst. Hoewel vervolgonderzoek hier zeer welkom is, kan toch gezegd worden dat *straightlinescores* hier als goede maat dienden ter aanduiding van *satisficing* gedragingen, omdat het verwachte verschil tussen de hoge en lage betrokkenheid in satisfice gedrag werd bevestigd door de *straightlinescore*. Hoe lager de betrokkenheid, des te meer satisfice gedrag.

4.6 Tot slot

Aanhangers van het scepticisme en het constructivisme vroegen zich het al af: bestaat de echte werkelijkheid of is dat wat wij met de hersenen ervaren onze eigen 'werkelijkheid'? Het verschil tussen werkelijkheid en waargenomen werkelijkheid in de gedragswetenschap uit zich in de zogeheten meetfouten. *Biased response behavior* is een dergelijke meetfout. Door de aanwezigheid van meetfouten als de *biased response behavior* zal de gerapporteerde attitude nooit overeenkomen met de daadwerkelijke attitude. Plato filosofeerde er al over: "Beauty is in the eye of the beholder". In dit onderzoek door middel van "Tracking the eye of the beholder" hebben we dat bevestigd gezien: de eye tracker zag een andere werkelijkheid dan de werkelijkheid die door de proefpersonen zelf werd ervaren.

5. Referenties

- Ajzen, I., Brown, T. C. & Rosenthal, L. H. (1996). Information bias in contingent valuation: Effects of personal relevance, quality of information, and motivational orientation. *Journal of Environmental Economics and Management*, 30, 43-57.
- Chang, L. & Krosnick, J. A. (2009). National Surveys Via RDD Telephone Interviewing vs. the Internet: Comparing Sample Representativeness and Response Quality. *Public Opinion Quarterly*, 73, 641-678.
- Cialdini, R. (1993). *Influence - The Psychology of Persuasion*. New York: QuLill William Morrow
- Cooke, L. (2004). Improving Usability Through Eye Tracking Research. *IEEE International Professional Communication Conference*, pp. 195-198.
- Cooke, L. (2006). Is Eye Tracking the Next Step in Usability Testing? *IEEE International Professional Communication Conference*, pp. 236-242.
- Couper, M. P., Traugott, M. & Lamias, M. (2001). Web Survey Design and Administration. *Public Opinion Quarterly*, 65, 30-253.
- Couper, M. P. (2008). *Designing Effective Web Surveys*. New York: Cambridge University Press.
- Dillman, D. A. (2007). *Mail and Internet Surveys: The Tailored Design Method*. Hoboken, NJ: Wiley.
- Ganassali, S. (2008). The influence of the Design of Web Survey Questionnaires on the Quality of Responses. *Survey Research Methods*, 2, pp. 21-32.
- Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: Hancock, P. A. & Meshkati, N. (Eds.), *Human Mental Workload*, pp. 139-183. North-Holland: Elsevier Science.
- Holbrook, A. L., Green, M. C., & Krosnick, J. A (2003). Telephone versus Face-to-Face Interviewing of National Probability Samples with Long Questionnaires: Comparisons of Respondent *Satisficing* and Social Desirability Response Bias. *The Public Opinion Quarterly*, 67, 79-125.
- Jäckle, A., Roberts, C. & Lynn, P. (2006). Telephone versus Face-to-Face Interviewing: mode Effects on Data Quality and Likely Causes. Report on Phase II of the ESS-Gallup Mixed Mode Methodology Project. *ISER Working Paper, 2006-41*. Colchester: University of Essex.

