

Sylvie Vermeend

Experts en gebruikers inzetten bij het testen van een gepersonaliseerde website?

Een onderzoek naar de waarde van het testen van de usability door zowel gebruikers als experts uitgevoerd op een gepersonaliseerde website.

**Master thesis
Toegepaste Communicatie
Wetenschappen
Universiteit Twente
Faculteit GW
Specialisatie: Nieuwe Media
Technische & Professionele
Communicatie (TPC)
Studentnummer: S0013935**

**Begeleiders:
Marieke Welle Donker-Kuijer
Lex van Velsen**

UNIVERSITEIT TWENTE.

Inhoudsopgave

Hoofdstuk 1 Inleiding	7
Hoofdstuk 2 Theoretisch kader	9
2.1. Personaliseren	9
2.1.1 Personaliseren, definitie	9
2.1.2 Vormen van personaliseren.....	10
2.1.3 Personalisatie, voor- en nadelen	12
2.2 Usability	15
2.3 Testen usability gepersonaliseerde websites.....	16
2.4 Expertonderzoek.....	16
2.5 Gebruikerstesten.....	17
2.5.1 Welke gebruikerstest heeft de voorkeur	17
2.5.2 Hardopdenkmethode	19
2.6 Expertonderzoek versus gebruikersonderzoek	20
2.7 Hoofdvraag en subvragen.....	24
H3. Methodologische verantwoording	26
3.1 Onderzoeksontwerp	26
3.2 Testobject	26
3.3 Respondenten.....	30
3.3.2 Selectieprocedure respondenten.....	30
3.4 Onderzoeksinstrumenten	32
3.4.1 Materiaalgebruik bij gebruikersonderzoek	32
3.4.2 Expertonderzoek	34
3.5 Experimentele procedure	37
3.5.1 Experimentele procedure gebruikers.....	37
3.5.2 Experimentele procedure experts.....	37
3.7 Data analyse	39
Hoofdstuk 4 Resultaten	42
4.1 Analyse	42
4.2 Gebruikersonderzoek.....	44
4.2.1 Pre-interview.....	44
4.2.2 Bijdrage per doelgroep	45
4.2.3 Problemen per pagina	46
4.2.4 Post-interview	48
4.3 Expertonderzoek	48
4.3.1 Pre-interview.....	49
4.3.2 Gevonden problemen vanuit heuristieken	49
4.3.3 Bijdrage per expert	50
4.3.4 Heuristieken Nielsen en Jameson	50
4.3.5 Experts en hoge scores op ernst.....	52
4.4 gebruikers en experts.....	53
4.4.1 Analyse experts en gebruikers	53
4.4.2 Mate van ernst.....	55
4.4.3 N-corrected effect size.....	57
Hoofdstuk 5 Conclusie & discussie	58

Voorwoord

Deze scriptie is geschreven ter afsluiting van de Master Toegepaste Communicatie Wetenschappen aan de Universiteit Twente. Aangezien ik op het moment van het starten met deze scriptie bij een internetbedrijf werkte, sprak mij het onderwerp van deze scriptie direct aan. Ik dacht de kennis die ik tijdens dit onderzoek zou opdoen perfect te kunnen toepassen in mijn dagelijkse werkzaamheden. Tijdens mijn scriptie ben ik van baan gewisseld en ben mij bij een uitgeverij bezig gaan houden met de online marketing. Door de werkzaamheden die ik hier verrichtte kwam het onderwerp van deze scriptie gedurende het onderzoek steeds meer tot leven. Ik zag dat mijn werkgever zowel usability-onderzoek door experts als door gebruikers onder externe verantwoordelijkheid liet uitvoeren, maar dat de problemen die uit deze onderzoeken naar voren kwamen vervolgens niet gebundeld werden tot unieke problemen. Ook bleek dat de waarde van de beide onderzoeken niet op de juiste wijze werd ingeschat. Ten gevolge hiervan is het voor een volgende periode niet vast te stellen of er meer of minder budget voor een bepaald type onderzoek begroot moet worden. Ik wil dan ook graag, na het afronden van deze scriptie, met de in dit onderzoek opgedane kennis in mijn werk aan de slag. Het schrijven van deze scriptie en het uitvoeren van kwalitatief onderzoek is voor mij daarmee erg leerzaam geweest en nuttig voor mijn toekomst.

Ik wil dan ook hierbij graag van de gelegenheid gebruik maken om een aantal mensen te bedanken.

Ten eerste wil ik mijn beide begeleiders, Marieke Welle Donker-Kuijer en Lex van Velsen, bedanken. Zij hebben keer op keer mijn concepten nagekeken en van goede feedback voorzien, zodat ik weer een stapje verder kwam in mijn onderzoek. Ook kwamen zij met goede ideeën waardoor ik mij nog verder in de stof kon verdiepen en vanuit andere invalshoeken en literatuur naar de materie kon kijken. Daarnaast heeft Joyce Karreman, als meelezer, mij met haar feedback erg geholpen.

Ten tweede wil ik mijn vader bedanken voor zijn hulp en steun bij mijn afstuderen. Zijn feedback en aanmoedigen zijn van onmisbare waarde geweest voor mijn eindresultaat.

Daarnaast wil ik Frank Croes bedanken voor zijn enthousiaste feedback. En niet te vergeten wil ik alle gebruikers en experts bedanken die vrijwillig hebben deelgenomen aan mijn onderzoek.

Sylvie Vermeend

Groningen, Augustus 2010

Samenvatting

Dit onderzoek is uitgevoerd om inzicht te krijgen in de waarde van het testen van usability door zowel gebruikers als experts van een gepersonaliseerde website. Uit literatuuronderzoek blijkt dat het vergelijken van beide methodes van onderzoek veelvuldig onderwerp van onderzoek geweest. Er is echter geen onderzoek bekend dat zich specifiek richt op met name de waarde van beide onderzoeksmethodes ten aanzien van gepersonaliseerde websites. De hoofdvraag van dit onderzoek luidt als volgt: *“Wat is de waarde van het testen door gebruikers en het testen door experts bij het verbeteren van de usability van gepersonaliseerde websites?”*. Het onderzoek in deze scriptie richt zich specifiek op de gepersonaliseerde website van Ziggo. Als onderzoeksmethoden worden naast elkaar gezet enerzijds de heuristische evaluatie door experts, waarbij de algemene heuristieken van Nielsen zijn aangevuld met de heuristieken van Jameson (heuristieken specifiek voor personalisatie) en anderzijds als gebruikerstest de hardopdenkmethode aangevuld met een post-interview. Aan het onderzoek hebben 6 experts en 18 gebruikers meegewerkt.

Uit de resultaten van dit onderzoek komt naar voren, dat het toevoegen van heuristieken specifiek gericht op het vinden van problemen verbonden aan personalisatie, in een expertonderzoek geen meerwaarde oplevert ten aanzien van het gevonden aantal problemen, de gemiddelde mate van ernst van de problemen, de thoroughness en de validiteit van deze evaluatiemethode. Verder blijkt dat problemen gevonden door experts niet hoger scoren op de mate van ernst dan de problemen gevonden door gebruikers. Binnen de groep gebruikers dragen de doelgroepen niet in even grote mate bij aan de gevonden echte problemen en verschilt de gemiddelde mate van ernst van de gevonden problemen.

In zijn totaliteit geven de resultaten van dit onderzoek ondersteuning aan de veronderstelling dat de waarde van heuristische evaluatie bij het testen van een gepersonaliseerde website relatief lager is dan bij het testen van een niet gepersonaliseerde website.

Abstract

This study was conducted gain insight into the value of usability testing by both users and experts of a personalized website. Literature study shows that comparing the two methods of research has often been a subject of research. There is no research known that specifically focuses in particular on the value of both research methods regarding personalized websites. The main question of this study is: "What is the value of user testing and testing by experts to improve the usability of personal websites? ". The research in this thesis focuses specifically on the personalized website of Ziggo. As research methods are compared on the one hand heuristic evaluation by experts, in which of the common heuristics of Nielsen are supplemented with heuristics Jameson (heuristics specifically for personalization) and on the other hand as a user test the thinking aloud method supplemented by a post-interview. Six experts and 18 users have participated in the investigation.

The results of this study show that adding heuristics specifically aimed at finding problems associated with personalization in an expert examination adds no value in relation to the number of found problems, the average degree of severity of the problems, the thoroughness and validity of this evaluation method. Furthermore, it appears that problems found by experts don't score higher on the degree of severity of the problems found by users. Within the group of users the target groups don't contribute in the same degree to the real problems found and the average degree of severity of the, the average of the detected problems difference.

In its totality, the results of this study support the hypothesis that the value of heuristic evaluation in the testing of a personalized website is relatively lower than when testing a non-personalized website.

Hoofdstuk 1 Inleiding

In 1920 sprak Henry Ford de legendarische woorden: “Any customer can have a car painted any colour that he wants as long as it is black”. Het was een uitspraak in overeenstemming met de toenmalige verhouding producent consument, een passende marktstrategie (Da Silveria et al., 2001). Nu, bijna een eeuw later, zien we dat de verhoudingen volledig zijn veranderd. De huidige consument is goed opgeleid, heeft toegang tot een rijk arsenaal aan vergelijkende gegevens en overziet via internet de markt vanuit regionaal, landelijk en veelal ook globaal perspectief.

Het document ‘De digitale economie’ (CBS, 2009) illustreert hoe belangrijk internet is: “In 2008 beschikte 88 procent van de Nederlandse huishoudens over een computer. Bijna twaalf miljoen mensen in Nederland maken regelmatig gebruik van internet”. Het CBS (2009) ziet in dit verband eenzelfde trend in het e-shopperen: “In 2009 is het aandeel e-shoppers sterk gegroeid tot 74 procent van de internetgebruikers. Nederland behoort hiermee tot de Europese top.” Identieke bevindingen kunnen worden gevonden in de resultaten uit het onderzoek van Synovate (2008) onder communicatiebureaus. Internet en printreclame eindigden bovenaan als profijtelijke vorm van communiceren met de verwachting van een toenemend belang van internet. Internet is een communicatiemiddel met bijzondere mogelijkheden. Het biedt de ‘zender’ de kans om te kijken of er een ‘ontvanger’ is van de boodschap en de kans om te komen tot interactiviteit (Punselie, 2003).

Interactiviteit is het sleutelbegrip geworden voor succes. De tijd is voorbij dat een leuke website met een gelijke inhoud voor iedereen en zonder mogelijkheid tot interactie, vanzelf verkeer trok. Het web wordt nu ingezet om een-op-een relaties te creëren en de band tussen consument en bedrijven te versterken. Een stevige band vergroot immers de concurrentiepositie, veelvuldig contact biedt de mogelijkheid om nieuwe zaken onder de aandacht te brengen en is een actuele bron van informatie over consumentengedrag. Het is dan ook van het grootste belang dat de consument bij een bezoek aan de site het idee heeft van een persoonlijke benadering, hij moet zich welkom voelen, het moet leiden tot een herhaald bezoek. De content die wordt geboden moet daarom aansluiten bij de context van de individuele webbezoeker. Het zoeken moet eenvoudig en zonder vragen verlopen en de gezochte informatie zal aangepast aan de voorkeuren van het individu moeten worden gepresenteerd. Kortom de website zal ten behoeve van een optimale communicatie optimaal moeten zijn gepersonaliseerd.

Het gaat om een optimale en niet een maximale personalisatie. Optimaal uitgaande van de bedoeling van de website. Het zal duidelijk zijn dat bij websites van non-profit organisaties de mogelijkheden tot personalisatie anders zijn ingevuld dan bij een website met commerciële doelstellingen. Maar om veel bezoekers te trekken en te behouden zal een website ook vooral gebruiksvriendelijk moeten zijn, eenvoudig en intuïtief te gebruiken, in het algemeen aangeduid met usability. Een bezoeker aan een website moet ervaren dat hij zijn doel bereikt, dat dit op een efficiënte manier is gebeurd en dat hij zich tevreden voelt over de manier waarop hij zijn taak heeft uitgevoerd (ISO, 1998).

De voorkeuren en wensen van de bezoekers zijn echter niet stabiel. De technische mogelijkheden worden steeds groter, dus om in verbinding te blijven is het voor beheerders van websites van belang om permanent verbetering na te streven. Websites worden daarom regelmatig getest op usability, gebruiksvriendelijkheid. Dat er gezien de grote belangen zal moeten worden getest staat niet ter discussie. Wel blijft de vraag door wie. Het testen kan gebeuren door experts aan de hand van een lijst van kenmerken waarvan men heeft vastgesteld dat ze van invloed zijn op de gebruiksvriendelijkheid. Maar technisch van belang hoeft niet noodzakelijkerwijs overeen te komen met het belang dat de gebruiker er aan hecht. Het testen kan ook gebeuren door de gebruiker te bevragen. Elke beheerder van een website streeft uiteraard naar een zo groot mogelijk bereik richting de gewenste doelgroep. Het is dus van groot belang om te weten welke test welke informatie op kan leveren. Bij een gepersonaliseerde website speelt deze vraag des te meer daar de website per definitie niet voor elke gebruiker een identiek beeld geeft. In dit onderzoek zal worden onderzocht wat de waarde is van het inzetten van experts c.q. gebruikers bij het verbeteren van de usability van gepersonaliseerde websites.

Hoofdstuk 2 Theoretisch kader

In dit hoofdstuk wordt personalisatie gedefinieerd en in een breder perspectief geplaatst van de huidige tijd. Personalisatie van websites wordt gezien als een oplossing om de kwaliteit van de communicatie via websites te verbeteren. Het belang van websites als communicatiemiddel neemt toe en daarmee het belang van personalisatie. Om bezoekers te trekken zullen ook gepersonaliseerde websites moeten voldoen aan de wensen ten aanzien van gebruiksvriendelijkheid, usability. De concurrentie op het web is groot, het is dus voor elke beheerder van belang om de kwaliteit van de site te monitoren. Het testen van gepersonaliseerde websites kan met de inzet van experts maar ook met de inzet van gebruikers. Bovendien zijn er verschillende methodes. Dit leidt tot de onderzoeksvraag en bijbehorende subvragen.

2.1. Personaliseren

2.1.1 Personaliseren, definitie

Personaliseren komt als woord van oorsprong uit het Frans en betekent zoiets als 'een persoonlijk tintje geven'. Zo gaf Graficus, het oudste Nederlandse weekblad voor de grafische industrie, in 1995 de gepersonaliseerde roman *Moord op de Drupa* uit als cadeau voor zijn abonnees. De lezers zelf speelden in het boekje de hoofdrol tenminste als zij hun persoonlijke gegevens, zoals naam, adres, automerk, favoriete eten en vakantiebestemming, beschikbaar hadden gesteld. Dat was 1995. Nu, in 2010 biedt Nokia op haar website haar klanten de kans om zelf het design voor de telefoon te ontwerpen. Kopers van de Smart kunnen de buitenkant van hun auto aan laten passen aan hun wensen. Voor de consument lijkt het of de producten speciaal op maat worden gemaakt, terwijl er sprake is van massaproductie, ook wel bekend onder de naam mass customization (Davis, 1993). Beide producenten geven op hun eigen manier antwoord op de vraag naar maatwerk. In het geval van Nokia is er sprake van een interactief systeem dat vanuit opgedane kennis uit eerdere contacten met de koper, de koper optimaal ondersteunt om tot het gewenste resultaat te komen. In het geval van de Smart kan de klant kiezen uit beschikbare mogelijkheden, het systeem ondersteunt niet. Personalisering van sites en productpersonalisering zijn twee uitingen van dezelfde ontwikkeling, een toenemende individualisering. Dit past binnen de trend van Customer Relationship Management waarbij bedrijven klantcontacten optimaliseren en inspringen op de wensen van de individu (Kincaid, 2003).

Zaphiris et al. (2007) geven aan dat gepersonaliseerde websites de inhoud, de indeling van de inhoud en de presentatie van informatie aanpassen aan de voorkeuren, behoeften, interesses op basis van opgeslagen gegevens vanuit eerdere bezoeken aan de website en het doel van de gebruiker. Kobsa et al (2001) definiëren een gepersonaliseerde website als een systeem dat zich aanpast aan de kenmerken, het gedrag en de omgeving van de individuele gebruiker.

2.1.2 Vormen van personaliseren

Personaliseren gaat om een oplossing voor het individu dat dagelijks overspoeld wordt met oneindig veel informatie en geholpen moet worden om zijn weg daarin te vinden. Bij het helpen gaat het om twee hoofdlijnen.

Ten eerste zijn er sites waar bezoekers slechts incidenteel komen zoals bijvoorbeeld de site van een museum. Omdat herhaald bezoek niet wordt verwacht en dus investeren in het aangaan van een binding geen doel is, ligt personaliseren bij deze sites minder voor de hand. Aandacht voor een goede navigatie zou hier prioriteit moeten krijgen.

Ten tweede zijn er sites waar een bezoeker gezien het soort aanbod vaker gebruik van maakt bijvoorbeeld de site van Wehkamp. Bij dergelijke sites kan personaliseren een uitkomst bieden.

Er komen nu twee vraagstukken naar voren. Ten eerste hoe komt de beheerder van een website aan de benodigde informatie over de bezoeker? En ten tweede, wie heeft de controle over de veranderingen in het systeem? Goh en Foo (2008) beschrijven in hun boek 'Social information retrieval systems' twee werkwijzen waarop personalisatie van sites, het verkrijgen van informatie, tot stand kan komen. Ten eerste 'Explicit Personalization' waarin de gebruiker nadrukkelijk gevraagd wordt om gegevens in te vullen in een format ten behoeve van een persoonlijk profiel. Het gaat minimaal om het e-mailadres om zo communicatie mogelijk te maken. De informatie voor de personalisatie komt dus in dit geval van de gebruiker, de beheerder van de site interpreteert. Ten tweede wordt 'Implicit Personalization' genoemd. Het profiel van de gebruiker wordt in dit geval samengesteld op basis van verkregen informatie uit cookies. Daarbij kan onder andere gedacht worden aan klikken met de muis, gebruik van het toetsenbord, het gebruik van de functie scrollen. "The main purpose of it is to identify users and possibly prepare customized Web pages for them. Cookies are used for storing a user's ID en password for specific Web sites. Each time the user logs on to the Web site, the browser will check to see if the user has any pre-defined preferences (via cookies)" (Goh en Foo, 2008) De gebruiker wordt op basis van de geconstateerde voorkeuren door de beheerder van de website gekoppeld aan een referentiegroep. Deze koppeling bepaalt de manier waarop de website reageert op de inbreng van de gebruiker (collaboration).

Naast de expliciete en impliciete benadering is er nog sprake van een vorm waarin beide gecombineerd worden. Dit wordt aangeduid als 'hybrid' personaliseren.

Bezien vanuit de vraag "Wie heeft de zeggenschap over de veranderingen in het systeem?" kan een onderscheid gemaakt worden in twee vormen, adaptable en adaptive, zoals weergegeven door Karwowski (2006) in tabel 1.

Tabel 1 **Adaptable en adaptive systeem van Karwowski (2006)**

	adaptable system	adaptive system
presentatie van de informatie	gebruiker verandert format informatie	gebruiker en systeem veranderen format informatie
functionaliteit	gebruiker beheerst de functies	gebruiker en systeem beheersen de functies

Bij een adaptable systeem kunnen interface en content enkel aangepast worden door de gebruiker, bij een adaptive systeem beschikken gebruiker en systeem beide over deze mogelijkheid.

Beide systemen vallen onder de paraplu van gepersonaliseerde systemen (Van Velsen, 2010). Bij een adaptable systeem ligt de mogelijkheden tot sturen enkel bij de gebruiker. De gebruiker kiest zonder ondersteuning uit de aangeboden mogelijkheden. Bij een adaptive systeem komt er een wisselwerking tot stand tussen de informatie die bekend wordt of al is verkregen van de gebruiker en de reactie daarop van het systeem. De indeling is van belang omdat met name het gebruik van adaptive systemen vragen oproept over risico's van sturen, van niet gewenst gebruik van informatie zoals beschreven door Jameson (2007) en Van Velsen (2010). In paragraaf 2.1.3 wordt hier nader op ingegaan.

Er bestaan verschillende vormen van personalisatie. Wu et al (2003) hebben de meest bekende vormen geïdentificeerd. Dit zijn 'control personalization', 'content personalization', 'link personalization', 'customized screen personalization' en 'anthropomorphic personalization'. Deze vormen worden kort aangestipt.

Er wordt gesproken over 'control personalization' als de vraag centraal staat in hoeverre de gebruiker zelf de controle heeft over het proces van personalisatie. Daarbij staan vragen centraal als; "Kan de gebruiker sturen op de manier en de diepgang van het inwinnen van informatie bij de expliciete personalisatie?" en "Kan de gebruiker zelf invloed uitoefenen op de informatie die impliciet wordt verzameld of al verzameld is?" (Wu et al.,2003).

Er is sprake van 'content personalization' wanneer de inhoud wordt geselecteerd met behulp van filtertechnieken, zoals 'contextual targeting' en 'behavioral targeting'. Bij de eerste variant wordt de inhoud van de webpagina, de content, getoond op basis van de context van de betreffende pagina. Een bekend voorbeeld hiervan is 'google ad sense' (Rossi et al., 2001). Contextual targeting wordt gewoonlijk gebruikt voor expliciet gepersonaliseerde content. Bij behavioral targeting wordt de inhoud getoond op basis van het online bezoekersgedrag. Daarbij kan gedacht worden aan de clicks die een bezoeker maakt op de website, de bestanden die hij download, de zoekopdrachten en de online herkomst van de bezoeker, het IP-adres. Deze filtertechniek wordt gebruikt voor impliciet gepersonaliseerde content (Anfuso, 2006).

Een andere vorm van personalisatie is 'link personalization'. Het personaliseren van links kan zowel impliciet als expliciet zijn (Rossi et al.,2001) Voorbeelden van expliciete gepersonaliseerde links zijn "favorieten" in webbrowsers of de links in een "My" portal. De gebruiker kan zelf bepalen welke links hij/zij opslaat. Links die door de website zelf worden aangedragen op basis van een zoekopdracht, zijn een voorbeeld van impliciete personalisatie.

Wanneer een bezoeker de look en feel van een website naar eigen wens kan aanpassen, is er sprake van 'customized screen personalization' (Wu et. al., 2003).

Van 'anthropomorphic personalization' wordt gesproken wanneer een website zich voordoet alsof het een mens is. De bezoeker van de website wordt bijvoorbeeld begroet met 'Hallo Jan, hoe gaat het met je?' (Sadasivam, 2008).