- Jacob, R. J. K., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In Radach & Deubel (Eds.). *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, pp. 573-604. Oxford: Elsevier.
- Jemmot, J. B., Ditto, P. H. & Croyle, R. T. (1986). Judging health status: effects of perceived prevalence and personal relevance. *Journal of personality and social psychology*, 50, 899-905.
- Just, M. A. and P. A. Carpenter (1976). Eye Fixations and Cognitive processes. *Cognitive Psychology*, 8, 441-480.
- Krosnick, J. A. (1991). Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys. *Applied Cognitive Psychology*, 5, 213-236.
- Krosnick, J. A., Holbrook, A. L., Berent, M. K., Carson, R. T., Hanemann, M., Kopp, R. J., Mitchell, R. C., Presser, S., Ruud, P. A., Smith, V. K., Moody, W. R., Green, M. C. & Conaway, M. (2002). The Impact of "No Opinion" Response Options on Data Quality: Non-Attitude Reduction or an Invitation to Satisfice? *Public Opinion Quarterly*, 66, 371-403.
- Lozar Manfreda, K., Z. Batagelj, and V. Vehovar. 2002. Design of Web Survey Questionnaires: Three Basic Experiments. *Journal of Computer-Mediated Communication* 7 (3). <http://www.ascusc.org/jcmc/vol7/issue3/vehovar.html> (verkregen op 1 maart, 2010).
- Mittal, B. (1995). A Comparative Analysis of Four Scales of Customer Involvement. *Psychology and involvement*, 12, 663-682.
- Palmer, S. (1992). Common region: A New Principle of Perceptual Grouping. *Cognitive Psychology*, 24, 436-447.
- Poynter, R. (2001). A Guide to Best Practice in Online Quantitative Research. In Westlake, A., Sykes, W., Manners, T. & Rigg, M. (eds). *The challenge of the internet: Proceedings of the ASC International Conference on Survey Research Methods*. London: Association for Survey Computing, pp. 3-19.
- Redline, C. D., & Dillman, D. A. (1999). The Influence of Auxiliary, Symbolic, Numeric, and Verbal Languages on Navigational Compliance in Self-Administered Questionnaires." Paper presented at the International Conference on Survey Nonresponse, Portland, OR.
- Schuman, H., & Presser, S. (1981). *Questions and Answers in Attitude Surveys*. New York: Academic Press.

- Schwarz, N. & Hippler, H. J. (1995). Subsequent Questions May Influence Answers to Preceding Questions in Mail Surveys. *Public Opinion Quarterly*, 59, 93-97.
- Schwarz, N., & Scheming, B. (1992). Frequency Reports of Psychosomatic Symptoms: What Respondents Learn From Response Alternatives. *Zeitschrift fur Klinische Psychologie*, 22, 197-208.
- Smyth, J. D., D. A. Dillman, L. M. Christian, and M. J. Stern. 2006. Comparing Check-all and Forced-choice Question Formats in Web Surveys. *Public Opinion Quarterly* 70, 66-77.
- Sudman, S., Bradburn, N. M. & Schwarz, N. (1996). Thinking about Answers: The Application of Cognitive Processes to Survey Methodology. San Francisco: Jossey-Bass.
- Toepoel, V., Das, M. & van Soest, A. (2005). Design of web Questionnaires: A Test for Number of Items per Screen. Tilburg University: CentERdata Discussion Paper No. 2005-114.
- Tourangeau, R., Couper, M. P. & Conrad, F. G. (2004). Spacing, position, and Order: Interpretive Heuristics for Visual Features of Survey Questions. *Public Opinion Quarterly*, 68, 368-393.
- Tourangeau, R., Rips, L.J., & Rasinski, K. (2000). *The Psychology of Survey Response*. Cambridge: Cambridge University Press.
- Ware, Colin. 2000. Information Visualization: Perception for Design. San Francisco: Morgan Kaufman.
- Wojtowicz, T (2001). Designing Lengthy Internet Questionnaires: Suggestions and Solutions. In Westlake, A., Sykes, W., Manners, T. & Rigg, M. (eds.). *The challenge of the internet: Proceedings of the ASC International Conference on Survey Research Methods*. London: Association for Survey Computing, pp. 25-32.

Bijlage I: Vragenlijsten

Hoge betrokkenheid

Geef aan in hoeverre je het eens/oneens bent met de volgende stellingen (1-5)

Matrix 1: Geld om te leven

1. Als student wil ik zo min mogelijk op mijn geld hoeven te letten.
2. Ik kan goed inschatten welke risico's een studieschuld heeft voor mijn financiële toekomst.
3. Om goed rond te kunnen komen, zal ik wel geld móeten lenen.
4. Het aflossen van een studieschuld tot 5000 euro stelt niets voor.
5. Het idee dat ik een studieschuld van meer dan 5000 euro moet gaan aflossen, beangstigt mij.
6. Vergeleken met mijn niet-studerende leeftijdgenoten vind ik dat ik een luizenleventje heb.
7. Als de basisbeurs wordt omgezet in een lening, zal ik minder geld gaan lenen.
8. Als de basisbeurs wordt omgezet in een lening, verwacht ik meer geld te krijgen van mijn ouders.
9. Als de basisbeurs wordt omgezet in een lening, ga ik meer werken.
10. Ik heb veel over voor het zo laag mogelijk houden van mijn studieschuld.

Matrix 2: Geld om te leven

1. Het grootste deel van mijn studiefinanciering gaat op in de kroeg.
2. Mijn basisbeurs is te laag om van rond te kunnen komen.
3. Ik zou best mijn ov-kaart in willen leveren als ik daarmee een hogere basisbeurs krijg.
4. Ter aanvulling op mijn huidige basisbeurs zou ik een bijbaantje moeten hebben.
5. Naast een lening bovenop de basisbeurs zijn er voor mij genoeg andere manieren om rond te kunnen komen.
6. Ik leer minder goed om te gaan met geld omdat ik studiefinanciering ontvang.
7. Ik heb professionele begeleiding nodig bij het lenen van geld.
8. Als ik zou willen, zou ik alleen van mijn basisbeurs rond kunnen komen.
9. De studiefinanciering vormt mijn voornaamste bron van inkomsten.

Matrix 3: Geld van de overheid

1. Ik vind het terecht dat mensen met een slecht betaalde baan door het afdragen van belasting indirect meebetalen aan mijn studie.
2. Als afgestudeerd academicus zal ik een belangrijke bijdrage gaan leveren aan de maatschappij.
3. De oudere generaties moeten de jongere generaties in staat stellen de studie te kunnen behalen.
4. Het omzetten van de basisbeurs in een lening gaat niet ten koste van de kennismaatschappij.

5. Studenten van ouders met een gezamenlijk jaarinkomen van meer dan € 200.000 zouden geen basisbeurs mogen ontvangen.
6. De samenleving heeft geen profijt van het indirect meebetalen aan mijn studie.
7. Bedrijven die de studie (bachelor of master) van haar medewerkers betalen, mogen de medewerker verplichten 10 jaar bij dat bedrijf te blijven werken.
8. Ik vind het terecht dat als ik halverwege mijn studie stop, ik de reeds ontvangen basisbeurs niet hoeft terug te betalen.
9. Studenten kosten de overheid veel geld.
10. Het omzetten van de basisbeurs in een lening levert de overheid uiteindelijk weinig geld op.

Matrix 4: Geld van de overheid

1. Het omzetten van de basisbeurs in een lening komt de Nederlandse economie ten goede.
2. De overheid kan beter op andere punten bezuinigen dan op de studiefinanciering.
3. Door het huidige studiefinancieringssysteem is studeren voor mij toegankelijker dan destijds voor mijn ouders.
4. Mocht de basisbeurs worden omgezet in een lening, dan is de kenniseconomie het grootste slachtoffer.
5. Ik zou mijn studiefinanciering best willen afstaan aan mensen in probleemwijken.
6. Ik zou het onterecht vinden dat de overheid de kosten van de crisis gaat betalen met het geld dat vrijkomt uit de afschaffing van de studiefinanciering.
7. Ik hoef niet per se studeren om mij te kunnen handhaven in de maatschappij.
8. Ik zou mijn basisbeurs best willen inleveren als de overheid dan de huur van mijn kamer betaalt.
9. Studenten kosten de maatschappij minder dan ze opleveren.
10. Alleen studenten zonder ouders hebben recht op het ontvangen van de basisbeurs.

Matrix 5: Geld en studie

1. Als de basisbeurs zou worden omgezet in een lening, zou ik vaker naar college gaan.
2. Studenten die minder naar college gaan, zouden een lagere basisbeurs moeten ontvangen dan studenten die vaak naar college gaan.
3. Studeren is een investering in jezelf.
4. Door het ontvangen van een basisbeurs voel ik mij verplicht hard te studeren.
5. Studenten die vaker naar college gaan, hebben recht op een hogere basisbeurs.
6. Ik vind het onterecht langer een basisbeurs te ontvangen als ik, door het volgen van andere vakken naast mijn studie, vertraging oploop.
7. Ik word niet tekort gedaan als de basisbeurs wordt omgezet in een lening.
8. Ik besteed minder dan 40 uur per week aan mijn studie.
9. De stof die studenten moeten leren is, gezien de tijd die er voor staat, te veel.