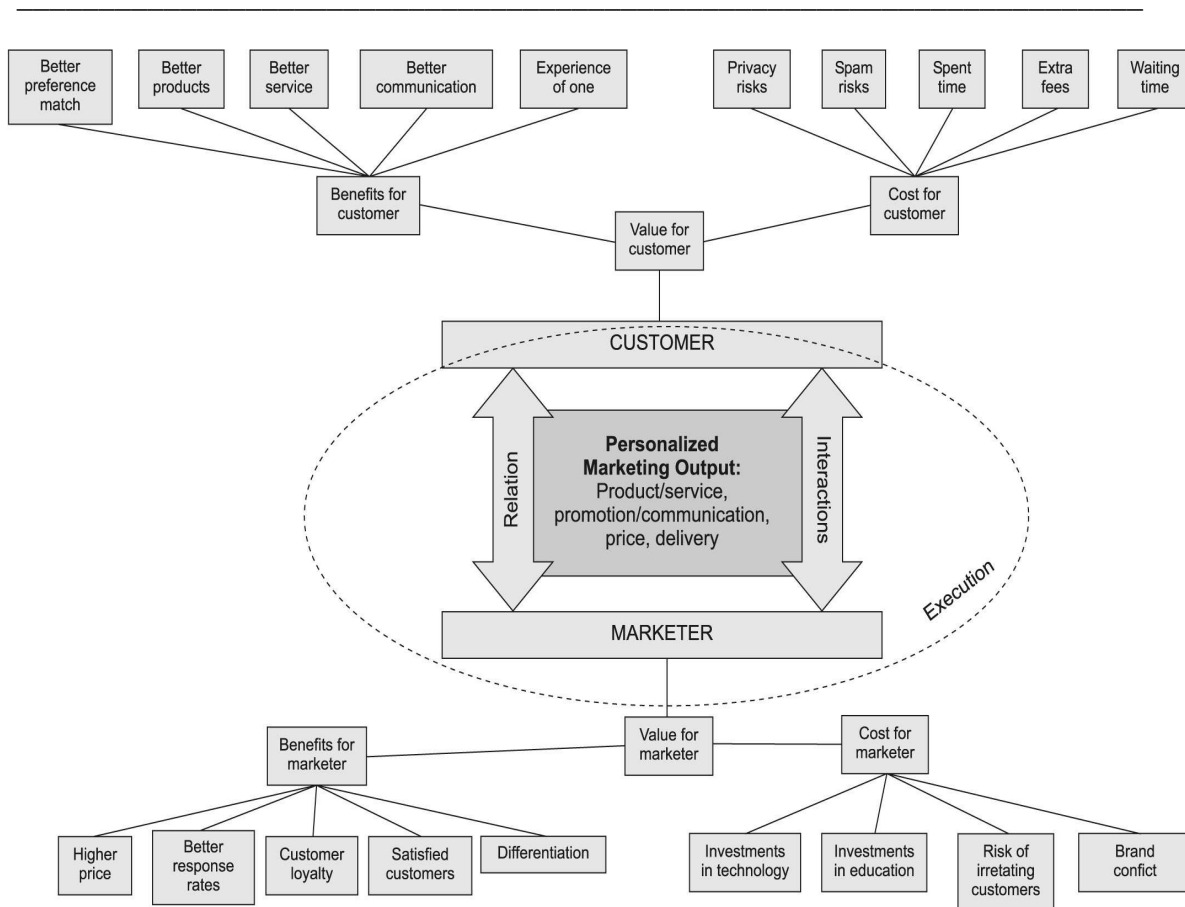
In het kader van dit usability onderzoek naar een gepersonaliseerde website richt ik mij op 'content personalization'. De reden hiervoor is dat content één van de belangrijkste items is op websites om te personaliseren (Wu et. al, 2003).

2.1.3 Personalisatie, voor- en nadelen

Waarom zijn individuen bereid om persoonlijke gegevens kenbaar te maken op een onpersoonlijke website? Waarom zijn de bedrijven en organisaties die sites inzetten als communicatiemiddel bereid om extra in dit medium te investeren door er voor elke gebruiker een persoonlijk profiel van te maken? Klaarblijkelijk verwachten gebruiker en eigenaar van de site beiden zo veel voordelen dat ze bereid zijn om hun aandeel in de investering te doen.

De een in de vorm van de bereidheid om actief persoonlijke informatie te leveren dan wel toe te staan dat er door in te loggen informatie via zijn / haar pc door de beheerder van de site wordt ingewonnen, de ander door te investeren in apparatuur, besturingssoftware, databases inclusief een bijbehorende organisatie. Het overzicht door Versanen (2007) biedt inzicht in de 'cost' en 'benefits' voor zowel de gebruiker als de beheerder van de website en de verhouding waarin ze tot elkaar staan.

Figuur 1 Voordelen en nadelen gebruiker en beheerder websites (Versanen, 2007)



Het overzicht plaatst personalisatie van websites in een breed perspectief. Voor bedrijven wordt het economisch belang in beeld gebracht door enerzijds de voordelen te benoemen (benefits for marketer) en anderzijds aan te geven dat het ook een kostenpost vormt (cost for marketer). Het in de lucht houden van een website is een zakelijke afweging en daarbij zullen ook de kosten voor onderzoeken naar usability een rol spelen. Voor de gebruikers geldt niet zozeer een zakelijke afweging als wel de afweging van onder andere het inleveren van tijd en privacy (cost customer) tegenover een meer op maat toegesneden dienstverlening (benefits for customer).

De punten aangeduid met 'nadelen voor gebruikers' (cost) kunnen ook worden gezien als mogelijke risico's, risico's die men liever niet loopt. Kobsa (2007) constateert dat gebruikers van internet het belangrijk vinden om te weten wat er met hun persoonlijke informatie wordt gedaan en hoe zij hier de controle over kunnen houden. In het onderzoek (Kobsa,2007) geeft 63% van de ondervraagden aan dat zij onjuiste informatie heeft verstrekt, maar wel bereid is om juiste informatie te verstrekken als men weet wat er mee gebeurt en men daar mee in kan stemmen. Jameson (2007) benoemt eveneens risico's van personalisatie maar beschrijft deze risico's vervolgens als eisen die, naast de algemene richtlijnen behorend bij usability, specifiek gesteld zouden moeten worden aan usability van met name gepersonaliseerde systemen. Voor een nadere invulling wordt verwezen naar tabel 2. Deze tabel onderstreept het belang van usability van met name gepersonaliseerde websites.

Tabel 2 **Eisen Jameson t.a.v. usability gepersonaliseerde website (Jameson, 2007)**

Eisen Jameson t.a.v. usability gepersonaliseerde website	
Voorspelbaarheid	Gebruikers moeten in staat zijn om te voorspellen wat de bijdrage zal zijn van hun acties aan het te realiseren resultaat.
Begrijpelijkheid	Gebruikers moeten in staat zijn om te begrijpen hoe het systeem werkt ten aanzien van hun profilering en ten aanzien van het op hen toesnijden van informatie.
Controleerbaarheid	Gebruikers moeten in staat zijn om controle uit te oefenen op hun profilering en het genereren van de gepersonaliseerde output.
Hinderlijkheid	Gebruikers moeten in staat zijn om de taak af te maken zonder daarbij gehinderd te worden door bijwerkingen van de personalisatie.
Privacy	Gebruikers moeten niet het gevoel hebben dat de profilering niet in overeenstemming is met hun behoefte aan privacy.
Het missen van informatie	Het vastgelegde profiel mag er niet toe leiden dat de gebruikers de mogelijkheid wordt ontnomen om wat nieuws aangeboden te krijgen omdat dit niet past in dit van hen vastgelegde profiel.
Systeembevoegdheid	De gebruiker moet niet de indruk krijgen dat het systeem werkt met een onjuist profiel of output niet herkenbaar personaliseert.

2.2 Usability

Er zijn meerdere definities van usability maar in alle beschrijvingen komen de termen effectiviteit, efficiëntie en tevredenheid van de gebruiker voor. Soms is het perspectief van een definitie breder dan alleen websites en gaat het ook over producten in het algemeen zoals bijvoorbeeld de definitie van ISO 9241 (1998). Omdat het in deze scriptie om websites gaat wordt er gekozen voor de definitie van Van der Put (2006), een specificatie van de ISO-definitie; “Website-usability is de mate waarin een website, webpagina of een ander onderdeel daarvan kan worden benut door specifieke gebruikers om specifieke doelen effectief, efficiënt en naar tevredenheid te bereiken in een specifieke gebruikerscontext.”

In de definitie van Van der Put (2006) is drie maal het woord ‘specifiek’ opgenomen en zoals Van der Put (2006) schrijft, “heel bewust”. “Bij het ontwerpen van websites en webpagina’s moet worden uitgegaan van specifieke gebruikers met specifieke doelen in een specifieke gebruikerscontext.” Met andere woorden het gaat om de individuele bezoeker met zijn eigen wensen en vanuit de eigen context. Sites die hier geautomatiseerd rekening mee houden, heten gepersonaliseerd en dus is de genoemde definitie ook van toepassing op usability van gepersonaliseerde sites.

De beschreven definitie bevat echter nog een aantal termen die om een nadere concretisering vragen zeker als het gaat om het testen van usability. Zo wordt onder effectiviteit de nauwkeurigheid en volledigheid verstaan waarmee gebruikers gespecificeerde doelen bereiken (ISO, 1998). Als iemand bijvoorbeeld online een koelkast wil bestellen dan is zijn handeling alleen effectief als het inderdaad lukt om de koelkast te bestellen en te betalen. Als het bestelproces mislukt, dan is het bezoek van de gebruiker aan de site niet effectief. Bij het begrip efficiënt gaat het om de inspanning die de gebruiker moet verrichten om het doel te bereiken. Uiteindelijk gaat het om een maximaal resultaat met een zo gering mogelijke inspanning. Stel dat bij elk bezoek aan dezelfde webwinkel telkens opnieuw alle NAW-gegevens moet worden ingevoerd, dan is dat niet efficiënt. Tevredenheid is zonder meer het meest subjectieve begrip uit de definitie. Het gaat erom of de persoon de handeling op de website als prettig ervaart en of hij geen problemen ondervindt met het uitvoeren van zijn taken of zoals het formeel wordt omschreven; het moet leiden tot vrijwaring van ongemak en een positieve houdingen ten opzichte van het product (ISO, 1998). Tot slot wordt bij gebruikerscontext in de meeste gevallen aan een browser of een apparaat zoals een pc, Macintosh, pda, mobiele telefoon, enzovoort gedacht.

“Het web legt werkelijk alle macht bij de consument. Degene die met de muis klikt, beslist over alles. Het is zo eenvoudig om ergens anders heen te gaan; alle concurrenten over de hele wereld zijn slechts een muisklik verwijderd ” (Nielsen, 2001). De tijd is voorbij dat de klant eerst wat kocht, betaalde en vervolgens thuis ontdekte dat het toch niet zo werkte als hij had verwacht. Bij het bezoek aan een website stellen klanten nu eerst vast of deze voor hen functioneel is, datgene biedt dat ze ervan verwachten, voordat ze de site werkelijk gaan gebruiken. Usability is dus van belang voor de tevredenheid van de gebruiker van de site en speelt een belangrijke rol in een al of niet herhaald bezoek. Het bindend vermogen van de site ten opzichte van de gebruiker bepaalt mede de concurrentiekracht en kan zo rechtstreeks gevolgen hebben voor het economisch rendement van een bedrijf. In het binden van de gebruiker kan, zoals eerder besproken, personalisatie een belangrijke rol spelen.

2.3 Testen usability gepersonaliseerde websites

De usability van websites is op vele manieren te testen. Kuniavsky (2003) geeft aan dat er ongeveer vijftig methoden en technieken zijn voor het testen van de usability van websites. Volgens Sherman (2005) zijn daar tien technieken bij die veel gebruikt worden door meer dan de helft van alle usabilityprofessionals. De techniek die het meest gebruikt wordt voor het testen van de usability is de heuristische evaluatie, een vorm van expertonderzoek (Nielsen, 1993). In paragraaf 2.4 wordt verder in gegaan op expertonderzoek. Een andere vorm van het testen van usability die veel gebruikt wordt is gebruikersonderzoek (Sherman, 2005). Hierbij wordt een gebruikerstest uitgevoerd met een proefpersoon. Hier zijn meerdere vormen van, in paragraaf 2.5 wordt verder in gegaan op gebruikersonderzoek.

2.4 Expertonderzoek

Bij het werken met experts komen verschillende methodes naar voren. Er kan gewerkt worden met heuristieken (Nielsen, 1994), een pluralistic walkthrough of een cognitive walkthrough (Hollingshed & Novick, 2007) en gedrag kan op verschillende manieren geobserveerd worden. De verschillende methoden worden kort aangestipt.

Een heuristische evaluatie is een vorm van een expert-review, een analyse en beoordeling door een expert (Van der Put, 2006). Bij deze analyse wordt gebruik gemaakt van heuristieken. Heuristieken zijn lijsten met evaluatiecriteria waar een goede website aan zou moeten voldoen. Ze worden geformuleerd in de vorm van ijkpunten. Heuristieken, checklijsten, kunnen betrekking hebben op verschillende aspecten van websites, zoals de gebruiksvriendelijkheid, navigatiestructuur en visueel design. De evaluatie vindt plaats door de website te beoordelen aan de hand van deze lijsten met criteria.

Een 'pluralistic walkthrough' is een bijeenkomst waar zowel gebruikers als ontwikkelaars en experts met behulp van taak scenario's samen de website doorlopen. De ontwikkelaars en experts bekijken de website vanuit de rol van de typische gebruiker (Nielsen, 1994).

Bij een 'cognitive walkthrough' lopen experts met behulp van taken de website door. Terwijl zij dit doen, verplaatsen zij zich in de gedachten van gebruikers die voor het eerst de website bezoeken (Hollingshed & Novick, 2007). Om enig gevoel te krijgen bij dit laatste wordt experts aangeraden om gebruikers tijdens een usability test te observeren. Hierdoor kan veel informatie worden verkregen over de problemen waar een gebruiker tegenaan loopt bij het bezoeken van een website (Hollingshed & Novick, 2007).

Er is gekozen om gebruik te maken van de heuristieken omdat het een goed instrument is als een onderzoek in een kort tijdbestek moet plaatsvinden en er weinig financiële middelen zijn, zoals bij dit onderzoek het geval is. Geschoolde experts kunnen in een kort tijdbestek kwalitatief waardevolle bevindingen doen (Kantner & Rosenbaum, 1997).

Het onderzoek dat onderdeel uitmaakt van deze scriptie richt zich op usability van met name gepersonaliseerde websites. Daarom worden aan de gebruikelijke heuristieken (Nielsen & Mack, 1994), de speciale heuristieken passend bij personaliseren (Jameson, 2009) toegevoegd. De uitwerkingen van deze heuristieken zoals beschreven in paragraaf 2.1.3 kunnen worden gevonden in bijlage 8. Om inzicht te krijgen in de bijdrage van beide lijsten van heuristieken aan de evaluatie door experts worden ze afzonderlijk in de analyse verwerkt.

2.5 Gebruikerstesten

2.5.1 Welke gebruikerstest heeft de voorkeur

Bij de keuze voor de methode van gebruikersonderzoek die het beste paste bij het onderzoek naar usability van gepersonaliseerde websites is gebruik gemaakt van de literatuurstudie van Van Velsen (Van Velsen et al., 2008). Deze studie biedt o.a. een overzicht van de frequentie waarmee bij in de literatuur gevonden onderzoek naar gebruikerstesten, een aantal evaluatie methodes wordt ingezet ten behoeve van het meten van verschillende variabelen.

Tabel 3 Methods used for UCE in the collected studies (n = 63)
(Van Velsen User-centered evaluation of adaptive and adaptable systems)

Evaluation method or instrument	Times applied	Variables most frequent assessed
Questionnaires	47	Usability (21) Perceived usefulness (19) Intention to use (14)
Interviews	27	Intention to use (9) Appopriateness of adaption (7) Usability (7)
Data log analysis	24	User behavior (17) User performance (6)
Focus groups & group discussions	8	Intention to use (4) Perceived usefulness (2) Trust & privacy issues (2)
Think-aloud protocols	7	Usability (2) User behavior (2)
Expert reviews	6	Usability (3)

Bezien vanuit dit onderzoek zijn de methodes van belang die gebruikt zijn om usability van met name gepersonaliseerde websites te meten. In algemene zin worden met name vragenlijsten, interviews, de hardopdenkmethode en experts reviews genoemd. Van Velsen et al. (2008) gaat vervolgens in op de mogelijkheden die er zijn ten aanzien van het vaststellen van problemen bij usability van met name gepersonaliseerde websites. Zo wordt van vragenlijsten opgemerkt dat deze vooral geschikt zijn voor kwantitatieve metingen ten aanzien van usability waarbij gedacht kan worden aan onderzoek naar tevredenheid over de personalisatieopties, of algemene tevredenheid. Het zijn onderzoeken die worden verricht vanuit behoefte aan gegevens voor benchmark.

Interviews worden door de mogelijkheid van verdieping in de vraagstelling wel geschikt geacht in de zoektocht naar problemen met de usability van gepersonaliseerde websites. Interviews worden echter vooraf of na afloop afgenomen, wat ervoor kan zorgen dat de informatie mogelijk enigszins vertekend werd weergegeven.

Over de hardopdenkmethode wordt opgemerkt dat deze geschikt is voor het vaststellen van de problemen bij usability van gepersonaliseerde websites, omdat de gebruiker zelf aan het woord is en direct antwoord geeft aan wat er wordt ondervonden. Het is wel van belang om de omstandigheden en werkwijze nauwkeurig vast te leggen en er voor te zorgen dat de te testen gebruiker niet geheel opgaat in het doel dat er bereikt moet worden met het bezoek aan de site. Het ligt hiermee voor de hand dat in het onderzoek naar de usability van gepersonaliseerde websites de hardopdenkmethode wordt ingezet als gebruikerstest.

In het onderzoek van Van Velsen (2010) komt naar voren dat Jameson (2008) stelt dat thinking-aloud weliswaar de beste methode is voor het opsporen van ernstige fouten maar dat dit niet geldt voor fouten behorend bij items als privacy en voorspelbaarheid. Deelnemers aan onderzoek lijken enkel hun opvatting over deze items te uiten als ze daar gericht op worden bevraagd, wat leidt tot het inzetten van vragenlijsten of een interview.

Op basis van deze opmerking zal in het onderzoek bij deze scriptie als gebruikerstest gebruik worden gemaakt van de hardopdenkmethode gevolgd door een post-interview in de vorm van een vragenlijst gericht op de specifieke personalisatie items van Jameson (2009).

2.5.2 Hardopdenkmethode

Omdat de hardopdenkmethode wordt toegepast in het onderzoek als onderdeel van de scriptie wordt er kort ingegaan over de in te zetten vorm. Het is een methode waarbij de proefpersoon moet verwoorden wat hij of zij doet en denkt. Dat kan op twee momenten. Alle problemen worden mondeling verwoord tijdens het uitvoeren van de taken (concurrent think aloud) of na het uitvoeren van de taken tijdens het terugzien van de opname van hun handelingen (retrospective think aloud). Bij het inzetten van deze methode wordt de onderzoeksessie vastgelegd met bijvoorbeeld een videocamera. Bij beide uitvoeringen van de hardopdenkmethode is er een onderzoeker aanwezig die eventueel notuleert, de afname controleert en begeleidt. Hoewel 'concurrent think aloud' en 'retrospective think aloud' verschillende methoden zijn, lijken ze een vergelijkbaar resultaat te geven. Het zijn beide goede methoden, waarbij het aantal en typen geconstateerde problemen gelijk is (van der Haak & de Jong, 2003).

Bij personalisatie draait het om het aanpassen van inhoud of interface op het moment dat het van toepassing is. Dit kan zowel door middel van 'concurrent think aloud' als 'retrospective think aloud' bekeken worden. Beide methodes zijn geschikt om te gebruiken. Er is bij deze scriptie echter om praktische redenen voor 'concurrent think aloud' gekozen. De 'retrospective think aloud' is, doordat er twee verschillende opnames gemaakt moeten worden, een veel lastiger methode om uit te voeren. Daarnaast is 'concurrent think aloud' minder belastend voor de gebruiker, het kost minder tijd en moeite.

2.6 Expertonderzoek versus gebruikersonderzoek

2.6.1 Vergelijken

Het hoofddoel van het vergelijken van evaluatiemethoden is om vast te stellen welke het beste is, in dit geval het beste in het vinden van usability problemen voor een gepersonaliseerde website. Indien één methode voldoende informatie voor het gewenste doel oplevert dan kan er immers tijd en geld worden bespaard. Expertonderzoek met behulp van heuristieken en gebruikersonderzoek zijn twee algemeen erkende methodes voor het om vinden van usability problemen en het is dan ook logisch dat deze met elkaar worden vergeleken (de Kock et al, 2009). Het vergelijken van beide methodes is veelvuldig onderwerp van onderzoek geweest, soms gericht op enkelvoudige items als bijvoorbeeld effectiviteit soms als literatuurstudie waarin de resultaten van meerdere onderzoeken met elkaar werden vergeleken. Onder andere Batra en Bishu (2007) en Tan et al. (2009) concluderen dat beide methodes verschillende soorten fouten aan het licht brengen, dat de overlap in gevonden fouten niet bijzonder groot is en dat testen met gebruikers en de heuristische evaluatie gericht zijn op verschillende usability-problemen en wel zo significant dat beide methodes als complementair kunnen worden beschouwd.

Er is geen reden om te veronderstellen dat de uitspraken over het complementair zijn niet ook zouden gelden voor gepersonaliseerde websites. Het zou wel mogelijk kunnen zijn dat de verhouding in de bijdragen van beide methodes zou kunnen wijzigen. De set in te zetten heuristieken van de experts is immers uitgebreid met de specifieke heuristieken behorend bij personaliseren (Jameson, 2009) wat mogelijk leidt tot het vinden van meer problemen door de experts. Dit zouden vooral om meer overlappende problemen met de gebruikers gaan. Dit mogelijk effect zal in deze scriptie worden onderzocht door de gevonden problemen te onderscheiden naar de bijbehorende sets van heuristieken, de heuristieken van Nielsen en Mack (1993) en de heuristieken van Jameson (2009).

Voor de gebruikers geldt niet dat de criteria waarmee zij de usability van de gepersonaliseerde website beoordelen veranderen. De opdrachten zijn echter nadrukkelijk gericht op met name de gepersonaliseerde eigenschappen. Er bestaat geen inzicht of dit al of niet zal leiden tot een positief of negatief effect op het aantal en de ernst van de gevonden problemen. Dit vraagstuk maakt geen deel uit van het onderzoek van deze scriptie.

Om een beeld te krijgen of er een mogelijke wijziging optreedt in de verhouding tussen beide methodes van onderzoek, zal de te verwachten verhouding worden vergeleken met de verhouding berekend op basis van de resultaten van het onderzoek. De te verwachten verhouding wordt weergegeven met gebruik van een score op thoroughness voor het expertonderzoek met 6 experts en het gebruikersonderzoek met 18 gebruikers waarbij de scores kunnen worden afgelezen uit een grafiek. Hiervoor wordt de meta-analytische benadering van Hwang and Salvendy (2009) gebruikt.

2.6.2 Criteria

Hornbaek (2010) beschrijft in zijn literatuurstudie naar onderzoeken waarin methodes met elkaar werden vergeleken uitgevoerd tussen 1999 en 2004, dat het verloop van deze onderzoeken een vast verloop kende. Er worden op een website meerdere evaluatiemethodes ingezet wat leidt tot even zoveel lijsten met problemen. De problemen worden gematched en meestal voorzien van een code voor de mate van ernst. De analyse is gericht op het aantal unieke problemen per methode en de overlap. Vervolgens wordt aan de hand van dogma's helder waarom deze benadering niet voldoet. Hij stelt onder andere vast dat de ingebrachte problemen niet altijd usability-problemen zijn, dat elk probleem even zwaar telt en dat gebruiker en expert vanuit een verschillende context werken waardoor belangrijke informatie verloren gaat door het aanleveren van een simpel getal en een omschrijving. Daarnaast wordt gesteld dat meerdere geconstateerde problemen vaak onder één noemer worden gebracht en procedures om problemen vast te stellen, niet eenduidig zijn. Vanwege deze constatering wordt in deze scriptie de aanpak van het eenvoudig tellen van de problemen en rubriceren naar de mate van ernst aangevuld met kwalitatieve metingen en een verbinding met resultaten van eerder verricht onderzoek.

Om inzicht te krijgen in de mate waarin de problemen gevonden door experts verbonden zijn met de heuristieken van Nielsen en Mack (1994) en Jameson (2007) wordt gebruik gemaakt van de benadering van Ling en Salvendy (2009). Zij nemen als criteria het aantal echte problemen, het aantal false positives, de (gemiddelde) mate van ernst van het probleem, de 'thoroughness', de validiteit en de effectiviteit per expert per set heuristieken. Deze criteria zullen worden opgenomen in het onderzoek, evenals de mate van ernst van de gevonden problemen per set heuristieken. Daarmee verschilt deze scriptie van eerdere onderzoeken, waarbij websites in het algemeen als uitgangspunt voor onderzoek genomen zijn en niet specifiek naar gepersonaliseerde websites gekeken is

Op de genoemde begrippen wordt kort ingegaan.

Echte problemen

Echte problemen zijn de problemen gevonden door gebruikers. Het zijn usability problemen waar de gebruiker last van heeft wanneer men een website bezoekt (Ling & Salvendy, 2009).

False positives

False positives zijn problemen die enkel door experts zijn gevonden. Het zijn mogelijke problemen, geen echte problemen (Desurvire et al, 1991). Ook problemen die door experts worden beoordeeld met een score 0 op ernst, worden tot de false positive problemen gerekend (Ling & Salvendy, 2009). Het zijn wel problemen, maar geen usability problemen. Aangezien er geen vertaling voor de false positives is die de lading van het begrip in deze scriptie voldoende dekt, is het Engelse woord 'false positives' in deze scriptie gebruikt.

Mate van ernst

Niet alle fouten die door een expert worden geconstateerd zijn een probleem gezien vanuit usability van de gebruiker. Daarom is het enkel al uit overwegingen van kostenbeheersing van belang om te weten hoe de inschatting is van de ernst van de fout gezien vanuit het doel, het vergroten van usability voor gebruikers. Het zijn de experts die na afloop van een evaluatie de gevonden fouten beoordelen op de mate van ernst. Voor het ranken wordt in het algemeen gebruik gemaakt van de vijfpuntsschaal van Nielsen (Nielsen, 1993). De schaal is opgesteld naar rato van de inschatting dat herstel van de fout al dan niet leidt tot grote usabilityproblemen. De codering loopt van 0 (= geen usability probleem) tot 4 (= een ramp voor usability; herstel absoluut noodzakelijk voordat de site weer kan worden ingezet).

Thoroughness

Thoroughness wordt gedefinieerd als het vermogen, de kwaliteit van de methode om alle aspecten van een interface te kunnen evalueren. Thoroughness wordt berekend door het aantal echte problemen dat is gevonden te delen door het totaal aantal echte problemen dat bestaat (Bastien & Scapin, 1995). Het aantal echte problemen zijn problemen gevonden door alle gebruikers te samen. Hoe meer echte problemen er gevonden worden, hoe hoger de score op thoroughness, de score varieert van 0 tot 1.

De Nederlandse vertaling voor 'thoroughness' is grondigheid. Aangezien deze vertaling de lading van het begrip in deze scriptie niet voldoende dekt, is het Engelse woord 'thoroughness' in deze scriptie gebruikt.

Validiteit

Een methode is valide als een methode meet wat die moet meten. De validiteit kan berekend worden door het aantal echte problemen (problemen gevonden door de gebruikers) te delen door het totaal aantal gevonden problemen (Bastien & Scapin, 1995). Hoe minder false positives er door de experts gevonden worden, hoe hoger de validiteit. De score kan variëren van 0 tot 1.

Effectiviteit

Effectiviteit bevat informatie over zowel de validiteit als de thoroughness, het geeft de mate aan waarin de gebruikte methode is geoptimaliseerd. De effectiviteit kan berekend worden door de score op validiteit te vermenigvuldigen met de score op thoroughness (Sears, 1997).

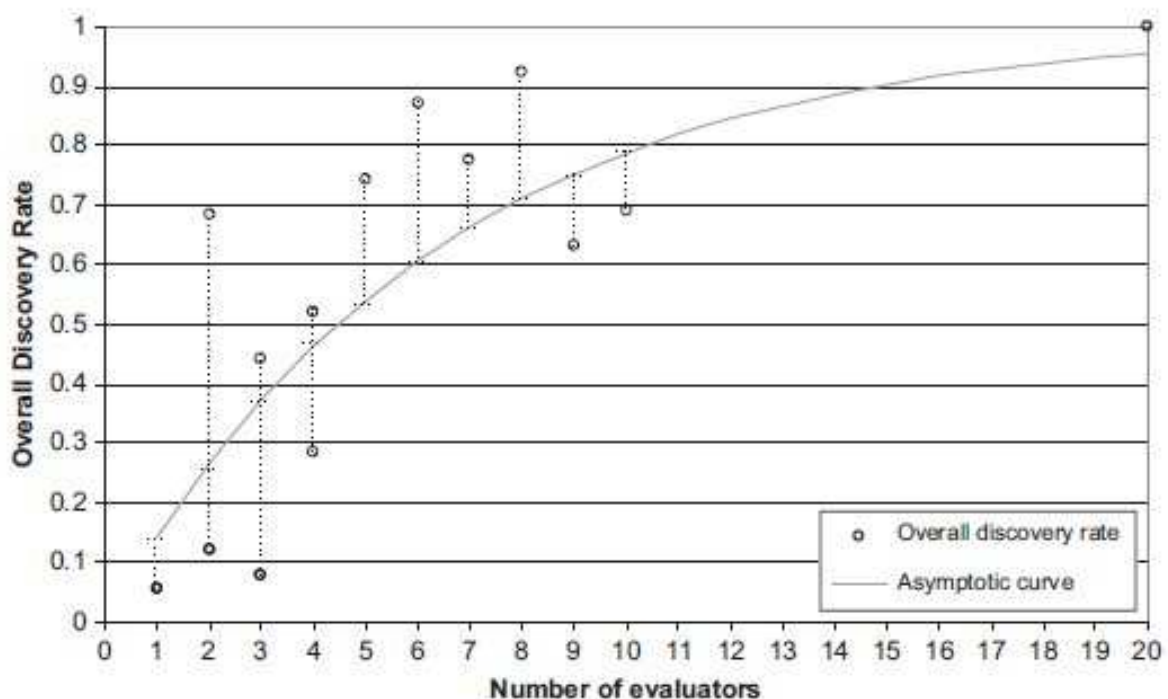
Hwang en Salvendy (2009) hebben gezocht naar een mogelijkheid om resultaten van verschillende onderzoeken met elkaar te kunnen verbinden en gelijk minder last te hebben van al die verschillende omstandigheden waaronder het onderzoek is verricht. Door bekende resultaten uit heuristische evaluaties en gebruikerstesten te vergelijken zijn ze gekomen tot een model bestaande uit een grafiek en een apart overzicht van omgevingsfactoren. In de grafiek wordt op de horizontale as de aantallen gebruikers cq. experts weergegeven en op de verticale as de mogelijke scores op thoroughness. Uit de analyse van de onderzoeken blijkt dat de relatie tussen beide variabelen kan worden weergegeven door een asymptotische kromme.

De omgevingsfactoren waartoe onder andere de gebruikte methode, de ervaring van de tester en de betrouwbaarheid van het systeem worden gerekend, worden gezien als onafhankelijke variabelen. De data worden op een eenvoudige manier beschreven zoals wel / niet aanwezig, wel / geen tijdsbeperking zodat vergelijking later mogelijk is.

De grafiek wordt in het onderzoek bij deze scriptie gebruikt als referentie. Ten eerste zal worden gekeken welke afzonderlijke score op thoroughness mag worden verwacht bij het gegeven aantal experts (6) en bij het gegeven aantal gebruikers (18), de te verwachten waarden. Het onderzoek leidt tot een berekende waarde voor thoroughness voor zowel het onderzoek met experts als voor gebruikerstest. Zo kan worden vastgesteld in hoeverre de werkelijk gevonden scores op thoroughness voor beide methodes van onderzoek afwijken van de te verwachten scores, de *n*-corrected effect sizes. Het zo gemeten verschil tussen de resultaten van beide methodes van onderzoek kan een indicatie zijn voor het wijzigen van het aandeel van beide methodes in het totaal aantal gevonden usability problemen bij gepersonaliseerde websites.

De grafiek van Hwang en Salvendy (2006) wordt weergegeven in figuur 2. Als voorbeeld kan dienen de beide onderzoeken met 4 gebruikers of experts. In de figuur is af te lezen dat bij 4 gebruikers of experts een waarde op de overall discovery rate verwacht mag worden van 0.47. Er is een eerste onderzoek te zien we bij 4 gebruikers of experts een berekende score op de overall discovery rate van 0.3. Dit is 0.17 ($0.47 - 0.3$) minder dan verwacht mag worden. En er is een tweede onderzoek te zien met 4 experts of gebruikers met een berekende score op de overall discovery rate van 0.52. Dit is 0.05 ($0.52 - 0.47$) hoger dan verwacht mag worden.

Figuur 2. Voorbeeld van de n-corrected effect size



2.7 Hoofdvraag en subvragen

Op basis van het theoretisch kader is de volgende hoofdvraag voor dit onderzoek opgesteld:

“Wat is de waarde van het testen door users en het testen door experts bij het verbeteren van de usability van gepersonaliseerde websites?”

Met als deelvragen:

1.1 Levert het toevoegen van heuristieken specifiek gericht op het vinden van problemen verbonden aan personalisatie, in een expertonderzoek een meerwaarde op ten aanzien van het gevonden aantal problemen en de gemiddelde mate van ernst van de problemen?

- 1.2 Levert het toevoegen van heuristieken specifiek gericht op het vinden van problemen verbonden aan personalisatie, in een expertonderzoek een meerwaarde op ten aanzien van de 'thoroughness en de validiteit van deze evaluatiemethode?
- 1.3 Blijkt uit de beoordelingen van de echte problemen op mate van ernst dat de problemen die gevonden zijn door experts hoger scoren op ernst dan de problemen gevonden door gebruikers?
- 1.4 Levert het onderscheiden van doelgroepen binnen de groep gebruikers een meerwaarde op ten aanzien van het gevonden aantal echte problemen en ten aanzien van de gemiddelde mate van ernst van deze problemen?

H3. Methodologische verantwoording

In dit hoofdstuk wordt de methode van onderzoek besproken die toegepast is om de onderzoeksvraag te beantwoorden. In de eerste paragraaf wordt aangegeven met welk onderzoeksontwerp gewerkt wordt. In paragraaf 3.2 wordt vastgesteld welke website is geselecteerd voor dit onderzoek. In paragraaf 3.3 komen alle gemaakte keuzes rondom de respondenten aan bod. Vervolgens wordt in paragraaf 3.4 uitgelegd welke onderzoeksinstrumenten worden ingezet. In paragraaf 3.5 wordt gesproken over de gevolgde procedures. In paragraaf 3.6 wordt de dataverzameling besproken en in de laatste paragraaf wordt ingegaan op de verwerking van de gegevens.

3.1 Onderzoeksontwerp

Ten behoeve van het vinden van een antwoord op de probleemstelling van dit onderzoek is gekozen voor een kwalitatief onderzoek met een kwantitatieve component. Voor dit onderzoek zijn twee groepen respondenten ondervraagd. Aan zowel gebruikers als experts werd een gepersonaliseerde website voorgelegd. Uit beide groepen hebben alle respondenten de mate waarin zij de onderzochte website vonden voldoen aan de kwaliteit usability beschreven.

3.2 Testobject

De website die in het kader van dit onderzoek is geëvalueerd, is de website van Ziggo (www.ziggo.nl). Deze website bestaat uit verschillende pagina's waar vormen van personalisatie te vinden zijn.

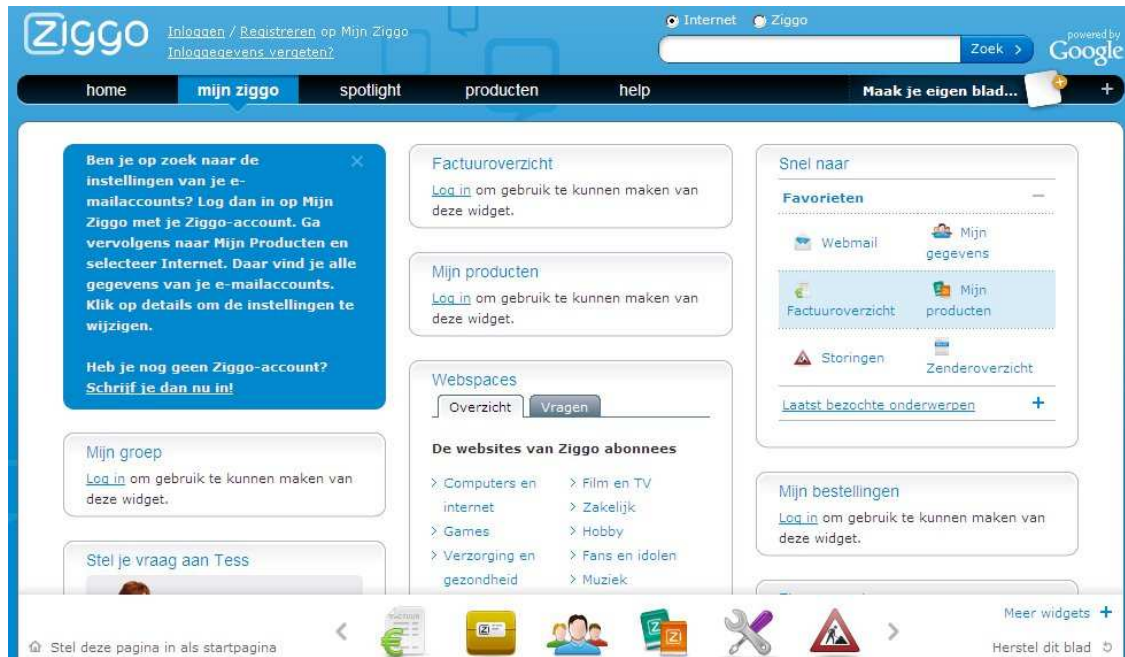
De website van Ziggo heeft een homepage die is ingedeeld in blokken. In figuur 2 is te zien hoe de homepagina van Ziggo er uit ziet. Er zijn verschillende menu-items naar eigen voorkeur in te richten. Zo is bij het nieuws aan te geven welke nieuwsgenres (binnenland, buitenland, economie, cultuur, sport) zichtbaar moeten zijn in het nieuwsblok. Bij het blok met verkeersinformatie is aan te geven hoe groot de lijst moet zijn (5 berichten, 10 berichten, 25 berichten), daarnaast is aan te geven van welke wegen je meldingen wilt krijgen. In het weerblok is aan te geven over welke regio de weerinformatie in het blok moet verschijnen. In het blok webradio zijn je persoonlijke favoriete radiozenders in te stellen. In het favorietenblok is het mogelijk de laatst bezochte onderwerpen weer te geven. Wanneer men ingelogd heeft, is het mogelijk de blokken naar een andere plek op de pagina te schuiven of

de blokken van de pagina te verwijderen. Onderaan de pagina zijn nog meer items te vinden, die met een klik op het item, op de homepagina kunnen worden gezet. Bijvoorbeeld een Hyvesblok of Flickrblok.



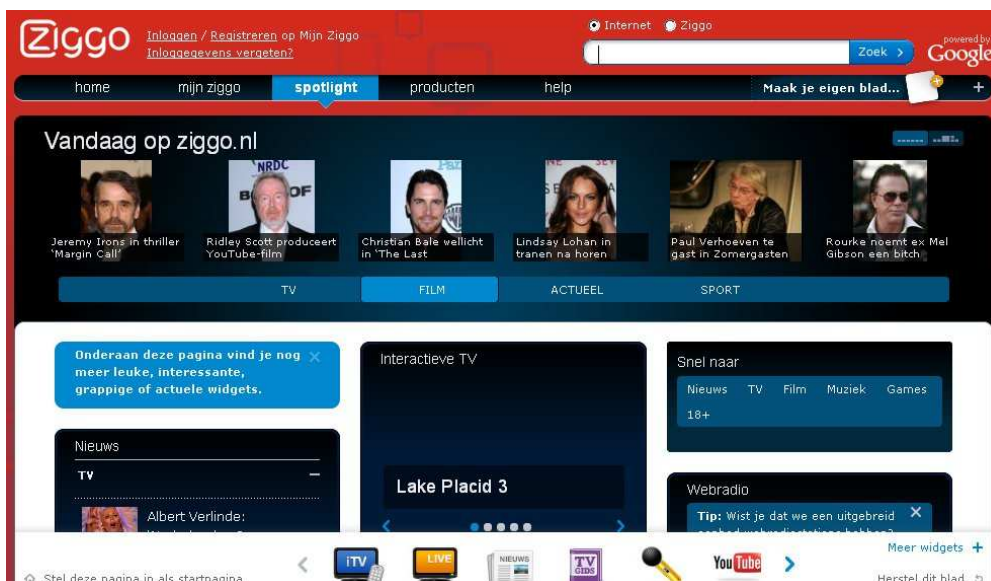
Figuur 3 Homepagina Ziggo

Daarnaast heeft de website een 'Mijn Ziggo' gedeelte. Om met 'Mijn Ziggo' te kunnen werken, moet de bezoeker ingelogd zijn. In figuur 3 is te zien hoe de pagina 'Mijn Ziggo' eruit ziet. Ook deze pagina is in blokken verdeeld. Er is een 'Factuuroverzichtblok' waar alle facturen van de ingelogde klant te zien zijn. In 'Mijn Groep' zijn alle ziggo-accounts zichtbaar, die aangemaakt zijn in 'Mijn Ziggo'. Het 'Mijn Productenblok' geeft weer welke producten de bezoeker afneemt bij Ziggo. Er is aan te geven of alle producten zichtbaar moeten zijn of dat er geen producten zichtbaar moeten zijn, wanneer men 'Mijn Ziggo' bezoekt. Daarnaast zijn de e-mailaccounts zelf te beheren en kunnen producten toegevoegd, gewijzigd en besteld worden. Ook administratieve wijzigingen zijn door te geven. In het 'Mijn Bestellingenblok' is te zien of en welke bestellingen er open staan. Mocht de bezoeker widgets aan de 'Mijn Ziggo' pagina willen toevoegen dan is er de keuze uit: storingen en meldingen, handig van Ziggo en contact met Ziggo. Een widget is de algemene naam voor de elementen waaruit de grafische interface kan worden opgebouwd (vensters, menuknoppen, keuzerondje). Net als op de homepagina zijn deze blokken op de pagina te verplaatsen, niet alle blokken zijn te verwijderen. De blokken die niet te verwijderen zijn, zijn blokken waar persoonlijke gegevens in staan.



Figuur 4 Pagina 'Mijn Ziggo'

Op de pagina 'Spotlight' zijn items te zien die met de producten die Ziggo verkoopt, te maken hebben. In figuur 4 is een screenshot van de pagina 'Spotlight' te zien. In het blok 'Interactieve tv' is aan te geven welke favoriete zenders onder andere zelf instellen welke genres je wilt zien via interactieve tv. Daarnaast kan bij de functionaliteit 'webradio' de favoriete zenders worden gekozen. Ook de digitale tv-gids is geheel naar eigen voorkeur op te maken.



Figuur 5 Pagina Spotlight

Verder biedt Ziggo de mogelijkheid een 'eigen blad' aan te maken op de website. Deze procedure is zichtbaar in figuur 5. De bezoeker kiest zelf wat voor soort blad hij/zij wil toevoegen, bijvoorbeeld een leeg blad, een WK 2010 blad of een eredivisie-blad. Daarna kan aan het gekozen blad een naam gegeven worden en kiest de bezoeker nog een achtergrondkleur. Wanneer er wordt gekozen voor de optie 'leeg blad' kan de content voor het blad gekozen worden uit de beschikbare widgets. De bezoeker sleept widgets die hij/zij op deze pagina wil hebben naar het lege blad. De overige bladen die gekozen kunnen worden zijn qua content al gevuld. Wel kan men hier de content weer zelf aanpassen, zoals ook op de overige pagina's van Ziggo het geval is.



Figuur 6 Eigen blad aanmaken

Het is overigens op elke pagina mogelijk om zelf extra widgets toe te voegen. De mogelijkheden verschillen per pagina. Voorbeelden van widgets zijn: Hyves, verkeersinformatie, weerinformatie, nieuwsberichten enzovoort. In figuur 6 is hiervan een voorbeeld te vinden.