Matrix 6: Geld om te studeren

1. Studenten die een moeilijke studie kiezen hebben recht op een hogere basisbeurs.
2. Ik vind het terecht dat ik, als ik een bestuursjaar doe, voor dat jaar geen recht heb op de basisbeurs.

3. Het is niet erg aannemelijk dat ik met mijn studie stop als tijdens mijn studie de basisbeurs wordt omgezet in een lening.
4. Als ik tijdens mijn studie besluit een jaar te emigreren vind ik het onterecht dat ik het recht op de basisbeurs kwijtraak voor dat jaar.
5. Als de basisbeurs wordt omgezet in een lening, zou ik het logisch vinden dat het collegegeld gehalveerd wordt.
6. Als de basisbeurs wordt omgezet in een lening, word ik door de dreigende studieschuld niet gestimuleerd eerder af te studeren.
7. Zonder studiefinanciering zou ik nooit zijn begonnen met studeren.
8. Ik vind het belangrijker goed rond te kunnen komen dan mijn studie te behalen.

Lage betrokkenheid

Geef aan in hoeverre je het eens bent met de volgende stellingen (1-5)

Matrix 1: Tangibles

1. Een pensioenfonds moet een aantrekkelijk bedrijfspand hebben.
2. Ik vind het belangrijk dat een pensioenfonds modern uitziende faciliteiten heeft.
3. De medewerkers van een pensioenfonds moeten er goed verzorgd uitzien.
4. De documenten rond de dienstverlening (brochures, formulieren) van een pensioenfonds moeten er fraai uitzien.
5. De website van een pensioenfonds moet er fatsoenlijk uitzien.
6. Ik wil niet te veel informatie ontvangen van mijn pensioenfonds.
7. Het logo van een pensioenfonds moet mij aanspreken.
8. De informatie op de website van een pensioenfonds moet duidelijk zijn.
9. De informatie op de website van een pensioenfonds mag niet te veel ingewikkelde termen gebruiken.
10. Ik vind het belangrijk dat een pensioenfonds moeite besteedt aan correcte administratie.

Matrix 2: Tangibles

1. De transacties bij een pensioenfonds moeten veilig zijn.
2. De premie bij een pensioenfonds moet laag zijn.
3. Een pensioenfonds moet een breed assortiment hebben.
4. Een pensioenfonds moet binnen een bepaalde productcategorie veel keuze hebben.
5. Ik vind het belangrijk dat ik het uitkeringsbedrag bij mijn pensioenfonds zelf kan bepalen.
6. De prijs-kwaliteitverhouding van een pensioenfonds moet slecht zijn.
7. Een pensioenfonds hoeft niet goed bereikbaar te zijn.
8. De locatie van de vestigingen van het pensioenfonds moeten dichtbij zijn.
9. Ik verwacht dat een pensioenfonds veel gebruik maakt van e-mail.

Matrix 3: Tangibles

1. Ik wil dat een pensioenfonds gebruik maakt van het internet om contact met mij te onderhouden.

2. Een pensioenfonds moet voorlichtingsavonden organiseren.
3. Ik wil graag het idee hebben dat ik mijn geld afsta aan een uniek pensioenfonds.
4. Ik vind het belangrijk dat een pensioenfonds een dochteronderneming is van een grotere organisatie.
5. De directeur van mijn (aanstaande) pensioenfonds wil ik persoonlijk leren kennen.
6. In de communicatie van een pensioenfonds vind ik het belangrijk dat er leuke kleuren gebruikt worden.
7. Het telefoonnummer van een pensioenfonds hoeft niet gratis te zijn.
8. Het telefoonnummer van een pensioenfonds hoeft niet eenvoudig te onthouden zijn.
9. Als ik met een pensioenfonds bel, wil ik graag steeds iemand anders aan de lijn krijgen.