Figuur 7 Voorbeeld van widgets op de website van Ziggo

3.3 Respondenten

Over het aantal gebruikers dat voor een gebruikersonderzoek moeten worden ingezet bestaat discussie. In het begin van de jaren negentig ging men ervan uit dat drie tot vijf gebruikers conclusies opleveren die valide zijn (Virzi,1992; Nielsen & Landauer,1993; Dumas en Redish,1993). Auteurs uit latere perioden pleitten ervoor meer gebruikers in een gebruikersonderzoek in te zetten. Op basis van eigen onderzoek en statistische berekeningen heeft Faulkner (2003) vastgesteld dat bij 20 testers minimaal 95% van de problemen wordt gevonden. De bevindingen van Faulkner worden in een latere studie van Tan et al. (2008) ondersteunt. Een belangrijke beweegreden om meer gebruikers in te zetten is de representativiteit van de respondenten, de samenstelling van de doelgroep van de steekproef is van groot belang (Tullis, 2008; Put,2006).

Aangezien er geen specifieke literatuur is waarbij onderzocht is of de opvattingen die in deze paragraaf zijn beschreven ook gelden voor websites met gepersonaliseerde onderdelen, is in dit onderzoek uitgegaan van de theorie over de samenstelling van een steekproef zoals in deze paragraaf beschreven. Omdat met het oog op representativiteit de groep gebruikers is onderverdeeld in drie groepen gebruikers, is voor 18 gebruikers gekozen (3 groepen van 6 gebruikers). Dit sluit aan bij de opvattingen die Faulker (2003), Tullis (2008) en van der Put (2006) hebben over aantallen gebruikers.

Ten aanzien van het aantal experts dat vereist is voor een kwalitatieve heuristische test merken Tan et. al. (2009) op dat het aantal experts afhankelijk is van de gebruikte heuristieken en de mate van ervaring van de experts met heuristieken. Er wordt aanbevolen 2-3 zeer ervaren experts te gebruiken of 3-5 iets minder ervaren experts.

In het onderzoek wordt gebruik gemaakt van 6 experts die ervaring hebben met het evalueren van websites. Dit past binnen de eisen die Tan et al. (2009) hebben geformuleerd ten aanzien van het aantal in te zetten experts.

3.3.2 Selectieprocedure respondenten

3.3.2.1 Experts

Aan de heuristische evaluatie werken 6 experts mee. Voor dit onderzoek wordt een expert gedefinieerd als iemand die afgestudeerd is in een studie op het gebied van webdesign of interactie design en minimaal 1 jaar werkzaam is in de ICT-branche op het gebied van usability en interaction-design (Desurvire et al.,1992). Op basis van deze selectiecriteria werden 6 experts per e-mail benaderd om aan dit onderzoek mee te werken. Alle 6 experts zegden toe om mee te willen werken aan het onderzoek.

De 6 experts zijn actieve deskundigen uit het zakelijk netwerk van de onderzoeker en zijn werkzaam bij verschillende bedrijven in de regio Noord-Nederland. De experts zijn voor het meewerken aan het onderzoek geen beloning in het vooruitzicht gesteld. De experts zijn geen bezoekers van de website van Ziggo. In onderstaand profiel wordt onder werkervaring verstaan het aantal jaren dat de expert werkzaam is in de ICT-branche. Zij zijn allen van het mannelijk geslacht. De gemiddelde leeftijd is 29.5. Teneinde de geldigheid en betrouwbaarheid van dit onderzoek te borgen is er bij de selectie van de experts gekozen voor professionele vertegenwoordigers uit verschillende bedrijven die geen afhankelijkheidsrelatie met elkaar en de onderzoeker onderhouden.

3.3.2.2 Gebruikers

Aan de gebruikerstesten werken 18 gebruikers mee. Voor dit onderzoek wordt een gebruiker gedefinieerd als iemand die geen ervaring heeft met het evalueren van websites en minimaal 2 jaar internetervaring. Daarnaast moet de respondent aan één van de volgende profielen, uitgewerkt in de drie doelgroepen voldoen:

Doelgroep A (studenten):

De groep gebruikers die onder persona A valt studeert aan het HBO of WO en is tussen de 18 en 30 jaar. Zij gebruiken internet dagelijks, zowel op hun laptop/pc als op hun mobiele telefoon. Ze zijn erg handig met internet en gebruiken het voor al hun dagelijkse bezigheden, dit kan zowel op hun studie, thuis, bij hun bijbaan als onderweg zijn.

Doelgroep B (werkende MBO-ers):

De groep gebruikers die onder persona B valt werkt en is tussen de 30 en 40 jaar. Zij hebben een Mbo-opleiding gevolgd. Zij gebruiken internet op hun werk en thuis af en toe om hun mail te checken of websites te bezoeken.

Doelgroep C (ouderen):

De groep gebruikers die onder persona C valt is tussen de 50 en 60 jaar. Zij werken en gebruiken internet voornamelijk voor werk- en hobbygerelateerde zaken. Zij zijn niet opgegroeid met internet en zien internet ook niet als een medium waar ze niet zonder kunnen.

Er is gekozen om te werken met persona's om het totale spectrum van de Nederlandse gebruikersgroepen te kunnen belichten. Persona's zijn profielen die worden opgesteld aan de hand van een archetypische gebruiker van de website (Mulken et al., 1998). Daarnaast is er gekozen voor deze groepen, aangezien de respondenten uit deze groepen in het dagelijks leven behoren tot potentiële bezoekers van de onderzochte website. Daarbij vertegenwoordigen de groepen de extremen in internetgebruik. De studenten uit dit onderzoek zijn met internet opgegroeid en gebruiken internet vrijwel overal voor. Zij zijn erg vaardig met websites en de navigatie hierop. Werkende MBO-ers tussen de 30 en 40 zijn in dit onderzoek de middengroep qua internetgebruik. Dit is ook zo in de criteria beschreven. De groep ouderen in dit onderzoek is het minst vaardig met internet.

Op basis van de in deze paragraaf genoemde selectiecriteria werden per doelgroep 6 respondenten per e-mail benaderd. Deze respondenten zijn relaties van collega's die voldoen aan de profielkenmerken van de drie onderscheiden persona's.

Alle 18 respondenten zegden toe om aan het onderzoek te willen deelnemen. De reden hiervoor is zeer waarschijnlijk dat zij via een voor hen bekende relatie zijn gevraagd voor dit onderzoek en affiniteit hebben met het onderwerp. De respondenten zijn afkomstig uit de regio Noord-Nederland. De respondenten zijn representatief voor de doelgroep van de website: het zijn gebruikers die in het dagelijks leven de onderzochte website zouden kunnen bezoeken. De respondenten zijn geen bezoekers van de website van Ziggo. De gebruikers is voor het meewerken aan het onderzoek geen beloning in het vooruitzicht gesteld.

3.4 Onderzoeksinstrumenten

3.4.1 Materiaalgebruik bij gebruikersonderzoek

De gebruikers beoordeelden de websites aan de hand van een aantal opdrachten. Een gebruikerstest bestaat uit 4 delen: de uitleg, het pre-interview, de uit te voeren opdrachten en een postinterview (Tan, 2003). Om in het gebruikersonderzoek de usability te meten is de hardopdenkmethode toegepast waarbij uitvoering van de taken en hardopdenken parallel lopen (concurrent think aloud).

In de uitleg is de testprocedure uitgelegd. De uitleg is in bijlage 1 te vinden. In het pre-interview is kort ingegaan op de internetervaring van de respondent en of de respondent bekend was met de website die onderzocht wordt. De vragen die hierbij gesteld worden, zijn te vinden in tabel 5.

Tabel 5 Pre-interview gebruikersonderzoek

Vraag	Onderwerp
1	Wat is uw leeftijd?
2	Wat is uw hoogst genoten opleiding?
3	Hoeveel uur per week surft u op internet?
4	Wat voor een type internetgebruiker bent u? Kies uit: Beginner / Gevorderd / Meer dan gevorderd / Expert
5.	Bent u bekend met ziggo.nl? Zo ja, hoe vaak per maand bezoekt u de website ziggo.nl?

Vraag 1 en 2 zijn algemene vragen. Vraag 3 geeft een indicatie van de ervaring van de gebruiker met internet. Het antwoord op vraag 4 geeft een beeld van hoe de respondent zichzelf ziet als gebruiker. Het antwoord op deze vraag is uiteraard subjectief maar kan mogelijk eventuele verschillen in resultaten verklaren. Vraag 5 geeft aan of de gebruiker een frequente gebruiker is van de ziggo.nl en daarmee al ervaring heeft met de site. Dit is een controlevraag ten aanzien van de gestelde voorwaarde dat een gebruiker geen bezoeker mag zijn van de site van Ziggo.

Vervolgens ontving de respondent een aantal opdrachten. Deze zijn hardopdenkend uitgevoerd. De opdrachten waren geformuleerd op basis van de verschillende onderdelen die de website www.ziggo.nl bevat. De opdrachten zijn te vinden in tabel 6.

Tabel 6 Opdrachten gebruikers

Vraag	Onderwerp
1	Je wilt een persoonlijk profiel aanmaken. Laat dit zien
2	De website bevat een pagina waar persoonlijke gegevens, facturen enzovoort te vinden zijn. Je wilt deze pagina naar eigen wens opmaken. Laat dit zien.
3	Je hebt een alles-in-1 pakket, maar je wilt een alles-in-1 plus abonnement. Hoe zou je dit doen?
4	Op de website van Ziggo wil je zelf een overzicht van je favoriete tv-zenders selecteren. Hoe zou je dit doen?
5	Je wilt je adres wijzigen, maar je kunt op de website niet vinden hoe dit moet. Hoe zou je dit doen?

Vraag 1 is gesteld om te testen of de gebruiker begreep waar je op de website van Ziggo een profiel kunt aanmaken en wanneer dit was gevonden, of de gebruiker in staat bleek om een persoonlijk profiel aan te maken. De pagina waar deze vraag om draait is de pagina 'Home' en eventueel, afhankelijk van hoe de gebruiker navigeerde, 'Mijn Ziggo'. Vraag 2 is opgenomen om te onderzoeken of de gebruiker begreep dat persoonlijke gegevens, facturen enzovoort te vinden zijn onder 'Mijn Ziggo'. Daarnaast was het interessant om te onderzoeken of de bezoeker op de pagina 'Mijn Ziggo' ook begreep dat de pagina naar eigen voorkeur in te richten is en gegevens zijn aan te passen. De pagina waar deze vraag om draait is 'Mijn Ziggo'. Vraag 3 moest aangeven of de gebruiker in staat bleek om een voorkeur te veranderen. De pagina waar deze vraag om draait is 'Mijn Ziggo'. Vraag 4 was er op gericht om te onderzoeken of de gebruiker een onderdeel van de website wat is te personaliseren, ook zelfstandig kon aanpassen. Daarnaast is er gekeken onder welk menu-item de gebruiker een onderdeel als de tv-gids verwachtte. De pagina waar deze vraag om draait is 'Spotlight'. Vraag 5 is gericht op de gebruiksvriendelijkheid van de zoek- en helpfunctie. De pagina waar deze vraag om draait is 'Help'.

De vragen uit het postinterview worden weergegeven in bijlage 4. Na een eerste algemene open vraag naar de mening van de gebruiker over de site kwamen in een meer gesloten vorm de items van Jameson aan bod zoals privacy, hinderlijkheid, begrijpelijkheid, voorspelbaarheid.

3.4.2 Expertonderzoek

Het expertonderzoek is uitgevoerd met behulp van een heuristische analyse. Experts onderzochten, onafhankelijk van elkaar, de website.

De experts beoordeelden de websites aan de hand van heuristieken. Het expertonderzoek bestond uit 3 delen: de uitleg, het pre-interview en het beoordelen van de website aan de hand van de heuristieken. In de uitleg is de testprocedure uitgelegd. De uitleg is in bijlage 5 te vinden.

In het pre-interview is kort ingegaan op de ervaring van de expert. Daarnaast kwam aan de orde of de expert bekend was met de website die onderzocht zou worden. De vragen die hierbij gesteld werden, zijn te vinden in tabel 7.

Tabel 7 Pre-interview expertonderzoek

Vraag	Onderwerp
1	Wat is uw leeftijd?
2	Welke vooropleiding heeft u gehad?
3.	Welk beroep oefent u uit?
4.	Hoeveel jaar houdt u zich met usability-onderzoek bezig?
5.	Heeft u ervaring met het gebruik van heuristieken?
6.	Bent u bekend met ziggo.nl? Zo ja, hoe vaak per maand bezoekt u de website ziggo.nl?

Vraag 1,2 en 3 zijn algemene vragen. De vragen 4 en 5 geven informatie over de ervaring die een expert heeft met usability-onderzoek. De antwoorden kunnen mogelijk achteraf gebruikt worden om eventuele verschillen in resultaten te verklaren. Vraag 6 is een controlevraag op de eis dat deelname aan het onderzoek slechts kan als men geen gebruiker is van de site van Ziggo.

Vervolgens is de expert gevraagd met de heuristieken aan de slag te gaan.

De heuristieken die voor dit onderzoek gebruikt worden zijn een combinatie van de heuristieken van Nielsen en Mack (1994) en Jameson (2007). De eerste tot en met de tiende heuristiek zijn de heuristieken van Nielsen. De elfde tot en met de zeventiende heuristiek zijn van Jameson.

Tabel 8 Heuristieken expertonderzoek

<p>1. Gebruiksgemak</p> <ul style="list-style-type: none"> - De bezoeker van de website krijgt optimaal feedback, snel en op het gewenste moment. - De bezoeker weet wat er gebeurt of wat er is gebeurd, wanneer hij/zij een actie uitvoert op de website.
<p>2. Helderheid</p> <ul style="list-style-type: none"> - De bezoeker begrijpt het taalgebruik van de website en is vertrouwd met de woorden, zinnen en concepten. - De bezoeker krijgt de content op een natuurlijke en in een logische volgorde gepresenteerd
<p>3. Gebruikerscontrole en vrijheid</p> <ul style="list-style-type: none"> - Bezoekers kunnen gebruik maken van een duidelijk herkenbare 'nooduitgang' wanneer zij achteraf gezien een verkeerde keuze hebben gemaakt. - De bezoeker is in staat om m.b.v. navigatie verkeerde handelingen ongedaan te maken of ongedaan gemaakte handelingen opnieuw uit te voeren
<p>4. Consistentie ontwerp gebruikersinterface</p> <ul style="list-style-type: none"> -Bezoekers hoeven zich niet af te vragen of verschillende woorden, situaties of acties dezelfde betekenis hebben, de presentatie is eenduidig en transparant.

5. Foutenpreventie

- Het ontwerp van de website garandeert dat de bezoeker probleemloos door de website kan navigeren.

6. Herkenning in plaats van herinnering

- Objecten, acties en opties zijn voor de bezoeker duidelijk herkenbaar, zodat de bezoeker weinig hoeft te onthouden.
 - De bezoeker hoeft geen informatie te onthouden als hij op zoek naar het gewenste resultaat de verschillende zoekstappen doorloopt.
 - Instructies over het gebruik van de website zijn duidelijk herkenbaar voor de bezoeker of kunnen makkelijk gevonden worden als de bezoeker dat wenst.

7. Flexibiliteit en efficiëntie

- De bezoeker hoeft geen onnodige handelingen te verrichten.
 - Onervaren bezoekers worden op de website anders tegemoet getreden dan ervaren bezoekers
 - Ervaren bezoekers worden in staat gesteld veel voorkomende handelingen te automatiseren door middel van bijvoorbeeld snelkoppelingen en sneltoetscombinaties.

8. Esthetische en minimalistische vormgeving

- De bezoeker wordt niet geconfronteerd met overbodige informatie.
 - Het ontwerp van de gebruikersinterface zorgt er voor dat de bezoeker niet met te veel informatie wordt geconfronteerd.
 - De vormgeving past bij de bezoeker.

9. Hulp bij herkennen, vaststellen, en herstellen van fouten

- De bezoeker begrijpt foutmeldingen.
 - De bezoeker wordt niet onnodig geconfronteerd met "error-codes".

10. Goede helpfunctie en documentatie

- De bezoeker krijgt voldoende informatie aangeboden wanneer hij / zij dit wenst.
 - De informatie is op een eenvoudige manier beschikbaar voor de bezoeker.
 - De bezoeker kan de website gebruiken zonder documentatie en helpschermen.
 - De documentatie is voor de bezoeker overzichtelijk weergegeven.

11. Voorspelbaarheid

De bezoeker is in staat om te voorspellen wat de bijdrage zal zijn van zijn/haar actie aan het te realiseren resultaat.

12. Begrijpelijkheid

De bezoeker is in staat zijn om te begrijpen hoe het systeem werkt ten aanzien van zijn/haar profilering en ten aanzien van de informatie die hij / zij ontvangt

13. Controleerbaarheid

De bezoeker is in staat om controle uit te oefenen op zijn/haar profilering en op het genereren van de gepersonaliseerde output

14. Hinderlijkheid

Een bezoeker is in staat om zijn/haar taak af te maken zonder daarbij gehinderd te worden door bijwerkingen van de personalisatie.

15. Privacy

De bezoeker krijgt niet het gevoel dat de profilering door het systeem, zijn/haar behoefte aan privacy schaadt.

16. Missen informatie i.v.m. profiel

De bezoeker wordt niet de mogelijkheid ontnomen om nieuwe informatie aangeboden te krijgen omdat deze informatie niet past in het van hem/haar vastgelegde profiel

17. Systeem bevoegdheid

De bezoeker krijgt niet de indruk dat het systeem werkt met een onjuist profiel of dat output niet herkenbaar personaliseert.

3.5 Experimentele procedure

Het onderzoek is afgenomen in Groningen, in een testruimte met een laptop en een videocamera. Op deze manier kon de test geregistreerd worden.

Zowel gebruiker als expert kregen aan het begin van het onderzoek uitleg over wat zij konden verwachten en wat verwacht werd van het uitvoeren van de taak. Vervolgens is het onderzoek gestart. Het onderzoek omvatte 24 sessies, alle uitgevoerd op dezelfde locatie. Er zijn 18 sessies met gebruikers uitgevoerd en 6 sessies met experts.

3.5.1 Experimentele procedure gebruikers

Bij de start van het onderzoek ontvingen de gebruikers een uitleg. In de uitleg is de testprocedure toegelicht. Daarin zijn benoemd: het doel van het onderzoek, de testmethode, de procedure, de aanwezigheid van apparatuur, anonimiteit en het uitschakelen van de mobiele telefoon. Vervolgens kreeg de respondent de mogelijkheid om vragen te stellen. Deze uitleg is in bijlage 1 te vinden. Wanneer eventuele vragen van de respondent waren beantwoord, startten de opnamen en het pre-interview.

In het pre-interview is kort ingegaan op de internetervaring van de respondent en of de respondent bekend was met de website die onderzocht werd. Vervolgens kreeg de respondent een aantal opdrachten. Tijdens de opdrachten is de respondenten gevraagd om hardop uit te spreken welke gedachten zij hadden tijdens het uitvoeren van de opdrachten. In het postinterview, in de vorm van een vragenlijst, zijn de gebruiker bevraagd op de specifieke items behorend bij met name personalisering van websites. Na het postinterview zijn de gebruiker bedankt en is het onderzoek afgesloten.

3.5.2 Experimentele procedure experts

Bij de start van het onderzoek ontvingen de experts een uitleg. In de uitleg is de testprocedure uitgelegd. De uitleg is in bijlage 5 te vinden. Vervolgens kreeg de expert de mogelijkheid om vragen te stellen. Wanneer eventuele vragen van de expert waren beantwoord, startten de opnamen en het pre-interview.

In het pre-interview is kort ingegaan op het beroep van de expert, hoeveel jaar de expert zich met usability-onderzoek bezig hield en wat zijn ervaring met heuristischeken was. Daarnaast kwam aan de orde of de expert bekend was met de website die onderzocht werd. De vragen die hierbij gesteld worden, zijn te vinden in bijlage 6.

Vervolgens is de expert gevraagd met de heuristieken aan de slag te gaan. De expert ontving van de onderzoeker een tabel waarin de heuristieken uit tabel 1 zijn opgenomen. Met behulp van deze tabel kunnen usability-problemen worden opgespoord. De expert ontving ook een beschrijving van de persona (bijlage 10,11,12), van waaruit hij met de heuristieken aan de slag ging. De groep gebruikers is onderverdeeld in 3 persona's. Er zijn twee experts die vanuit persona A naar de website gekeken hebben. Twee experts hebben vanuit persona B naar de website gekeken. En twee experts hebben vanuit persona C naar de website gekeken

De expert nam daarvoor plaats achter twee laptops. Op de ene laptop kon de website bekeken worden, op de andere laptop konden de gevonden problemen genoteerd worden. De expert noteerde elk probleem in een worddocument op een nieuwe bladzijde, het document was opgemaakt in een vast format. De expert noteerde per probleem onder welke heuristiek deze viel en op welk tijdstip het probleem geconstateerd was. Tabel 9 geeft de verdeling per persona weer.

Tabel 9 **Verdeling persona's per expert**

Expert	Persona
E1	Persona A
E2	Persona B
E3	Persona C
E4	Persona A
E5	Persona B
E6	Persona C

Nadat de expert de problemen genoteerd had, is de expert bedankt en werd het onderzoek afgesloten.

Van alle door de experts geconstateerde problemen en de problemen geconstateerd door de gebruikers is door de onderzoeker in 3 dagen één lijst gemaakt. Nadat de dubbelingen waren verwijderd zijn de unieke problemen door de zes experts voorzien van een code ten aanzien van de mate van ernst. De zes experts zijn hiervoor benaderd per e-mail, zij ontvingen een lijst met daarop de unieke problemen beschreven, waarachter zij een code ten aanzien van de mate van ernst konden plaatsen. Het gewicht van de mate van ernst is geclassificeerd met behulp van een vijf punts Likert Schaal (Nielsen, 1993):

0 = geen usability probleem

1 = enkel een cosmetisch probleem; enkel herstellen als er voldoende tijd beschikbaar is

2 = klein usability probleem; herstel heeft een lage prioriteit

3 = groot usability probleem; herstel heeft een hoge prioriteit

4 = een ramp voor usability; herstel absoluut noodzakelijk voordat de site weer kan worden ingezet.

3.6 Dataverzameling

De data zijn verzameld gedurende een periode van twee weken. De data bestaan uit door gebruikers ingevulde vragenlijsten, door experts gerapporteerde problemen, opnames van de gebruikers- en experttesten en observaties van gedrag. Voor elke sessie van zowel de expert als de gebruiker was maximaal 30 minuten beschikbaar.

De dataverzameling bestaat uit 393 minuten opnamemateriaal van het gebruikersonderzoek. De dataverzameling bestaat uit 180 minuten opnamemateriaal van het expertonderzoek.

3.7 Data analyse

Aangezien de data-analyse van Norgaard en Hornbaek (2006) voor wat betreft de inhoud, het analyseren van data uit de hardopdenkmethode, goed aansluit op deze scriptie, is deze data-analyse als uitgangspunt gebruikt.

Documenteren opnames

De opnames van de gebruikerstesten zijn systematisch in Word vastgelegd. Bij de opdrachten zijn tijdstippen, pagina's en de output van het hardopdenken verwerkt. Gedrag van de gebruiker is in cursieve tekst weergegeven. Het hardopdenken van de gebruiker is tussen dubbele aanhalingstekens weergegeven. Opmerkingen en constatering van de onderzoeker zijn met rode tekst aangegeven.

De experttesten zijn door de experts zelfstandig systematisch vastgelegd. Hierbij zijn tijdstippen, pagina's en de output per heuristiek door de expert genoteerd.

Aanmaken unieke lijst en Interbeoordelingsbetrouwbaarheid

Om tot een unieke lijst met problemen te komen zijn de gelijksoortige problemen door 2 onderzoekers bij elkaar geplaatst en verdeeld over 11 categorieën. Bijvoorbeeld informeren, sturen en privacy. Als tweede onderzoeker is een collega afgestudeerd in communicatie wetenschappen opgetreden.

De interbeoordelingsbetrouwbaarheid tussen beide onderzoekers is vastgesteld met Cohen's kappa.

Vervolgens zijn binnen elke categorie subcategorieën tot stand gebracht door wederom gelijksoortige problemen bij elkaar te plaatsen. Bijvoorbeeld privacy 1, 2 en 3.

En is het overeenstemmingspercentage tussen de onderzoekers berekend.

Nadat de onderzoekers tot consensus waren gekomen over de verschillen, kwam er een lijst van 75 unieke problemen (bijlage 13) tot stand.

Mate van ernst bepalen

De lijst met enkel de unieke problemen is per e-mail aan alle experts verstuurd. Zij hebben de mate van ernst vastgesteld. De mate van ernst is vastgesteld met behulp van het toekennen van een codering. Dit is beschreven in paragraaf 3.5.2. De codering loopt van 0 (= geen usability probleem) tot 4 (= een ramp voor usability; herstel absoluut noodzakelijk voordat de site weer kan worden ingezet).

Voorspellen resultaten

Om een voorspelling te doen over de te behalen resultaten door gebruikers en door experts afzonderlijk is gebruik gemaakt van de berekeningen van Hwang and Salvendy (2009).

Dit biedt een mogelijkheid om resultaten van verschillende onderzoeken met elkaar te kunnen verbinden en gelijk minder last te hebben van de verschillende omstandigheden waaronder het onderzoek is verricht. Door bekende resultaten uit heuristische evaluaties en gebruikerstesten te vergelijken is een model bestaande uit een grafiek en een apart overzicht van omgevingsfactoren ontstaan. In de grafiek wordt op de horizontale as de aantallen gebruikers cq. experts weergegeven en op de verticale as de mogelijke scores op thoroughness. De relatie tussen beide variabelen wordt weergegeven door een asymptotische kromme. Hiermee wordt op basis van data uit voorgaand onderzoek een verband aangegeven tussen het aantal gebruikers of experts en de daarbij te verwachten score op 'thoroughness' (=overall discovery rate). Als het aantal gebruikers en experts bekend is, dan kan hun bijbehorende score op thoroughness worden afgelezen. Ten eerste is gekeken welke afzonderlijke score op 'thoroughness' (=overall discovery rate) mocht worden verwacht bij het gegeven aantal experts (6) en bij het gegeven aantal gebruikers (18), de te verwachten waarden. Het onderzoek leidde tot een berekende waarde voor thoroughness voor zowel het onderzoek met experts als voor gebruikerstest. Zo kon worden vastgesteld in hoeverre de werkelijk gevonden scores op thoroughness voor beide methodes van onderzoek afweken van de te verwachten scores, de n-corrected effect size. Deze grafiek is gepresenteerd in paragraaf 2.6.

Vaststellen resultaten

Om de resultaten van de experts en gebruikers ten aanzien van het vinden van usability problemen met elkaar te vergelijken is gekeken naar mate waarin de in het onderzoek berekende score op thoroughness voor gebruikers en experts afzonderlijk afweek van de afzonderlijk voorspelde scores op thoroughness van gebruikers en experts. Deze zogenaamde n-corrected effect size van beide methodes werden met elkaar vergeleken. Om vast te stellen of in het onderzoek van de experts het toevoegen van de heuristieken van Jameson had geleid tot een stijging van de gemiddelde mate van ernst, de validiteit en de effectiviteit (Ling and Salvendy, 2009) zijn deze statistische maten voor zowel de heuristiek van Nielsen als van Jameson berekend. De gemiddelde mate van ernst per probleem is berekend door de scores van de zes experts op een probleem op te tellen en te delen door 6. De gemiddelde mate van ernst gaf een indicatie van hoe ernstig de problemen gemiddeld zijn en dus de mate van noodzaak van repareren. De validiteit geeft aan of de experts en gebruikers in staat zijn geweest om de problemen te identificeren die ook echt impact hebben op de usability. De effectiviteit wordt vastgesteld door validiteit en thoroughness met elkaar te vermenigvuldigen, het geeft de mate aan waarin de methode is geoptimaliseerd. De maximumscore is 1. De manier waarop deze statistische maten worden berekend zijn beschreven in paragraaf 2.6.

Hoofdstuk 4 Resultaten

In hoofdstuk 3 is beschreven welke methoden en technieken zijn gebruikt om te komen tot een antwoord op de centrale vraagstelling en de deelvragen. In dit onderzoek zijn twee deelpopulaties onderscheiden, gebruikers en experts. De belangrijkste resultaten van gebruikers en de experts zijn zowel apart als in een vergelijking beschreven. Voordat hier op wordt ingegaan, zal eerst besproken worden hoe de lijst met echte problemen tot stand is gekomen.

4.1 Analyse

De gebruikers en experts hebben gezamenlijk 556 problemen gevonden, waarvan 444 gevonden door gebruikers en 112 door experts. Om tot een unieke lijst met problemen te komen zijn de gelijksoortige problemen door 2 onderzoekers bij elkaar geplaatst en verdeeld over 11 categorieën. Over de indeling van 52 problemen verschilden de onderzoekers van mening. De interbeoordelingsbetrouwbaarheid tussen beide onderzoekers is vastgesteld met Cohen's kappap en bleek van hoog niveau ($k = 0,881$).

Na het bereiken van consensus over de verschillen zijn er vervolgens binnen elke categorie subcategorieën tot stand gebracht door wederom gelijksoortige problemen bij elkaar te plaatsen. Nu bleken de onderzoekers over de indeling van 80 problemen van mening te verschillen. Over 476 van de 556 waren ze het eens, dit is een overeenstemmingspercentage van 86%, een hoge score. Nadat de onderzoekers tot consensus waren gekomen over de verschillen, kwam er een lijst van 75 unieke problemen (bijlage 13) tot stand. In dit geval is geen gebruik gemaakt van Cohen's kappap omdat dit gezien de complexiteit voor de onderzoeker niet mogelijk bleek.

Van de 75 unieke problemen zijn 39 problemen gevonden door de gebruikers, 6 door de experts en 30 door zowel de gebruikers als de experts samen.

Een opsomming van enkele uitingen om een beeld te geven van de unieke problemen die alleen door de gebruikers zijn geconstateerd:

- “Bezoeker interpreteert op de website gebruikte termen op een andere manier dan Ziggo deze termen bedoeld heeft. Dit is bijvoorbeeld het geval bij: overstapservice, zenderoverzicht, interactieve tv.”
- “Bezoeker denkt met een ander onderdeel van de website bezig te zijn dan werkelijk het geval is. De bezoeker ziet bijvoorbeeld het zenderoverzicht aan voor de tv-gids.”
- “De lay-out van de tv-gids is onoverzichtelijk. Titels zijn niet helemaal zichtbaar en programma's kunnen beter van elkaar onderscheiden worden.”

Een opsomming van enkele uitingen om een beeld te geven van de unieke problemen die alleen door de experts geconstateerd zijn:

- “Het komt voor dat blokken blijven laden. Weghalen is niet mogelijk en er verschijnt geen foutmelding of hulp.
- “De website kent weinig call to action.”
- “Wegklikken blokken is duidelijk. Maar hoe komen deze weer terug?”

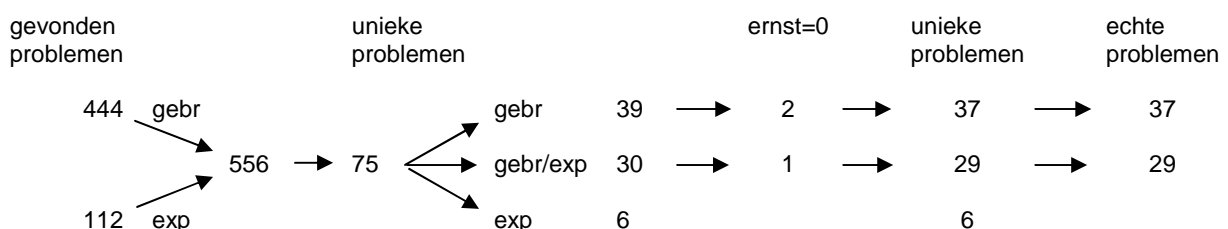
Een opsomming van enkele uitingen om een beeld te geven van de unieke problemen die door zowel de experts als de gebruikers geconstateerd zijn, de overlappings:

- “Op de website worden termen/iconen gebruikt die de bezoeker niets zeggen. Bijvoorbeeld: Spotlight, enkele iconen van de widgets.”
- “Er staat te veel informatie op de website en de site oogt erg druk en vol.”
- “Voor de bezoeker is het niet duidelijk hoe de navigatie van het hoofd/submenu werkt.”

Van de lijst van 75 unieke problemen zijn alle problemen door elke expert afzonderlijk beoordeeld op de mate van ernst. Van 3 problemen bleek dat de score 0 bedroeg, een score die het rechtvaardigde om aan te nemen dat het weliswaar om een probleem ging maar niet om een usability probleem van de onderzochte website, zogenaamde non-problems. Het betroffen uitgesproken verwachtingen van gebruikers die gebaseerd waren op ervaringen met andere websites. Er resteerden 72 problemen. Van deze 72 unieke usability problemen zijn er 6 enkel gevonden door experts. Echte problemen zijn de problemen die zijn gevonden door gebruikers. De 6 problemen door de experts gevonden waren mogelijke problemen en daarmee false positives.

Van de echte problemen zijn er 37 vastgesteld door de gebruikers en 29 door de gebruikers en experts samen. In figuur 7 is een overzicht te vinden van het proces waarin de lijst met echte problemen tot stand is gekomen.

Figuur 7 Tot stand komen echte problemen



4.2 Gebruikersonderzoek

Bij het gebruikersonderzoek zijn metingen afgenomen bij 18 gebruikers. Deze gebruikers behoorden tot één van de drie van tevoren vastgestelde doelgroepen, studenten (6), werkende MBO-ers (6) en ouderen (6). Nadat een pre-interview is afgenomen, heeft elke gebruiker 5 opdrachten uitgevoerd, gevolgd door een post-interview.

4.2.1 Pre-interview

Uit het pre-interview bleek dat de gemiddelde leeftijd van de gebruikers 38,6 jaar is. De leeftijd van de groep studenten varieerde van 23 tot en met 27 met een gemiddelde leeftijd van 25 jaar. De gemiddelde leeftijd van de groep werkende MBO-ers was 38 jaar met een minimum van 30 en een maximum van 48 jaar. De gemiddelde leeftijd van de ouderen was 53 jaar en varieerde van 50 tot 58 jaar. Aan het gebruikersonderzoek namen 8 mannen en 10 vrouwen deel. Van de 18 respondenten hebben er 10 (55%) een WO-opleiding, 2 (11%) een Hbo-opleiding en 6 (34%) een Mbo-opleiding afgerond. De gebruikers surfen gemiddeld 13,7 uur per week op internet waarbij de studenten het meeste aantal uren per week op internet te vinden waren, namelijk gemiddeld 15,8 uur. De werkende MBO-ers surfen gemiddeld 13,3 uur op het internet en de ouderen gemiddeld 9,2 uur. In tabel 10 zijn de profielgegevens van de doelgroepen te vinden.

Tabel 10 **Profielgegevens doelgroepen**

	Gemid. leeftijd	Surfen p/week ¹	man	vrouw	WO	HBO	MBO
Studenten	25	15,8	3	3	6	0	0
Werkende MBO-ers	38	13,3	2	4	0	0	6
Ouderen	53	9,2	3	3	4	2	0

¹ gemiddeld aantal uren

De meeste gebruikers classificeerden zichzelf als een gevorderde internetgebruiker (61%). Van de gebruikers noemde 5% zichzelf expert en 38% vergevorderd. Niemand gaf aan zichzelf als een beginner te zien. Van de 18 gebruikers gaven 10 respondenten aan bekend te zijn met Ziggo. Dit had verder geen invloed op het goed kennen van de website. Bijna alle respondenten die aangaven bekend te zijn met Ziggo, gaven aan nooit de website te bekijken (80%). De overige 20% van de respondenten bekeek gemiddeld 1 keer per jaar de website.

4.2.2 Bijdrage per doelgroep

De gebruikers waren verdeeld in drie doelgroepen: studenten (groep 1), werkende MBO-ers (groep 2) en ouderen (groep 3). Het aantal echte problemen bedroeg 66. Doelgroep 1 heeft 55 (85%) van de 66 problemen gevonden, voor doelgroep 2 was het aantal 37 (56%) en voor doelgroep 3 was het 48 (74%). Op basis van deze getallen kon worden gezegd dat doelgroep 1 de hoogste scores had behaald. Het gemiddelde waarbij rekening is gehouden met de standaarddeviatie en de informatie over de range hoog – laag moest informatie geven over hoe dit resultaat binnen de groep tot stand was gekomen. Het zou kunnen dat 1 of 2 gebruikers de uiteindelijke score voor de groep erg hadden beïnvloed. Bij groep 2 was zichtbaar dat het gemiddelde + de standaarddeviatie bijna overeenkwam met de hoogste score en de standaarddeviatie relatief laag was. Dit betekende dat er geen uitschieters naar boven voorkwamen en de scores dicht bij elkaar lagen. Bij groep 1 en 3 was de standaarddeviatie groter wat betekende dat de prestaties van de leden van de doelgroep verder uiteen liepen. De hoogste score in doelgroep 1 en de hoogste score in doelgroep 2 die beide meer dan een standaarddeviatie van het gemiddelde afweken, hebben het resultaat van deze groepen erg beïnvloed. De persoon met de hoogste score in groep 1 was alleen al goed voor 35 van 55 gevonden problemen, dit is 65%. Voor de persoon met de hoogste score in groep 3 ging het om ruim 50%. In tabel 11 worden de doelgroepen vergeleken op echte problemen, het gemiddelde aantal echte problemen wat per doelgroep gevonden is, de standaarddeviatie en de hoogste en laagste score op het aantal gevonden echte problemen.

Tabel 11 Doelgroepen vergeleken

	echte problemen	M	SD	hoogste score	laagste score
doelgroep 1	55	23,2	7,22	35	17
doelgroep 2	37	12,7	2,94	17	8
doelgroep 3	48	17	5,05	25	12

Om een beeld te krijgen van de relatie doelgroepen en mate van ernst van de gevonden problemen werden de scores 1 en 2 samengenomen en idem de scores 3 en 4. De scores 3 en 4 zijn samengevoegd omdat bij een beoordeling 3 en 4 op ernst er onmiddellijk moet worden ingegrepen. Bij een score 1 of 2 ligt de prioriteit lager. Uit de verdeling in procenten bleken de doelgroepen 1 en 3 tot een vergelijkbare score te komen.

Ongeveer 30% van de gevonden problemen waren van dien aard dat ingrijpen een hoge prioriteit had. Doelgroep 2 bleef al achter met het aantal gevonden problemen maar het bleek dat de gevonden problemen ook minder vaak (22%) dan bij de doelgroepen 1 en 2 om onmiddellijke actie vroegen. Van de in totaal 66 problemen zijn er 15 problemen gescoord 3 en 3 problemen met een 4. met In tabel 12 worden de doelgroepen vergeleken op de mate van ernst.

Tabel 12 Doelgroepen vergeleken op mate van ernst

	ernst 1 en 2		ernst 3 en 4	
	aantal	%	aantal	%
doelgroep 1	40	73%	15	27%
doelgroep 2	29	78%	8	22%
doelgroep 3	34	71%	14	29%

4.2.3 Problemen per pagina

De website was verdeeld over vijf pagina's: Home, Mijn Ziggo, Producten, Spotlight en Help. De gebruiker hebben per pagina verschillende soorten problemen gevonden.

De pagina 'Home' werd over het algemeen als een drukke pagina ervaren waar veel informatie op staat die de gebruiker niet allemaal als even relevant ervaart. Een opsomming van enkele uitingen om een beeld te geven van de problemen die door de gebruikers op de 'Home' geconstateerd zijn:

- "Ik vind 'm te druk, er zijn teveel hoofdbalken die mijn aandacht trekken. Bovenaan heb je de echte hoofdbalk, maar daaronder zitten ook allemaal balken die mijn aandacht trekken en dat leidt heel erg af. De navigatie vind ik niet heel logisch. Blokken zijn te groot en teveel aanwezig".
- "Er staat heel veel niet relevante content op, zoals nieuws, webradio, samen gamen. Ik vind het niet nuttig op deze pagina. Als ik nieuws wil ga ik wel ergens anders naar toe."
- "Less is more. Het moet gewoon veel simpeler, er zijn teveel kleine vakjes op het scherm, met teveel onzin info, waar je helemaal niet naar op zoek bent"

De vele functionaliteiten op het gebied van personalisering die de pagina 'Mijn Ziggo' biedt worden niet door alle gebruikers even goed begrepen. Daarnaast wordt ook deze pagina als een drukke pagina ervaren. Een opsomming van enkele uitingen om een beeld te geven van de problemen die door de gebruikers op 'Mijn Ziggo' geconstateerd zijn:

- “Ik zie hier staan 'meer widgets'. Ik weet niet wat een widget is.”
- “Maar er staat bij facturen ‘log in’ en bij producten ‘log in’ en overal op de pagina, ik snap niet wat het verschil daartussen is.”
- “Het is iets te druk op de Mijn Ziggo pagina vind ik, ik weet niet waar mijn aandacht naar toe moet, het is teveel in één keer”

Geen enkele gebruiker had echt een idee wat men bij de pagina ‘Spotlight’ zou moeten verwachten. Een opsomming van enkele uitingen om een beeld te geven van de problemen die door de gebruikers op de pagina ‘Spotlight’ geconstateerd zijn:

- “Ik heb geen idee wat ik mij bij Spotlight moet voorstellen. Waarom Spotlight? Dit had ik niet verwacht”.
- “Bij Spotlight zou ik nieuwe producten van Ziggo verwachten. Maar ik zie vandaag op Ziggo, nieuws, tv-gids, you-tube. Dit had ik niet verwacht”.
- “Er staat heel veel extra's op, waar mensen waarschijnlijk niet eens naar zoeken”

De pagina ‘Producten’ werd niet door alle gebruikers even goed begrepen. Een opsomming van enkele uitingen om een beeld te geven van de problemen die door de gebruikers op de pagina ‘Producten’ geconstateerd zijn:

- “Plus dat ik niet weet of dit nou reclame is of dat het erbij hoort zo'n blok bovenin”.
- “Het was wel verrassend wat ik af en toe tegenkwam als ik op een link klikte, het is een verrassing wat er komt. Ik wist niet gelijk aan de beschrijving te lezen wat ik kon verwachten”
- “Nieuwe producten bestellen is nu nog makkelijker als je inlogt. Kennelijk moet ik dan weer inloggen, maar ik ben toch al ingelogd?”

Op de pagina ‘Help’ wordt de werking van de zoekfunctie niet door alle gebruikers goed begrepen. Daarnaast is ook deze pagina als druk ervaren.

Hierna volgt een opsomming van enkele uitingen om een beeld te geven van de problemen die door de gebruikers op de pagina ‘Help’ zijn geconstateerd:

- “Bij de zoekknop verwacht je dat knopje niet dat je expliciet moet aangeven dat je op Ziggo wilt zoeken. Want je wilt liever niet honderdduizend andere sites in je overzicht hebben.”
- “608 resultaten gevonden op tv-gids, dat is wel heel veel”
- “De hele opsomming aan de rechterkant is heel vlak, ik weet echt niet waar ik moet beginnen.”

4.2.4 Post-interview

In het post-interview werden 9 vragen aan de gebruiker gesteld. De meeste respondenten konden weinig positieve aspecten opnoemen over de website van Ziggo. Wanneer men toch positieve aspecten aangaf dan hadden deze betrekking op Tess, de helpfunctie van Ziggo in de vorm van een assistente. Eén van de gebruikers merkte hierover op “Ik vind het wel aardig dat ze Tess inzetten, dat ze het persoonlijk maken”. De gebruikers konden meerdere negatieve aspecten opnoemen. Hierna volgt een opsomming van enkele uitingen om een beeld te geven van de negatieve aspecten:

- Over de opmaak van de pagina's:
“Ik vind het heel onoverzichtelijk, veel te vol, er staat veel te veel op één pagina.”
- Over de inhoud:
“Er staan bijvoorbeeld favorieten en zo, en allemaal dingen die je niet verwacht op een website van een internetprovider. Zoals het nieuws, webradio, het weer. Ik ga naar de website van een internetprovider om daar dingen voor m'n internet of tv te regelen en niet om daar het verkeer te bekijken.”
- Over de zoekfunctie:
“Ik vind het verwarrend dat er direct op internet gezocht wordt i.p.v. Op de Ziggo-site. Als je dat even niet door hebt, dan krijg je allemaal sites waar ik gewoon niks mee wil”.
- Over het taalgebruik en iconen:
“Sommige dingen waren een beetje raden, ik moest bij een aantal dingen toch wel eerst erop klikken en dan zag je pas wat het werkelijk was”.

De problemen die uit het post-interview naar voren kwamen, zijn verwerkt in het overzicht waarin alle problemen van de gebruikers weergegeven worden.