Matrix 4: Betrouwbaarheid van een pensioenfonds

1. Als een pensioenfonds mij iets belooft te doen, wil ik dat dat daadwerkelijk gebeurt.
2. Ik vind het belangrijk dat een pensioenfonds bij de eerste keer al goede service verleent.
3. Een pensioenfonds moet haar diensten verlenen op het beloofde tijdstip.
4. De medewerkers van een pensioenfonds moeten precies aan kunnen geven wanneer een dienst zal worden verricht.
5. Van een pensioenfonds verwacht ik advies op maat.
6. Een pensioenfonds moet snel gehoor kunnen geven aan mijn verzoeken.
7. Ik wil op eenvoudige wijze inzicht hebben in mijn kosten bij een pensioenfonds.
8. Bij een pensioenfonds wil ik niet voor financiële verrassingen komen te staan.
9. Een pensioenfonds hoeft de gemaakte beloften niet per se na te komen.
10. Een pensioenfonds hoeft niet per se zorgvuldig te zijn in de communicatie.
11. Ik vind het onbelangrijk of een pensioenfonds fraudeert of niet.

Matrix 5: De verantwoordelijkheid van een pensioenfonds

1. De medewerkers van een pensioenfonds moeten accurate service verlenen.
2. De medewerkers van een pensioenfonds moeten bereid zijn mij te helpen.
3. De medewerkers van een pensioenfonds mogen nooit te druk zijn.
4. Het pensioenfonds moet zich inzetten voor mijn belangen.
5. Een pensioenfonds moet regelmatig contact met mij onderhouden.
6. Een goed pensioenfonds doet alles om mij als klant te behouden.
7. De werknemers van een pensioenfonds moeten mij de indruk geven dat er hard gewerkt wordt.
8. Ik wil niet op de hoogte gehouden worden van de ontwikkelingen binnen een pensioenfonds
9. Als ik met een pensioenfonds bel, vind ik het niet erg om lang te wachten.
10. Binnen 24 uur moet een pensioenfonds in staat zijn mij te antwoorden als ik een e-mail stuur.
11. Mijn gegevens hoeven niet per se goed beveiligd zijn bij een pensioenfonds.

Matrix 6: De betrouwbaarheid van een pensioenfonds

1. De medewerkers van een pensioenfonds moeten vertrouwen wekken.

2. Ik verwacht van een pensioenfonds dat de medewerkers over een ruime hoeveelheid kennis met betrekking tot mijn pensioen beschikken.
3. Pensioenfondsmedewerkers moeten altijd beleefd zijn tegen mij.
4. De medewerkers van een pensioenfonds moeten minstens HBO geschoold zijn.
5. Als ik een pensioenfonds kies, mag de organisatie best doen denken aan een spinweb van regeltjes.
6. Een pensioenfonds moet niet de indruk wekken financieel stabiel te zijn.
7. Een pensioenfonds hoeft geen kwaliteit uit te stralen.
8. Een pensioenfonds hoeft niet per se vertrouwen te wekken.
9. Een goed pensioenfonds is een eerlijk pensioenfonds.

Matrix 7: Het inlevingsvermogen van een pensioenfonds

1. Een pensioenfonds moet oprechte belangstelling tonen als er problemen zijn.
2. Medewerkers van een pensioenfonds moeten bereid zijn mij persoonlijke aandacht te geven.
3. Ik wil van het pensioenfonds persoonlijke aandacht krijgen.
4. De medewerkers van een pensioenfonds moeten mijn specifieke behoeften begrijpen.
5. Ik wil mijzelf kunnen herkennen in een pensioenfonds.
6. Ik vind het belangrijk dat bij een pensioenfonds mijn vragen serieus worden genomen.
7. Ik vind het onbelangrijk dat de medewerkers van een pensioenfonds verstand hebben van hun vak.
8. Ik vind het onbelangrijk dat een pensioenfonds bruikbaar is voor mij.
9. Ik hoef bij het kiezen van een pensioenfonds niet te weten waar ik aan toe ben.