4.3 Expertonderzoek

Bij het expertonderzoek zijn metingen afgenomen bij 6 experts. De 6 experts hebben aan de hand van 17 heuristieken de website www.ziggo.nl op usability problemen beoordeeld. Van de heuristieken waren er 10 ontleend aan Nielsen en 7 aan Jameson. Aan de beoordeling van de experts is een pre-interview voorafgegaan.

4.3.1 Pre-interview

Uit het pre-interview bleek dat de gemiddelde leeftijd van de experts 29,5 jaar is. De leeftijd van de experts varieerde van 23 tot en met 42 jaar. Vier experts hebben een HBO-opleiding afgerond, de overige twee experts hebben een WO-opleiding afgerond. De beroepen van de experts varieerden van web/usability specialist tot interaction designer, maar allemaal hielden zij zich onder andere bezig met usability-onderzoek. De experts hadden gemiddeld 3,8 jaar ervaring met usability-onderzoek. De ervaring van de experts varieerde van 2 tot en met 6 jaar. Alle experts hadden ervaring met het gebruik van heuristieken. Van de 6 experts gaf 50% aan bekend te zijn met Ziggo. In tabel 13 worden de profielen van de experts beschreven.

Tabel 13 **Profiel Experts**

Experts	Titel	Ervaring met usability-onderzoek	Leeftijd	Vooropleiding
E1	Interaction designer, tester	6 jaar	25	HBO
E2	Webspecialist	4 jaar	42	HBO
E3	Interaction Designer	3 jaar	29	WO
E4	Webspecialist	2 jaar	27	HBO
E5	Usability Specialist	2 jaar	23	HBO
E6	Interaction, Webmarketeer	6 jaar	31	WO

4.3.2 Gevonden problemen vanuit heuristieken

Met behulp van de heuristieken hebben de experts 112 absolute problemen gevonden. Een opsomming van enkele uitingen om een beeld te geven van de gevonden problemen:

- Over het gebruiksgemak:
“Daarnaast wordt me niet duidelijk waar 'Mijn Ziggo' voor is. Waarom is het handig voor mij en zou ik er terug moeten komen?”
- Over de zoekfunctie:
“Ik vind stel een vraag aan Tess verwarrend want je gaat direct naar een FAQ. Terwijl je zou denken dat je zou gaan chatten omdat het een persoon is”
- Over de opmaak van de pagina's:
"Ik vind het hele systeem van bladen en widgets maar onduidelijk. Ik moet telkens zoeken om een overzicht van mijn producten of vragen beantwoord te kunnen krijgen.”
- Over het taalgebruik:
“geen goede duidelijk titels van items”

- Over de navigatie:
"Er zijn geen breadcrumbs op de pagina te vinden. Wanneer je ergens op klikt, ga je naar een andere pagina en weet je niet waar je in de site zit."
- Over het inloggen:
"je bent ingelogd en krijgt deze info 'Heb je nog geen Ziggo-account? Schrijf je dan nu in!' dat lijkt me niet goed"

4.3.3 Bijdrage per expert

In dit onderzoek is ook gekeken naar de bijdrage per expert. In tabel 15 is een overzicht te vinden van de bijdrage per expert op het aantal gevonden echte problemen en de gemiddelde mate van ernst van deze problemen.

Tabel 15 Resultaten per expert

experts	E1	E2	E3	E4	E5	E6	M
aantal echte problemen	15	7	10	3	13	11	9,83
gemiddelde ernst	2,20	2,14	2,10	2,00	2,23	2,18	2,14

Van de 66 echte problemen, zijn 29 problemen door zowel de experts als de gebruikers gevonden. Het gemiddelde aantal gevonden echte problemen door de experts is 9,8 met een standaarddeviatie van 2,14. De score van 15 gevonden problemen door expert E1 moet gezien de afwijking van het gemiddelde en de standaarddeviatie als een uitschieter worden beschouwd in positieve zin en de score 3 van expert E4 om dezelfde reden als een uitschieter in negatieve zin. Expert E1 heeft 15 van de 29 van de gevonden echte problemen op zijn naam staan, expert 3 van de 29 gevonden echte problemen. In tabel 16 is de verdeling van de echte problemen weergegeven.

Tabel 16 Echte problemen experts

	echte problemen	M	SD	Hoogste score	laagste score
experts	29	9,8	2,14	15	3

4.3.4 Heuristieken Nielsen en Jameson

De experts hebben de website van Ziggo beoordeeld met behulp van 2 soorten heuristieken, de heuristieken van Nielsen en de heuristieken van Jameson. De vergelijking tussen de resultaten van de experts op beide heuristieken was een punt van onderzoek van deze scriptie. Het totaal aantal unieke problemen dat is gevonden door de experts bedroeg 36.

Hiervan zijn er 6 alleen door de experts gevonden en vertonen 30 overlap met de problemen die ook door de gebruikers zijn gevonden.

Voor het vergelijken van de resultaten van beide heuristieken wordt in deze scriptie uitgegaan van de 29 door de experts gevonden echte problemen.

Om inzicht te krijgen in de mate waarin de problemen gevonden door experts verbonden zijn met de heuristieken van Nielsen en Mack (1994) en Jameson (2007) is gebruik gemaakt van de benadering van Ling en Salvendy (2009). Zij namen als criteria het aantal echte problemen, het aantal false positives, de (gemiddelde) mate van ernst van het probleem, de 'thoroughness', de validiteit en de effectiviteit per expert per set heuristieken. In tabel 17 zijn het aantal echte problemen, het aantal false positive problemen, de gemiddelde ernst, thoroughness, validiteit en effectiviteit bij zowel de heuristieken van Nielsen als Jameson weergegeven.

Tabel 17 Resultaten vergelijking resultaten heuristieken Nielsen en Jameson

	Heuristieken Nielsen	Heuristieken Jameson
Aantal echte problemen	29	7
aantal false positives	7	0
gemiddelde ernst	2,14	1,07
thoroughness	0,43	0,11
validiteit	0,81	1
effectiviteit	0,35	0,11

Van de 29 door de experts gevonden echte problemen zijn er 29 met de heuristieken van Nielsen gevonden en 7 met de heuristieken van Jameson. Het bleek dat van deze 29 echte problemen er 7 zowel gecodeerd waren als gevonden met de heuristiek van Jameson als gevonden met de heuristiek van Nielsen. Met andere woorden het aandeel van Jameson is een deelverzameling van het totaal van Nielsen. In die zin bleek dat de uitbreiding van de heuristieken van Nielsen met de heuristieken van Jameson, in elk geval in dit onderzoek, niet had geleid tot in totaal meer gevonden echte problemen.

Een voorbeeld van gevonden problemen met een heuristiek van Nielsen en een heuristiek van Jameson die in één echt probleem samenvielen : "Er zitten soms wat wel wat moeilijke termen bij voor de verschillende pakketten", gevonden met de heuristiek 'begrijpelijkheid' van Jameson en "Er staan alleen productnamen, die hebben iets meer uitleg nodig, althans sommige." gevonden met de heuristiek 'gebruiksgemak' van Nielsen.

Omdat het slechts om 7 problemen gaat werd een zinvolle bespreking van de andere statistische metingen niet zinvol geacht.

4.3.4.1 Bijdrage per heuristiek per expert

Voor de heuristieken van Nielsen gold dat het gemiddeld aantal gevonden echte problemen per expert 9,8 bedroeg met een maximum van 15 en een minimum van 3. Voor de heuristieken van Jameson bedroeg het gemiddelde 1,6 met een maximum van 6 en een minimum van 0. Twee experts (E2, E5) bleken geen enkel probleem gevonden te hebben dat gekoppeld kon worden aan de heuristieken van Jameson.

Op het niveau van de heuristieken van Nielsen bleek dat er in absolute zin relatief hoog gescoord werd op heuristieken gebruiksgemak en esthetische en minimalistische vormgeving. Op de heuristieken van Jameson, te weten missen informatie i.v.m. het profiel en systeembevoegdheid, is in het geheel niet gescoord. Ook de heuristiek privacy kende slechts 1 gevonden probleem. In tabel 18 is een overzicht te vinden van de bijdrage per expert. Het aantal echte problemen gevonden met de heuristieken van Jameson was echter zo gering dat een vergelijking niet zinvol bleek. Scores op gemiddelde ernst, thoroughness, validiteit en effectiviteit werden betekenisloze getallen.

Tabel 18 analyse van de bijdrage per heuristiek per expert

		E1	E2	E3	E4	E5	E6
Nielsen	aantal echte problemen	15	7	10	3	13	11
	aantal false positives	5	2	2	0	0	0
	gemiddelde ernst	2,20	2,14	2,10	2,00	2,23	2,18
	thoroughness	0,23	0,11	0,15	0,05	0,20	0,17
	validiteit	0,75	0,78	0,83	1,00	1,00	1,00
	effectiviteit	0,17	0,08	0,13	0,05	0,20	0,17
Jameson	aantal echte problemen	6	0	1	0	2	2
	aantal false positives	0	0	0	0	0	0
	gemiddelde ernst	2,00	0,00	2,00	0,00	2,25	2,00
	thoroughness	0,09	0,00	0,02	0,00	0,03	0,03
	validiteit	1,00	0,00	1,00	0,00	1,00	1,00
	effectiviteit	0,09	0,00	0,02	0,00	0,03	0,03

4.3.5 Experts en hoge scores op ernst

Om inzicht te krijgen in het antwoord op de vraag of experts de door hen zelf gevonden problemen gemiddeld een hogere score op ernst toekennen dan de problemen gevonden door de gebruikers werd niet alleen gekeken naar de echte problemen maar ook naar de false positive problemen. De 75 unieke usability problemen zijn door elke expert gewaardeerd met een score op ernst.

Per uniek probleem is een gemiddelde score berekend en deze is vervolgens afgerond op een geheel getal. Om aansluiting te vinden bij de indeling van Nielsen (Nielsen, 1993). Van de 75 unieke problemen werden 6 problemen aangemerkt als false positieve problemen omdat ze enkel gevonden waren door de experts. Daarnaast bleek van 3 problemen dat de score op ernst 0 bedroeg. Een van deze drie was gevonden door de experts.

Van de 66 echte problemen gevonden door gebruikers bleken er in totaal 18 gewaardeerd te zijn met een score 3 (15 problemen) of een score 4 (3 problemen), dit is 27%. De experts hadden 29 echte problemen gevonden waarvan 7 gewaardeerd waren met de score 3 op ernst, dat is 24%. In tabel 19 zijn deze gegevens weergegeven.

Tabel 19 Hoogste scores op ernst

	<u>totaal</u>	<u>ernst 3 en 4</u>	
gebruikers	66	18	27%
experts	29	7	24%

Er is ook gekeken naar de waardering op ernst van de 6 false positive problemen omdat juist deze problemen enkel door de experts zijn gevonden. Eén probleem kreeg de score 3 en vijf problemen de score 2. Ook deze problemen zijn niet extra hoog beoordeeld ten aanzien van de mate van ernst.

4.4 gebruikers en experts

4.4.1 Analyse experts en gebruikers

Voor zowel gebruikers als experts zijn het aantal echte problemen en het aantal false positive problemen vastgesteld en zijn de gemiddelde mate van ernst, thoroughness, validiteit en effectiviteit zijn berekend. Een overzicht hiervan is te vinden in tabel 20.

Tabel 20 Resultaten van gebruikers en experts

gebruikers	G1	G2	G3	G4	G5	G6	G7	G8	G9
aantal echte problemen	17	29	35	21	19	18	12	14	12
aantal false positives	0	1	0	1	0	0	0	0	1
gemiddelde ernst	2,12	2,10	2,06	2,05	2,00	2,06	1,92	1,93	2,08
thoroughness	0,26	0,44	0,53	0,32	0,29	0,27	0,18	0,21	0,18
validiteit	1,00	0,97	1,00	0,95	1,00	1,00	1,00	1,00	0,92
effectiviteit	0,26	0,42	0,53	0,30	0,29	0,27	0,18	0,21	0,17

gebruikers	G10	G11	G12	G13	G14	G15	G16	G17	G18
aantal echte problemen	17	8	13	18	12	15	25	12	20
aantal false positives	0	1	0	0	0	0	0	0	2
gemiddelde ernst	2,18	2,13	1,92	2,28	2,33	1,93	2,16	1,92	2,10
thoroughness	0,26	0,12	0,20	0,27	0,18	0,23	0,38	0,18	0,30
validiteit	1,00	0,89	1,00	1,00	1,00	1,00	1,00	1,00	0,91
effectiviteit	0,26	0,11	0,20	0,27	0,18	0,23	0,38	0,18	0,28
experts	E1	E2	E3	E4	E5	E6			
aantal echte problemen	15	7	10	3	13	11			
aantal false positives	5	2	2	0	0	1			
gemiddelde ernst	2,20	2,14	2,10	2,00	2,23	2,18			
thoroughness	0,23	0,11	0,15	0,05	0,20	0,17			
validiteit	0,75	0,78	0,83	1,00	1,00	0,92			
effectiviteit	0,17	0,08	0,13	0,05	0,20	0,15			

Er zijn t-toetsen uitgevoerd op het gemiddelde en de standaarddeviatie van het aantal echte problemen, het aantal false positive problemen, de gemiddelde ernst, thoroughness, validiteit en effectiviteit van gebruikers en experts. De resultaten in tabel 21 laten zien dat er, met uitzondering van gemiddelde ernst, sprake was van significante verschillen tussen gebruikers en experts.

Tabel 21 Effect van inzet gebruikers cq experts op meerdere statistische criteria

	gebruikers		experts		t-waarde	p-waarde
	M	SD	M	SD		
aantal echte problemen	17,61	6,71	9,83	4,31	2,64	0,015
aantal false positives	0,33	0,59	1,67	1,86	-2,75	0,012
gemiddelde ernst	2,07	0,12	2,14	0,08	-1,32	0,200
thoroughness	0,27	0,10	0,15	0,07	2,57	0,170
validiteit	0,98	0,04	0,88	0,11	3,94	0,001
effectiviteit	0,26	0,10	0,13	0,06	3,03	0,006

Het gemiddelde aantal echte problemen gevonden door de gebruikers (M=17.61, SD=6.71) was significant hoger ($t = 2.64$, $p = 0.015$) dan het gemiddelde aantal echte problemen gevonden door experts (M=9.84, SD=3.10). Dit betekent dat de gebruikers in staat waren om meer echte problemen te vinden dan de experts.

Het gemiddelde aantal false positives gevonden door de gebruikers (M=0.33, SD=0.59) was significant lager ($t = -2.75$, $p = 0.01$) dan het gemiddelde aantal false positives gevonden door de experts (M=1.67, SD=1.86). Daarmee kon geconcludeerd worden dat de gebruikers minder false positives realiseerden dan de experts.

Er was geen significant verschil ($t=-2.31$, $p=0.200$) tussen de gemiddelde mate van ernst van de gevonden problemen door de gebruikers ($M=2.07$, $SD=0.12$) en de experts ($M=2.14$, $SD=0.08$). Dit houdt in dat de problemen gevonden door de gebruikers en de experts gemiddeld even ernstig waren.

De gemiddelde thoroughness van de problemen gevonden door de gebruikers ($M=0.27$, $SD=0.10$) was significant hoger ($t=2.57$, $p=0.017$) dan de gemiddelde thoroughness van de problemen gevonden door de experts ($M=0.15$, $SD=0.07$). Daarmee kan gezegd worden dat de gebruikers beter in staat waren om alle aspecten van de interface te evalueren dan de experts.

De gemiddelde validiteit van de problemen gevonden door de gebruikers ($M=0.98$, $SD=0.04$) was significant hoger ($t=3.94$, $p=0.001$) dan de gemiddelde validiteit van de problemen gevonden door de experts ($M=0.88$, $SD=0.11$). Het inzetten van gebruikers blijkt een betere meetmethode dan het inzetten van experts.

De gemiddelde effectiviteit van de problemen gevonden door de gebruikers ($M=0.26$, $SD=0.10$) was significant hoger ($t=3.03$, $p=0.006$) dan de gemiddelde effectiviteit van de problemen gevonden door de experts ($M=0.13$, $SD=0.10$). Dit geeft aan dat het gebruikersonderzoek een meer optimale manier is voor het vinden van problemen dan het expertonderzoek.

4.4.2 Mate van ernst

Omdat meer inzicht in de mate van ernst en de onderzoeksmethode onderdeel uitmaakte van de vraagstelling van deze scriptie, werd nader ingezoomd op dit onderwerp.

Het gemiddelde aantal problemen met score "1" op ernst, aangeduid als cosmetische problemen, dat werd gevonden door de gebruikers ($M=5$, $SD=4.3$) was niet significant hoger ($t = -1.10$, $p = 0.299$) dan het gemiddelde problemen met score 1 gevonden door de experts ($M=1.5$, $SD=0.71$). Dit betekent dat de gebruikers en experts gemiddeld even veel problemen vonden die alleen hersteld moeten worden als er voldoende tijd beschikbaar is. Het gemiddelde aantal problemen met score 2 op ernst, kleine usability problemen, dat door gebruikers werd gevonden ($M=4.82$, $SD=3.76$) was significant hoger ($t= -2.84$, $p=0.006$) dan het gemiddelde aantal problemen met score 2 gevonden door experts ($M=2.21$, $SD=1.4$). Dit houdt in dat de gebruikers gemiddeld meer problemen vonden waarbij herstel een lage prioriteit heeft, dan de experts.

Het gemiddelde aantal problemen dat met score 3 voor ernst aangemerkt wordt als een groot usability probleem, werd gevonden door de gebruikers (M=5.23, SD=3.06) en was significant hoger ($t = -2.93$, $p = 0.009$) dan het gemiddelde aantal problemen met score 3 gevonden door experts (M=1.71, SD=0.95). Dit geeft aan dat de gebruikers gemiddeld meer problemen vonden waarbij herstel een hoge prioriteit heeft dan experts.

Problemen met de aanduiding dat ze een ramp zijn voor de usability (score 4) werden alleen door de gebruikers gevonden. De hier besproken resultaten zijn weergegeven in tabel 20.

Voor zowel de gebruikers als de experts is per score aangegeven wat de gemiddelde score op de mate van ernst was en welke standaarddeviatie gold.

Tabel 22 Aantal usability problemen met verschillende mate van ernst gevonden door gebruikers en experts

	Ernst score 1		Ernst score 2		Ernst score 3		Ernst score 4	
	M	SD	M	SD	M	SD	M	SD
gebruikers	5	4,3	4,82	3,76	5,23	3,06	3	2,65
experts	1,5	0,71	2,21	1,4	1,71	0,95	0	0
t-waarde		-1.10		-2.84		-2.93		
p-waarde		0,299		0,006		0,009		

Voorbeelden van problemen welke de score '1' toegekend kregen zijn:

- “Er moet veel gescrolld worden om alle informatie op de pagina's te kunnen bekijken.”
- “Facturen zijn direct zichtbaar op Mijn Ziggo. Gebruiker vindt dat privacy in geding komt.”
- “Bezoeker vindt info op Spotlight oninteressant en niet bij Ziggo passen.”

Voorbeelden van problemen welke de score '2' toegekend kregen zijn:

- “Op de website worden termen/iconen gebruikt die de bezoeker niets zeggen. Bijvoorbeeld: Spotlight, enkele iconen van de widgets.”
- “Bezoeker begrijpt niet hoe een onderdeel op de site werkt. Bijvoorbeeld 'eigen blad aanmaken', tv-gids, Tess, functie reclameblokken.”
- “Het hoofdmenu klapt zonder aanwijsbare reden af en toe uit, door de bezoeker wordt dit niet gesnapt en/of gewaardeerd.”

Voorbeelden van problemen welke de score '3' toegekend kregen zijn:

- “Het is de bezoeker niet duidelijk waar hij/zij op de website heen moet en wat hij/zij moet doen. Bezoeker mist sturing.”

- “Website bevat achterhaalde informatie”.
- “Website loopt vast wanneer bezoeker op de bestelpagina of pagina om producten aan te passen, met de navigatie van Internet Explorer terug gaat naar de vorige pagina.”

Voorbeelden van problemen met score ‘4’ op ernst:

- “Gebruiker is niet in staat het huidige pakket aan te passen.”
- “Gebruiker verwacht door design van een icoon of links hierop te kunnen klikken, maar deze is niet aanklikbaar.”
- “Website bevat informatie die gebruiker niet bij Ziggo vindt passen.”

4.4.3 N-corrected effect size

De n-corrected effect size bood de mogelijkheid om het resultaat van de gebruikers en van de experts op thoroughness te vergelijken met een norm die is vastgesteld op basis van eerdere resultaten van gebruikerstests en heuristische evaluaties. Hwang en Salvendy (2009) hebben een meta-analytische benadering toegepast op resultaten van eerder verrichte gebruikerstests en heuristische evaluaties. Dit heeft geleid tot een verband tussen het aantal experts / gebruikers enerzijds en een score op thoroughness anderzijds. Het verband wordt weergegeven door een asymptotische kromme in een grafiek waarin op de horizontale as het aantal ingezette experts / gebruikers staat en op de verticale as een verdeling van thoroughness tussen de waarden 0 en 1. In het hier besproken onderzoek zijn 6 experts ingezet wat zou moeten leiden tot een score van 0,6 op thoroughness. Voor de 18 ingezette gebruikers zou een score van 0,94 mogen worden verwacht. De berekende waarde voor thoroughness vanuit de gevonden waarden bedroeg voor gebruikers 0,92 en voor experts 0,49 wat leidde tot een n-corrected effect size van respectievelijk -0,02 en -0,11. De experts hebben ruim minder unieke problemen gevonden dan op basis van eerdere onderzoeken met 6 experts mocht worden verwacht.

Hoofdstuk 5 Conclusie & discussie

In dit hoofdstuk wordt antwoord gegeven op de subvragen en op de hoofdvraag van dit onderzoek. Daarnaast worden de beperkingen van dit onderzoek en aanbevelingen voor vervolgonderzoek besproken.

Levert het toevoegen van heuristieken specifiek gericht op het vinden van problemen verbonden aan personalisatie, in een expertonderzoek een meerwaarde op ten aanzien van het gevonden aantal problemen en de gemiddelde mate van ernst van de problemen?