Bijlage II: Technische uitleg installatie eye tracking en databestanden

Positioneren van de proefpersoon

De eerste stap in de procedure was het positioneren van de proefpersoon en het instellen van de eye tracking apparatuur en software. Deze installatie vond plaats door het programma FaceLab een foto te laten maken van het hoofd van de proefpersoon. De proefleider gaf vervolgens een referentiekader aan voor de locatie van de ogen. Om dit referentiekader samen te stellen, dienden de ogen, wenkbrauwen, mond en neus aangegeven te worden. Aan dit referentiekader werd door FaceLab een aantal punten toegekend. Deze score is op basis van het gemak dat de eye tracker heeft om de punten die het referentiekader aangeven te kunnen lokaliseren. Bij een puntenaantal hoger dan 80 werd het onderzoek voortgezet. Bij minder dan 80 punten zou er een nieuwe foto en een nieuw referentiekader gemaakt moeten worden. Bij geen van de proefpersonen hoefde deze procedure opnieuw plaats te vinden.

Calibratie en Screen Intersection Display (SID)

Na deze stap vond de calibratie plaats, zodat FaceLab de precieze locatie van de fixaties op het beeldscherm kon bepalen en de afstand tussen de ogen bekend was. Hiervoor werd de proefpersoon gevraagd eerst in de linker- en vervolgens in de rechtercamera te kijken. Hierdoor kon de afstand tussen de ogen nauwkeurig bepaald worden, tot op de tiende millimeter. Na deze calibratie werd de proefpersoon gevraagd een witte knipperende stip te volgen met de ogen, terwijl deze steeds op andere plekken op het scherm verscheen. Dit wordt door FaceLab de zogenaamde SID (Screen Intersection Display) genoemd. Hierdoor is door de eye tracker te achterhalen dat de proefpersoon kijkt waar FaceLab registreert waar de respondent zou moeten kijken. FaceLab geeft per oog aan hoe nauwkeurig de fixaties plaatsvonden op de plek van de witte stip. FaceLab geeft hierbij ook de gemiddelde afwijking per positie van een fixatie ten opzichte van de witte stip aan. Bij een standaard deviatie van kleiner dan 10mm werd besloten het onderzoek voort te zetten. Bij slechts een enkeling moest deze SID opnieuw worden gedaan om zeker te zijn van accurate data.

Eye tracker data

De data die de eye tracker leverde, bestonden per proefpersoon uit een Excelbestand. Deze bestanden verwerken elke fixatie. Per fixatie wordt er door de eye tracker gegeven waar deze fixatie plaatsvond aan de hand van X- en Y-coördinaten. Daarnaast lieten de Excelbestanden zien hoe lang de desbetreffende fixatie had geduurd. Deze tijdsduur werd uitgedrukt in seconden, tot 6 cijfers achter de komma. Van elke fixatie is de plaats bepaald die deze fixatie had op het beeldscherm. Dit gebeurde aan de hand van de X- en Y-coördinaten. Door de waarden van deze coördinaten, kon nauwkeurig bepaald worden of een fixatie plaatsvond in het vraagdeel van een vraag of juist in het antwoorddeel van een bepaalde vraag.

Bijlage III: semantische differentiaalschalen betrokkenheid en inspanning

Betrokkenheid

Studiefinanciering($\alpha=0,87$)/*pensioenfonds* ($\alpha=0,71$)

'Is voor mij belangrijk' versus 'is voor mij onbelangrijk' (1-7)

'Gaaf mij aan' versus 'gaaf mij niet aan' (1-7)

'Boeit mij' versus 'boeit mij niet' (1-7)

'Heeft veel betekenis' versus 'heeft geen betekenis' (1-7)

'Is voor mij van belang' versus 'is voor mij van geen belang' (1-7)

Inspanning

Hoe zorgvuldig hebt u deze vragenlijst ingevuld? (1-7)

Hoe grondig hebt u nagedacht alvorens een antwoord te geven? (1-7)

Hoeveel inspanning hebt u geleverd bij het invullen van de vragenlijst? (1-7)

Hoeveel mentale inspanning kostte het invullen van de vragenlijst? (1-7)

Hoe verveeld was u bij het invullen van deze vragenlijst? (1-7)

Na analyse van deze schaal is het vijfde item verwijderd om tot een betrouwbare schaal te komen. Na verwijdering van het vijfde item bedraagt Cronbach's α 0,81 voor de hoge betrokkenheidsvragenlijst en 0,65 voor de vragenlijst waarbij het lage betrokkenheidsonderwerp gebruikt werd.