De veronderstelling van een mogelijk betere score op gevonden aantal problemen komt voort uit de gedachte dat de toevoeging met de 7 heuristieken van Jameson, zou moeten leiden tot 7 meer mogelijkheden voor de experts om problemen aan te koppelen. De specifieke inhoud van deze heuristieken zou er toe moeten leiden dat er meer problemen zouden worden gevonden die aan sluiten bij de specifieke kwaliteiten van een gepersonaliseerde website. Op basis van de onderzoeksresultaten van het onderzoek in deze scriptie kan worden geconstateerd dat deze veronderstelling niet wordt ondersteund. Van de 66 echte problemen blijken er 29 gevonden te zijn door de experts. Van deze 29 zijn er 29 gevonden met de heuristieken van Nielsen en 7 met de heuristieken van Jameson. Met andere woorden, de 7 problemen gevonden met de heuristieken van Jameson kunnen worden gezien als overlap, er is geen sprake van een uitbreiding van het aantal gevonden problemen. Mogelijk kan in een eventueel vervolgonderzoek eenzelfde populatie afzonderlijk onderzocht worden met de heuristieken van Nielsen en de heuristieken van Jameson om te zien of dat leidt tot het vinden van andere problemen.

Bezien vanuit de individuele resultaten van de experts valt op dat 2 experts in het geheel geen resultaat hebben geboekt met de heuristieken van Jameson, 1 expert 1 probleem meldt, 2 experts 2 problemen melden en 1 expert 6 problemen meldt. Op 2 van de 7 heuristieken is in het geheel niet gescoord. Dit kan de vraag oproepen naar de mate waarin de experts inhoudelijk bekend zijn met de heuristieken van Jameson en / of er eerder mee hebben gewerkt. Mogelijk heeft, al dan niet in relatie tot kennis en ervaring, voor de experts ook de factor tijd een rol gespeeld. Allen hebben de beschikbare tijd van 30 minuten ten volle nodig gehad en aangegeven dat het naar hun mening te weinig was. Dit kan verklaren waarom mogelijk de heuristieken van Jameson minder aandacht hebben gekregen. Deze heuristieken waren in het overzicht van de experts namelijk achter de heuristieken van Nielsen geplaatst. Dit laatste zou, naast het beschikbaar stellen van meer tijd, in een mogelijk vervolgonderzoek eventueel opgelost kunnen worden door voor de helft van de experts deze volgorde om te draaien.

Levert het toevoegen van heuristieken specifiek gericht op het vinden van problemen verbonden aan personalisatie, in een expertonderzoek een meerwaarde op ten aanzien van de 'thoroughness en de validiteit van deze evaluatiemethode?

Een methode is valide als de methode meet wat hij moet meten. De veronderstelling van een mogelijk betere score op validiteit komt voort uit de gedachte dat de uitbreiding met 7 mogelijkheden, de 7 heuristieken van Jameson, zou moeten leiden tot een meer eenduidige koppeling van een probleem aan een heuristiek. De heuristieken van Jameson zouden door hun specifieke inhoud de expert, voor wat betreft het herkennen van problemen, meer op één lijn met de gebruiker hebben moeten brengen. Dit zou tot meer overlap met gebruikers kunnen leiden met als gevolg minder false positives.

Met de heuristieken van Jameson zijn 7 problemen gevonden en het bleken alle echte problemen wat een score van 1 op validiteit opleverde. De validiteit met de heuristieken van Nielsen bedroeg 0,81. Deze scores zouden aanleiding kunnen zijn om de veronderstelling te ondersteunen dat de toevoeging met de heuristieken van Jameson zou leiden tot een hogere score op validiteit. Maar omdat het om overlap ging is de score op validiteit van alle heuristieken ook 0,81 en kan op basis van de resultaten van dit onderzoek de veronderstelling niet worden ondersteund.

De veronderstelde hogere score op thoroughness berustte op de gedachte dat de toevoeging van de heuristieken van Jameson bij het onderzoeken van een gepersonaliseerde website zou leiden tot het vinden van relatief meer echte problemen. In de conclusies is eerder al aangegeven dat dit niet het geval is geweest en zijn mogelijke redenen aangegeven. De 7 echte problemen gevonden met de heuristieken van Jameson zijn ook gevonden met de heuristieken van Nielsen en leidden dus niet tot het uitbreiden van het aantal gevonden unieke problemen en daarmee niet tot het verhogen van de score op thoroughness. Daarmee wordt de veronderstelling door de resultaten in dit onderzoek niet ondersteund.

De effectiviteit wordt gemeten door de thoroughness te vermenigvuldigen met de validiteit. De veronderstelling dat de effectiviteit van de heuristische methode zou toenemen is gebaseerd op de gedachte dat als beide componenten waaruit effectiviteit is samengeteld, validiteit en thoroughness, toenemen ook effectiviteit zou toenemen. Dit is zoals eerder omschreven niet het geval en daarmee wordt de veronderstelling van het toenemen van de effectiviteit door de resultaten van dit onderzoek niet ondersteund.

Blijkt uit de beoordelingen van de echte problemen op mate van ernst dat de problemen die gevonden zijn door experts hoger scoren op ernst dan de problemen gevonden door gebruikers?

De veronderstelling komt voort uit de gedachte dat experts vooral oog hebben voor de meer technische problemen en problemen met een directe aansluiting aan de heuristieken. Zij zouden zich niet goed voor kunnen stellen dat andere problemen van meer invloed zijn op de gebruiksvriendelijkheid van de website voor gebruikers. Het empathisch vermogen van de experts zou tekort schieten (Van der Put, 2006). Bovendien leidt een hogere waardering voor de zelf gevonden problemen mogelijk tot een hogere waardering voor de methode heuristische evaluatie.

Uit de resultaten van dit onderzoek komt naar voren dat de waardering van problemen gevonden door experts en gebruikers niet uiteenloopt. De 3 problemen gewaardeerd met de score 4 op ernst zijn overigens ook gevonden door gebruikers. De 6 problemen die enkel zijn gevonden door experts en daarom als false positive problemen zijn benoemd, hebben bijna alle een waardering 2 op de mate van ernst.

In de resultaten van dit onderzoek is geen steun gevonden voor de veronderstelling.

Levert het onderscheiden van doelgroepen binnen de groep gebruikers een meerwaarde op ten aanzien van het gevonden aantal echte problemen en ten aanzien van de gemiddelde mate van ernst van deze problemen?

De achterliggende gedachte achter deze vraag is de veronderstelling dat de samenstelling van de groep gebruikers van invloed is op het resultaat. Uit de resultaten blijkt dat de groep studenten (55) en de groep ouderen (48) absoluut meer echte problemen gevonden hebben dan de groep werkende mbo-ers (37). In de groep studenten scoren alle personen hoger dan de persoon met de hoogste score in de groep mbo-ers.

De resultaten op de mate van ernst van de gevonden problemen lieten zien dat van de gevonden problemen door de groep mbo-ers bijna 80% was beoordeeld met een score 1 of 2 en 20% met een score 3 en geen enkel score 4. Voor de groepen studenten en ouderen lagen de scores ongeveer gelijk, 70% van de gevonden problemen was beoordeeld met een score 1 of 2 en 30% met een score 3 of 4.

De resultaten in dit onderzoek laten zien dat het onderscheiden van doelgroepen zinvol is omdat een toevallige samenstelling van de groep gebruikers kan leiden tot andere resultaten op aantal gevonden problemen en mate van ernst van deze problemen. Het blijft wel de vraag welke criteria van belang zijn om doelgroepen te onderscheiden. Het was opvallend dat in dit onderzoek alle ouderen hoger opgeleid waren. Vanuit de kenmerken van de gebruikers lijkt daarom in dit onderzoek mogelijk opleidingsniveau van invloed te zijn geweest. In vervolgonderzoek zal hier aandacht aan besteed moeten worden.

Eindconclusie

Met de in dit hoofdstuk genoemde beperkingen in het achterhoofd houdende, kan antwoord gegeven worden op de hoofdvraag; *“Wat is de waarde van het testen door gebruikers en het testen door experts bij het verbeteren van de usability van gepersonaliseerde websites?”*

Op basis van de gevonden resultaten in dit onderzoek blijkt dat gebruikers significant beter scoren op het gevonden aantal echte problemen, op false positive problemen, op thoroughness, op validiteit en op effectiviteit. Enkel op de mate van ernst van de gevonden problemen houden beide methodes elkaar in evenwicht. Dit ondanks de toevoeging van de heuristieken van Jameson aan de heuristische evaluatie. Maar zijn de verschillen nu groter of kleiner nu er onderzoek wordt gedaan met een gepersonaliseerde website?

Op basis van resultaten van gebruikertest en heuristische evaluaties uit het verleden vastgelegd in de besproken grafiek van Hwang & Salvendy (2009), is de voorspelling gedaan dat de score voor thoroughness 0,94 zou zijn voor de gebruikerstest met 18 gebruikers en 0,6 voor de heuristische evaluatie door 6 experts. De veronderstelling was dat deze verhouding zou wijzigen ten gunste van de heuristische evaluatie doordat de toevoeging van de heuristieken van Jameson bij het onderzoeken van een met name gepersonaliseerde website zou leiden tot meer gevonden usability problemen.

Het onderzoek laat als resultaten voor de berekende thoroughness voor experts een score van 0,49 zien en voor de gebruikers een score van 0,92. De vervolgens berekende resultaten op de n -correcte effect size (berekende score minus verwachte score), waren voor experts -0,02 en voor gebruikers -0,11. Dit betekent dat de experts het minder goed hebben gedaan dan mocht worden verwacht en dat de (gebruikers) het volgens verwachting hebben gedaan. Deze uitspraak geldt in verhouding tot het gemiddelde resultaat van heuristische evaluaties c.q. gebruikersonderzoek uit het verleden

Maar deze resultaten geven meer aan. Het verwachte verschil tussen de gebruikers en de experts op thoroughness bedroeg 0,34 (0,94 – 0,6). Het berekende verschil tussen gebruikers en experts op thoroughness bedraagt 0,43 (0,92 – 0,49). Dit betekent dat, in tegenstelling tot

de verwachting, de bijdrage in het vinden van usability problemen door de heuristische evaluatie ten opzichte van de bijdrage van de gebruikerstest is gedaald.

Beperkingen onderzoek

In dit hoofdstuk zijn al kort enige beperkingen besproken. Een aantal punten wil ik kort uitlichten. Bij de interpretatie van deze uitspraak moet rekening gehouden worden met vele facetten die specifiek waren voor dit onderzoek. Zo moet er rekening mee worden gehouden dat de bevindingen die uit dit onderzoek naar voren komen van toepassing zijn op de website van Ziggo. Aangezien iedere gepersonaliseerde website weer andere gepersonaliseerde onderdelen en een andere interface heeft, die in meer of mindere mate gebruiksvriendelijker is dan de in dit onderzoek onderzochte website, gelden de resultaten uit dit onderzoek niet zondermeer voor alle gepersonaliseerde websites.

Er is binnen dit onderzoek gekozen om één website te onderzoeken. Daarmee is het niet mogelijk de resultaten onderling te vergelijken. Er is gewerkt met 6 experts en 18 gebruikers gebaseerd op adviezen vanuit de literatuur. Mogelijk zijn de experts van onvoldoende kwaliteit geweest zeker ten aanzien van het gebruik van de heuristieken van Jameson. Maar het is ook mogelijk dat 6 experts te weinig is. Dit laatste wordt mede ingegeven door de resultaten van Hwang & Salvendy (2009) die voor het behalen van een bepaalde score op thoroughness geen onderscheid maken in het aantal gebruikers of experts. Over de gebruikers kan worden opgemerkt dat ze specifiek zijn gekozen passend bij de veronderstelde doelgroep van de onderzochte website. Daarmee zijn zij mogelijk niet representatief voor doelgroepen van andere gepersonaliseerde websites of gepersonaliseerde websites in het algemeen.

Aanbevelingen vervolgonderzoek

In dit hoofdstuk zijn al kort enige aanbevelingen voor vervolgonderzoek gedaan. Twee punten wil ik kort uitlichten. Voor vervolgonderzoek kan het mogelijk interessant zijn om een onderzoek uit te voeren met zowel 18 experts als 18 gebruikers. Volgens de inzichten van Hwang en Salvendy (2009) zal dit tot een gelijke score op thoroughness moeten leiden. Dit wijkt af van alle eerder in dit onderzoek gepresenteerde theorieën over het aantal te gebruiken experts. Het is daarnaast voor vervolgonderzoek aan te bevelen om niet met beperkende tijd te werken. Dit komt overeen met de opvatting van Hwang en Salvendy (2009) in de bespreking van de onafhankelijke omgevingsvariabelen.

Literatuur

- Anfuso, D. (2006). "Contextual vs. Behavioral Targeting. Gedownload via: <http://www.imediaconnection.com/content/8863.asp> (25 Maart, 2010)
- Bastien, J.M.C., Scapin, D.L. (1995). Evaluating a user interface with ergonomic criteria. *International Journal of Human-Computer Interaction*, 7, 105-121.
- Batra, S., Bishu, R.R. (2007). Web Usability and Evaluation: Issues and Concerns. *Computer Science*, 4559, 243-249.
- Centraal bureau voor de statistiek (2009). De Digitale Economie 2009, Den Haag.
- Da Silveira, G., Borenstein, D., & Fogliatto, F. (2001). Mass customization: Literature review and research directions. *International Journal of Production Economics*, 72 (1), 1-13.
- Desurvire, H. W., Kondziela, J. M., & Atwood, M.E. (1992). What is gained and lost when using evaluation methods other than empirical testing. In *People and Computers VII*, edited by Monk, A., Diaper, D., and Harrison, M. D., 89-102. Cambridge: Cambridge University Press.
- Desuivre, H.W., Lawrence D., & Atwood, M. (1991). Empiricism versus judgement: comparing user interface evaluation methods on a new telephone-based interface. *ACM SIGCHI Bulletin*, 23 (4), 58-59.
- Dumash, J.S., Redish, J.C. (1993). *A practical guide to usability testing*. Norwood, NJ: Ablex.
- Faulkner, L. (2003). Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments, & Computers*, 35(3), 379-383.
- Goh, D., Foo, S. (2008). *Social information retrieval systems: emerging technologies and applications for searching the Web effectively*. Idea Group Inc.
- Haak van den, M., De Jong, M. (2003). Exploring Two Methods of Usability Testing: Retrospective vs. concurrent think-aloud protocols. In: IEEE International Professional Communication Conference, IPCC 2003, 21-24 Sept., Orlando, Florida, USA.
- Hollingshed, T., & Novick, D. G. (2007). Usability Inspection Methods after 15 Years of Research and Practice. Proceedings of the 25th annual ACM international conference on Design of communication, 249-255. New York, NY: ACM.
- Hornbaek, K. (2010). "Dogmas in the Assessment of Usability Evaluation. Methods". *Behaviour & Information Technology*, 29 (1), 97 - 111.

- Hwang, W., Salvendy, G. (2009). Integration of Usability Evaluation Studies via a Novel Meta-Analytic Approach: What are Significant Attributes for Effective Evaluation? *Journal of Human-Computer Interaction*, 25(4), 282-306.
- ISO (1998). Ergonomic requirements for office work with visual display terminals (VDTs)-Part 11: guidance on usability—Part 11: guidance on usability (ISO 9241-11:1998).
- Jameson, A. (2007). *Adaptive interfaces and agents*. DFKI, German Research Center for Artificial Intelligence. International University in Germany, 107-108
- Kantner, L., Rosenbaum, S. (1997). *Usability Studies of WWW Sites: Heuristic Evaluation vs. Laboratory Testing*. SIGDOC 97 Proceedings, (Salt Lake City, UT), published by Association for Computing Machinery, Inc.,
- Karwowski, W. (2006). A Systemic-Structural Theory of Activity International. *Encyclopedia of ergonomics and human factors*, 1, 1893.
- Kincaid, J.W. (2003). *Customer Relationship Management*. Prentice Hall
- Kobsa, A. (2007). Privacy-Enhanced Personalization. *Communications of the ACM*, 50(8), 24-33.
- Kobsa, A., Koenemann, J., & Pohl, W. (2001). Personalised hypermedia presentation techniques for improving online customer relationships. *The Knowledge Engineering Review*, 16(2):111-155.
- Kock, de, E., Biljon, van, J., & Pretorius, M. (2009). Usability evaluation methods: mind the gaps. *Proceedings of the 2009 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists*. Vanderbijlpark, Emfuleni, South Africa ,128
- Kuniavsky, M. (2003). *Observing the User Experience: A Practitioner's Guide to User Research*. San Francisco, California: Morgan Kaufmann Publishers.
- Law, L.C., Hvannberg, E.T. (2002). *Complementarity and convergence of heuristic evaluation and usability test: A case study of universal brokerage platform*. Proceeding of the Second Nordic Conference on Human-Computer /interaction. New York: ACM.
- Ling, C., Salvendy, G. (2008). Effect of evaluators' cognitive style on heuristic evaluation: Field dependent and field independent evaluators. *Human Computer Studies*, 67, 382- 393
- Nielsen, J., Landauer, T.K. (1993). A mathematical model of the finding of usability problems, *Proceedings of the SIGCHI conference on Human factors in computing systems*, 206-21
- Nielsen, J. (1993). *Usability Engineering*. San Diego, California: Academic Press
- Mack, R. L., Nielsen, J. (1993). Usability inspection methods. *ACM SIGCHI Bulletin*, 25 (1), 28–33.

- Mulken, S., André, E., Muller, J. (1998). *The Persona Effect: How Substantial Is It?*
 German Research Center for Artificial Intelligence (DFKI), Saarbrücken.
- Nielsen, J. (1994). Heuristic Evaluation. In: J. Nielsen and R.L. Mack: *Usability Inspection Methods*. John Wiley & Sons, New York, 22-62.
- Nielsen, J. (2000). "Why you only need to test with 5 users". Gedownload via:
www.useit.com/alertbox/20000319.html (15 mei, 2010).
- Nielsen, J. (2001). *Functioneel webdesign*. New Riders, Pearson Education: Amsterdam.
- Norgaard, M., Hornbaek, K. (2006). *What Do Usability Evaluators Do in Practice? An Explorative study of Think-Aloud Testing*. Department of Computer Science.
 University of Copenhagen
- Punselie, R. (2003). *Websites die werken*. Kluwer, 84-87.
- Put, van der, W. (2006). *Website Usability, gebruiksvriendelijke websites in de praktijk*.
 Pearson Education: Benelux.
- Rossi, G., Schwabe, D., Guimaraes, R. (2001). Designing personalized web applications.
Proceedings of the 10th international conference on World Wide Web, 275-284, Hong Kong, Hong Kong
- Sadasivam, R. S., Sundar, G., Tanik, M. M., et al., (2006) .Process personalization framework for service-driven enterprises. *Journal of Integrated Design & Process Science, 12 (2), 31-41*
- Sears, A. (1997). Heuristics walkthroughs: finding the problems without the noise.
International Journal of Human-Computer Interaction, 9 (3), 213-234.
- Sherman, P. (2005). UPA 2005 Member and Salary Survey. Usability Professionals' Association, 5 augustus. Gedownload via: http://www.usabilityprofessionals.org/usability_resources/surveys/2005_upa_salary_survey.pdf
- Synovate (2008). *Ontwikkelingen in de communicatiemarkt*. Adformatie.
- Tan, W. (2003). *A comparison of user testing and heuristic analyses of web sites*.
 Department of Industrial and Management Systems Engineering, University of Nebraska.
- Tan, F., Xu, J., Tung, L. (2009). Attributes of Web Site Usability: A Study of Web Users with the Repertory Grid Technique. *International Journal of Electronic Commerce, 13 (4), 97-126.*
- Tan, D., Bishu, R. (2009). Web evaluation: heuristic evaluation vs. user testing.
International Journal of Industrial Ergonomics, 39 (4), 621-627
- Tullis T., Albert, W. (2008). *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*. San Francisco: Morgan Kaufmann.

- Velsen, van, L. (2008). Usercentered evaluation of adaptive and adaptable systems; a literature review. *The Knowledge Engineering Review*, 23(3), 261 – 281.
- Velsen, van, L. ,König, F.,Paramythis, A. (2009) *Assessing the Effectiveness and Usability of Personalized Internet Search through a Longitudinal Evaluation*. In: 6th Workshop on User-Centred Design and Evaluation of Adaptive Systems, UCDEAS, June 26, Trento, Italy.
- Velsen, van, L.,Geest, van der, T., Klaassen, R. (2010). Identifying usability issues for personalization during formative evaluations: A comparison of three methods. *International Journal of Human-Computer Interaction*.
- Versanen, J. (2007). What is personalization? A conceptual framework. *European Journal of Marketing*, 2007, vol. 41 (5-6), 409-418
- Virzi, R. A. (1992). Refining the test phase of usability evaluation: How many subjects is enough? *Human Factors*, 34 (4), 457-468.
- Vroom, B. (2008). *Websites testen bij gebruikers*. Uitgeverij Kluwer, serie Communicatie Memo
- Wu, D.I., Tremaine, M. Instone, K., et al. (2003). A framework for classifying personalization scheme used on e-commerce websites. *36th Annual Hawaii International Conference on System Sciences (HICSS'03), Hawaii*
- Zaphiris, P. Ang C.H. (2009). *Human Computer Interaction: Concepts, Methodologies, Tools, and Applications*. Idea Group Publishing.

Bijlage 1

Uitleg

Hartelijk dank voor het meewerken aan dit onderzoek. Ik doe dit onderzoek in het kader van mijn afstudeeropdracht: usability-test naar de website www.ziggo.nl

Ik ben een onafhankelijke onderzoeker ben, die niets te maken heeft met het design of de bedrijven van de website www.ziggo.nl.

Ik zou u daarom willen vragen vooral eerlijk te zijn. Ik wil graag precies weten wat u denkt en niet wat u denkt dat ik graag zou willen horen. Uw kunt anoniem deelnemen aan dit onderzoek.

Gedurende het onderzoek geef ik u verschillende opdrachten die u op deze websites moet uitvoeren. Ik zal u observeren terwijl u deze opdrachten uitvoert. Probeer de opdrachten precies zo uit te voeren als wanneer ik hier niet bij aanwezig zou zijn. Ik werk vanuit een script, om ervoor te zorgen dat elk onderzoek dat ik afneem hetzelfde zal zijn.

Tijdens de uit te voeren opdrachten wil ik u vragen hardop te denken. Spreek hardop uit wat u denkt terwijl u bezig bent.

Kunt u uw mobiele telefoon uit zetten? U kunt zich dan geheel op het onderzoek concentreren.

De gehele test duurt in totaal ongeveer 30 minuten.

Heeft u nog vragen?

Bijlage 2

Pre-interview gebruikers

Algemene gegevens

1. Wat is uw leeftijd?
2. Wat is uw hoogst afgeronde opleiding?

Internetervaring

3. Hoeveel uur per week surft u op internet?
4. Wat voor een type internetgebruiker bent u?
Kies uit: Beginner / Gevorderd / Meer dan gevorderd / Expert

Inventarisatie gebruik Ziggo.nl

1. Bent u bekend met ziggo.nl?

Zo ja:

1. Hoe vaak per maand bezoekt u de website ziggo.nl?

Bijlage 3 Opdrachten gebruikers

Homepage			
Testleider geeft de volgende informatie: Ga naar www.ziggo.nl. Wij hebben inloggegevens voor je klaar liggen. Voor dit onderzoek gebruiken we een testaccount. Er staan dus mogelijk andere instellingen in de testaccount dan je gewend bent van je eigen 'Mijn Ziggo', mocht je deze gebruiken.			
Vraag	Opdracht	Opmerkingen / oplossing	Logisch pad
1.	Je wilt een persoonlijk profiel aanmaken. Laat dit zien	Res zijn gang laten gaan en doorvragen op eventuele usability issues.	Van respondent afhankelijk.
Vraag	Opdracht	Opmerkingen / oplossing	Logisch pad
2.	De website bevat een pagina waar persoonlijke gegevens, facturen enzovoort te vinden zijn. Je wilt deze pagina naar eigen wens opmaken. Laat dit zien.	Zijn er functies die niet gebruikt worden?	Van respondent afhankelijk.
Vraag	Opdracht	Opmerkingen / oplossing	Logisch pad
3.	Je hebt een alles-in-1 pakket, maar je wilt een alles-in-1 plus abonnement. Hoe zou je dit doen?		Mijn Ziggo >Mijn producten
Vraag	Opdracht	Opmerkingen / oplossing	Logisch pad
4	Op de website van Ziggo wil je zelf een overzicht van je favoriete tv-zenders selecteren. Hoe zou je dit doen?		Spotlight > complete tv-gids
Vraag	Opdracht	Opmerkingen / oplossing	Logisch pad
5	Je wilt je adres wijzigen, maar je kunt op de website niet vinden hoe dit moet. Hoe zou je dit doen?		Helpfunctie bevindt zich in het hoofdmenu

Bijlage 4

Post-interview gebruikers

De proefpersoon mag de website bekijken, terwijl onderstaande vragen beantwoord worden:

1. Algemeen:

- Welke aspecten van de website zijn u opgevallen in positieve zin?
- Welke aspecten van de website zijn u opgevallen in negatieve zin?

2. Duidelijkheid

- In hoeverre was het duidelijk voor u hoe u een account moest aanmaken voor 'Mijn Ziggo'? Waarom wel/niet?

3. Controle

- In welke mate was het voor u gemakkelijk om de website aan te passen aan uw voorkeuren?
- Kunt u dit toelichten?

4. Helpfunctie

- Heeft u gebruik moeten maken van een zoek- of helpfunctie omdat u opdrachten anders niet had kunnen voltooien? Zo ja, waarom?
- Wat vindt u over het algemeen van de zoekfunctie op de website?

5. Privacy

- Bent u van mening dat er op www.ziggo.nl voldoende rekening wordt gehouden met uw privacy? Zo nee, wat zou er in uw optiek moeten worden aangepast?

6. Informatie onthouden

- In welke mate heeft u tijdens het onderzoek het idee gehad dat bepaalde informatie u niet bereikt heeft omdat deze informatie niet aansluit bij de voorkeuren die u in 'Mijn ziggo' aangegeven heeft?

7. Voorspelbaarheid

- In hoeverre was het u bij elke actie, die u uitvoerde op de website, duidelijk welke informatie in beeld zou verschijnen?
- In hoeverre was het u bij elke actie, die u uitvoerde op de website, op welke pagina u zou uitkomen?

8. Gevolgen onjuist profiel

- Heeft u het idee dat de website van Ziggo ook juist werkt wanneer de persoonlijke gegevens en voorkeuren in 'Mijn Ziggo' niet kloppen?

9. Hinderlijkheid

- Wat vindt u van de hoeveelheid tijd die u kwijt bent om een account aan te maken voor 'Mijn Ziggo' ?

Welke suggesties heeft u verder nog die de website kunnen verbeteren en die nog niet besproken zijn?

Bijlage 5

Uitleg experts

Hartelijk dank voor het meewerken aan dit onderzoek. Ik doe dit onderzoek in het kader van mijn afstudeeropdracht: usability-test naar de website www.ziggo.nl

Ik ben een onafhankelijke onderzoeker ben, die niets te maken heeft met het design of de bedrijven van de website www.ziggo.nl.

Eerst zullen we u een aantal vragen gesteld worden over uw kennis van usability-onderzoeken en uw huidige werkzaamheden.

Na deze opdracht wil ik u verzoeken met behulp van een aantal heuristieken die u van mij uitgereikt krijgt, de website www.ziggo.nl nogmaals te beoordelen. Opmerkingen kunt u op de laptop in een worddocument noteren.

Kunt u uw mobiele telefoon uit zetten? U kunt zich dan geheel op het onderzoek concentreren.

De gehele test duurt in totaal ongeveer 30 minuten.

Heeft u nog vragen?

Bijlage 6

Pre-interview experts

Algemene gegevens

1. Wat is uw leeftijd?
2. Wat is uw hoogst afgeronde opleiding?

Ervaring met usability-onderzoek

3. Welk beroep oefent u uit?
4. Hoeveel jaar houdt u zich met usability-onderzoek bezig?
5. Heeft u ervaring met het gebruik van heuristieken?

Inventarisatie gebruik Ziggo.nl

6. Bent u bekend met ziggo.nl?

Zo ja:

1. Hoe vaak per maand bezoekt u de website ziggo.nl?

Bijlage 7

Verdere uitleg experts

Testleider geeft de volgende informatie:

Ga naar www.ziggo.nl. Wij hebben inloggegevens voor u klaar liggen. Voor dit onderzoek gebruiken we een testaccount. Er staan dus mogelijk andere instellingen in de testaccount dan u gewend bent van uw eigen 'Mijn Ziggo', mocht u deze privé gebruiken.

Ik wil u vragen de website www.ziggo.nl op basis persona A/persona B/persona C (noemen wat van toepassing is) met behulp van de heuristieken, die ik aan u uitreik, te beoordelen.

U kunt één laptop gebruiken om de website op te bekijken. De andere laptop kunt u gebruiken om de gevonden problemen op te noteren. Ik wil u vragen elk geconstateerd probleem op een aparte pagina te documenteren. En bij het probleem aan te geven onder welke heuristiek het probleem valt. Daarnaast zou ik u ook willen vragen de tijdstippen te noteren waarop u gestart bent met het probleem te beschrijven.

Bijlage 8 Heuristische experts

<p>1. Gebruiksgemak</p> <ul style="list-style-type: none"> - De bezoeker van de website krijgt optimaal feedback, snel en op het gewenste moment. - De bezoeker weet wat er gebeurt of wat er is gebeurd, wanneer hij/zij een actie uitvoert op de website.
<p>2. Helderheid</p> <ul style="list-style-type: none"> - De bezoeker begrijpt het taalgebruik van de website en is vertrouwd met de woorden, zinnen en concepten. - De bezoeker krijgt de content op een natuurlijke en in een logische volgorde gepresenteerd
<p>3. Gebruikerscontrole en vrijheid</p> <ul style="list-style-type: none"> - Bezoekers kunnen gebruik maken van een duidelijk herkenbare 'nooduitgang' wanneer zij achteraf gezien een verkeerde keuze hebben gemaakt. - De bezoeker is in staat om m.b.v. navigatie verkeerde handelingen ongedaan te maken of ongedaan gemaakte handelingen opnieuw uit te voeren
<p>4. Consistentie ontwerp gebruikersinterface</p> <ul style="list-style-type: none"> -Bezoekers hoeven zich niet af te vragen of verschillende woorden, situaties of acties dezelfde betekenis hebben, de presentatie is eenduidig en transparant.
<p>5. Foutenpreventie</p> <ul style="list-style-type: none"> - Het ontwerp van de website garandeert dat de bezoeker probleemloos door de website kan navigeren.
<p>6. Herkenning in plaats van herinnering</p> <ul style="list-style-type: none"> - Objecten, acties en opties zijn voor de bezoeker duidelijk herkenbaar, zodat de bezoeker weinig hoeft te onthouden. - De bezoeker hoeft geen informatie te onthouden als hij op zoek naar het gewenste resultaat de verschillende zoekstappen doorloopt. - Instructies over het gebruik van de website zijn duidelijk herkenbaar voor de bezoeker of kunnen makkelijk gevonden worden als de bezoeker dat wenst.
<p>7. Flexibiliteit en efficiëntie</p> <ul style="list-style-type: none"> - De bezoeker hoeft geen onnodige handelingen te verrichten. - Onervaren bezoekers worden op de website anders tegemoet getreden dan ervaren bezoekers - Ervaren bezoekers worden in staat gesteld veel voorkomende handelingen te automatiseren door middel van bijvoorbeeld snelkoppelingen en sneltoetscombinaties.
<p>8. Esthetische en minimalistische vormgeving</p> <ul style="list-style-type: none"> - De bezoeker wordt niet geconfronteerd met overbodige informatie. - Het ontwerp van de gebruikersinterface zorgt er voor dat de bezoeker niet met te veel informatie wordt geconfronteerd. - De vormgeving past bij de bezoeker.
<p>9. Hulp bij herkennen, vaststellen, en herstellen van fouten</p> <ul style="list-style-type: none"> - De bezoeker begrijpt foutmeldingen. - De bezoeker wordt niet onnodig geconfronteerd met "error-codes".

10. Goede helpfunctie en documentatie

- De bezoeker krijgt voldoende informatie aangeboden wanneer hij / zij dit wenst.
- De informatie is op een eenvoudige manier beschikbaar voor de bezoeker.
- De bezoeker kan de website gebruiken zonder documentatie en helpschermen.
- De documentatie is voor de bezoeker overzichtelijk weergegeven.

11. Voorspelbaarheid

De bezoeker is in staat om te voorspellen wat de bijdrage zal zijn van zijn/haar actie aan het te realiseren resultaat.

12. Begrijpelijkheid

De bezoeker is in staat zijn om te begrijpen hoe het systeem werkt ten aanzien van zijn/haar profilering en ten aanzien van de informatie die hij / zij ontvangt

13. Controleerbaarheid

De bezoeker is in staat om controle uit te oefenen op zijn/haar profilering en op het genereren van de gepersonaliseerde output

14. Hinderlijkheid

Een bezoeker is in staat om zijn/haar taak af te maken zonder daarbij gehinderd te worden door bijwerkingen van de personalisatie.

15. Privacy

De bezoeker krijgt niet het gevoel dat de profilering door het systeem, zijn/haar behoefte aan privacy schaadt.

16. Missen informatie i.v.m. profiel

De bezoeker wordt niet de mogelijkheid ontnomen om nieuwe informatie aangeboden te krijgen omdat deze informatie niet past in het van hem/haar vastgelegde profiel

17. Systeem bevoegdheid

De bezoeker krijgt niet de indruk dat het systeem werkt met een onjuist profiel of dat output niet herkenbaar personaliseert.

Bijlage 9

Overzicht geconstateerde problemen m.b.v. heuristieken door experts

Heuristiek	Geconstateerde problemen
1. Gebruiksgemak	
2. Helderheid	
3. Gebruikerscontrole en vrijheid	
4. Consistentie ontwerp gebruikersinterface	
5. Foutenpreventie	
6. Herkenning in plaats van herinnering	
7. Flexibiliteit en efficiëntie	
8. Esthetische en minimalistische vormgeving	
9. Hulp bij herkennen, vaststellen en herstellen van fouten	
10. Goede helpfunctie en documentatie	
11. Voorspelbaarheid	
12. Begrijpelijkheid	
13. Controleerbaarheid	
14. Hinderlijkheid	

Testen gepersonaliseerde websites

15. Privacy	.
16. Missen informatie i.v.m. profiel	
17. Systeem bevoegdheid	

Bijlage 10

Beschrijving Persona A

Ik wil u vragen de website www.ziggo.nl op basis van persona A, met behulp van de heuristieken, die ik aan u uitgereikt heb, te beoordelen. Bij het beoordelen van de website wil ik u vragen te denken aan een gebruiker als persona A:

Persona A (Student)

Persona A studeert en is tussen de 18 en 25 jaar. Zij gebruiken internet dagelijks, zowel op hun laptop/pc als op hun mobiele telefoon. Ze zijn erg handig met internet en gebruiken het voor al hun dagelijkse bezigheden, dit kan zowel op hun studie, thuis, bij hun bijbaan als onderweg zijn.

Bijlage 11

Beschrijving Persona B

Ik wil u vragen de website www.ziggo.nl op basis van persona B, met behulp van de heursistieken, die ik aan u uitgereikt heb, te beoordelen. Bij het beoordelen van de website wil ik u vragen te denken aan een gebruiker als persona B:

Persona B (werkende MBO-ers):

Persona B werkt en is tussen de 30 en 40 jaar. Zij hebben een Mbo-opleiding gevolgd. Zij gebruiken internet op hun werk en thuis af en toe om hun mail te checken of websites te bezoeken.

Bijlage 12

Beschrijving Persona C

Ik wil u vragen de website www.ziggo.nl op basis van persona C, met behulp van de heursistieken, die ik aan u uitgereikt heb, te beoordelen. Bij het beoordelen van de website wil ik u vragen te denken aan een gebruiker als persona C:

PersonaC(ouderen)

Persona C is tussen de 50 en 60 jaar. Zij werken en gebruiken internet voornamelijk voor werk- en hobbygerelateerde zaken. Zij zijn niet opgegroeid met internet en zien internet ook niet als een medium waar ze niet zonder kunnen.

Bijlage 13**Unieke problemen**

1	Bezoeker interpreteert op de website gebruikte termen op een andere manier dan Ziggo deze termen bedoelt heeft. Dit is bijvoorbeeld het geval bij: overstapservice, zenderoverzicht, interactieve tv.
2	Op de website worden termen/iconen gebruikt die de bezoeker niets zeggen. Bijvoorbeeld: Spotlight, enkele iconen van de widgets.
3	Bezoeker denkt met een ander onderdeel van de website bezig te zijn dan werkelijk het geval is. De bezoeker ziet bijvoorbeeld het zenderoverzicht aan voor de tv-gids.
4	Bezoeker begrijpt niet hoe een onderdeel op de site werkt. Bijvoorbeeld 'eigen blad aanmaken', tv-gids, Tess, functie reclameblokken.
5	Sommige keuzes op de site roepen vraagtekens op bij de bezoeker. Voor hen zijn deze keuzes niet logisch. Bijvoorbeeld dezelfde content-blokken die meerdere keren voorkomen, een zoekfunctie die zowel op internet als op de website zoekt
6	Keuzes in de kleurstelling en het ontbreken van accenten zorgen ervoor dat voor de bezoeker onbelangrijke zaken eruit springen en belangrijke zaken nauwelijks opvallen, blokken vallen weg in het ontwerp, teksten zijn lastig te lezen en er is geen onderscheid tussen informatie van Ziggo en andere bronnen.
7	De lay-out is voor de bezoeker onaantrekkelijk. Er staat teveel tekst op de website.
8	De plaats waar ingelogd moet worden is niet voor alle bezoekers duidelijk.
9	Voor de bezoeker staan belangrijke items als 'contact' en 'storingen' niet prominent genoeg op de website.
10	De lay-out van de tv-gids is onoverzichtelijk. Titels zijn niet helemaal zichtbaar en programma's kunnen beter van elkaar onderscheiden worden.
11	De structuur van de website is, o.a. door de losse blokken die te verplaatsen zijn, onduidelijk.
12	Bezoeker snapt niet, of snapt niet waarom, de zoekfunctie binnen internet zoekt in plaats van gewoon op de website van Ziggo.
13	Op de website zijn geen 'breadcrumbs' gebruikt om de bezoeker inzicht te geven waar hij/zij zich op de website bevindt.
14	Er wordt geen gebruik gemaakt van 'hoovers' om de bezoeker bij knoppen, onderdelen of links feedback te geven.
15	Bezoeker ziet op de pagina Mijn Ziggo geen mogelijkheden om deze naar eigen wens op te maken.
16	Bezoeker is niet in staat de tv-gids op de website te vinden.
17	Het is voor de bezoeker niet duidelijk welke informatie wel en welke informatie niet gepersonaliseerd is.

18	Voor de bezoeker is het niet duidelijk hoe de navigatie van het hoofd/submenu werkt.
19	Het is de bezoeker niet duidelijk waar hij/zij op de website heen moet en wat hij/zij moet doen. Bezoeker mist sturing.
20	Gebruiker heeft het gevoel dat bepaalde informatie hem niet heeft bereikt omdat dit niet aansluit bij de voorkeuren in Mijn Ziggo.
21	Bezoeker ziet belangrijke links waarmee hij/zij producten kan bestellen of aanpassen over het hoofd.
22	Gebruiker heeft zoekfunctie moeten gebruiken om onderdelen op de site te vinden.
23	Bezoeker is niet in staat om het huidige pakket dat men bij Ziggo afneemt aan te passen.
24	De website is voor een beginnende internetgebruiker lastig te gebruiken.
25	Het is de bezoeker niet altijd duidelijk welke informatie/pagina in beeld verschijnt wanneer hij/zij een actie uitvoert.
26	Bezoeker snapt de toegevoegde waarde van 'Mijn Ziggo' niet.
27	Bezoeker wil balk met widgets onderaan pagina weg kunnen klikken. Deze balk stoort hem/haar.
28	Het hoofdmenu klappt zonder aanwijsbare reden af en toe uit. Door de bezoeker wordt dit niet gesnapt en/of gewaardeerd.
29	Bezoeker vindt zaken als het weer, verkeersinfo, ANP-nieuws niet op de 'Home-pagina' thuis horen.
30	Bezoeker vindt de reclameboodschappen in de banners vervelend en afleidend.
31	Bezoeker vindt info op Spotlight oninteressant en niet bij Ziggo passen.
32	Website bevat achterhaalde informatie.
33	Dezelfde content/blokken komt op meerdere plekken op de website terug.
34	Er staat te veel informatie op de website en de site oogt erg druk en vol.
35	De zoekfunctie genereert veel te veel zoekresultaten.
36	Er moet veel gescrolld worden om alle informatie op de pagina's te kunnen bekijken.
37	De widgets onderaan de pagina vallen nauwelijks op. Er wordt overheen gekeken.
38	Hulp per onderwerp op de Help-pagina wordt door de grote hoeveelheid tekst en lay-out niet als leesbaar ervaren.
39	Op de productenpagina wordt de reclamebanner als storend ervaren. Men wil direct de producten zien.
40	Info over Ziggo, welke belangrijk is, staat aan de rechterkant van de pagina, terwijl de gebruiker eerst de linkerkant van de pagina scant.
41	Facturen zijn direct zichtbaar op Mijn Ziggo, dit kan leiden tot problemen met privacy.
42	Wanneer de gebruiker de pagina wegklikt en vervolgens opnieuw www.ziggo.nl in typt, is de gebruiker nog steeds ingelogd.

43	De website moet veel laden, dit duurt te lang.
43	Het is onduidelijk welke prive-gegevens er voor anderen zichtbaar zijn en wat het beleid t.a.v. privacy is.
44	Website loopt vast wanneer bezoeker op de bestelpagina of de pagina om producten aan te passen, met de navigatie van Internet Explorer terug gaat naar de vorige pagina.
45	Op een aantal pagina's ontbreekt navigatie om terug te keren naar de vorige pagina.
46	Het komt voor dat blokken blijven laden. Weghalen is niet mogelijk en er verschijnt geen foutmelding of hulp.
47	Zoekfunctie kan ingevoerde zoekwoorden niet vinden.
48	Wanneer men in de banner op de pagina Spotlight op een item klikt, is het niet mogelijk terug te keren naar de openingspagina van Spotlight.
49	Tijdens surfen over de site wordt bezoeker geconfronteerd met melding 'er is een fout opgetreden'. De site loopt vast.
50	Gebruiker verwacht aanpassen van huidige pakket onder 'Producten', terwijl dit onder 'Mijn Ziggo' staat.
51	Gebruiker verwacht de tv-gids niet onder 'Spotlight' maar op een andere pagina.
52	Het is de gebruiker niet duidelijk welke informatie er onder 'opties' te vinden is bij Mijn Ziggo.
53	De website kent weinig call to action.
54	Het is de gebruiker niet duidelijk welke informatie er onder de icoontjes van vraagtekens te vinden is bij Mijn Ziggo.
55	Het is voor de gebruiker niet duidelijk dat het 'bestellen' op de productenpagina voor nieuwe klanten is i.p.v. bestaande klanten.
56	Bezoeker moet zelf handelingen verrichten om Ziggo als startpagina in te richten terwijl de link suggereert dat het met één klik op de link geregeld is.
57	Onduidelijk of men wel of niet ingelogd is, aangezien bijv. bij bestellen gevraagd wordt opnieuw in te loggen, terwijl men al ingelogd is.
58	Inloggen wordt als omslachtig ervaren. Bezoeker verwacht direct op Home een inlogknop waarna op een scherm (evt. pop-up) direct ingelogd kan worden.
59	Website bevat informatie die gebruiker niet bij Ziggo vindt passen.
60	Website bevat achterhaalde informatie
61	Bezoeker verwacht andere informatie onder 'Spotlight' dan nu het geval is.
62	Tess impliceert een applicatie om mee te chatten, maar is een FAQ,
63	Gebruiker verwacht door design van een icoon of links hierop te kunnen klikken, maar deze is niet aanklikbaar,
64	Gebruiker is niet in staat het huidige pakket aan te passen.
65	Gebruiker is niet in staat de zoekfunctie te vinden

66	Wanneer een gebruiker in de tv-gids genres kiest die niet in het getoonde tijdspad staan, verandert er niks aan het scherm (ook geen melding), waardoor bezoeker denkt dat de functie niet werkt.
67	Gebruiker twijfelt over werking zoekfunctie.
68	Wegklikken blokken is duidelijk. Maar hoe komen deze weer terug?
69	Werking van de viewwer is onduidelijk.
70	De blokken functioneren qua navigatie op verschillende manieren, er is hier geen eenheid in te vinden.
71	Snel naar' en 'Handig van Ziggo' worden op Help-pagina als overbodig gezien
72	instructiefilmpjes niet duidelijk vindbaar
73	Gebruiker verwacht naar analogie van andere websites functionaliteiten die er niet blijken te zijn
74	Onduidelijk of het noemen van andere bedrijfsnamen een vorm van sponsering is