

Weet u nog wat u hier dacht?

Een onderzoek naar de bijdrage van observaties en verbalisaties aan het detecteren van problemen en de mate waarin persoonskenmerken en de conditie RTE en RTA daar invloed op hebben.

Evelien van Riel
Studentnummer 3040151
Scriptie Master Communicatiestudies
12 juli 2010

Begeleider: Sanne Elling
Tweede begeleider: Leo Lentz
Universiteit Utrecht

Inhoudsopgave

Samenvatting	4
1. Inleiding.....	5
2. Theoretisch kader	7
2.1 Gebruikersonderzoek met behulp van de hardop denkmethode	7
2.2 Registratie van oogbewegingen	8
2.3 Onderzoek naar de hardop denkmethode met eye tracking.....	9
3. Onderzoeksvraag	12
3.1 Hoofd- en deelvragen	12
3.2 Hypothesen bij deelvraag 1	13
3.3 Hypothesen bij deelvraag 2	13
3.4 Hypothesen bij deelvraag 3 en 4	14
3.5 Hypothesen bij deelvraag 5	14
3.6 Hypothesen bij deelvraag 6	14
4. Onderzoeksmethode	16
4.1 Materiaal	16
4.2 Proefpersonen	16
4.3 Procedure	17
4.4 Analyse	18
4.4.1 Observatie-analyse	18
4.4.2 Probleemdetectie bij observaties	18
4.4.3 Verbalisatie-analyse	19
4.4.4 Probleemdetectie bij verbalisaties.....	19
4.5 Probleemcategorieën	20
4.6 Opzetten probleemsets.....	21
5. Resultaten	24
5.1 Algemene resultaten van de taakuitvoering.....	24
5.1.1 Duur van de taakuitvoering.....	24
5.1.2 Zelfstandigheid beantwoorden vraag	25
5.1.3 Correctheid beantwoorden vraag	25
5.1.4 Invloed persoonlijkheidskenmerken op taakuitvoering.....	26
5.2 Probleemdetectie met behulp van observaties en verbalisaties.....	28
5.2.1 Invloed van RTA en RTE op probleemdetectie.....	29

5.2.2 Invloed van persoonskenmerken op probleemdetectie	30
5.3 Bijdrage van observaties en verbalisaties aan probleemdetectie	32
5.4 Bijdrage van RTA en RTE aan probleemdetectie	34
6. Conclusie	36
7. Discussie	39
8. Literatuur	41
9. Bijlagen	44
Bijlage I	44
Bijlage II	45
Bijlage III	46
Bijlage IV	51

Samenvatting

Om websites te verbeteren en problemen op het spoor te komen, kan gebruikersonderzoek worden ingezet. Zo kunnen proefpersonen bijvoorbeeld aan de hand van een scenario een taak op de website uitvoeren en vlak na of tijdens de taak de gedachten verbaliseren. In dit onderzoek is gekeken naar de retrospectieve hardop denkmethodes RTA en RTE. Onderzocht is wat observaties en verbalisaties bijdragen aan het detecteren van problemen. Daarbij is gekeken of persoonskenmerken als leeftijd, geslacht en opleidingsniveau invloed hebben op de verbalisaties en de observaties. Ook is er onderzocht wat voor problemen proefpersonen in de RTA en RTE conditie detecteren en of daar een verschil te vinden is tussen geobserveerde en geverbaliseerde problemen. Geconcludeerd kan worden dat persoonskenmerken nauwelijks invloed hebben op de taakuitvoering en op de problemen die gedetecteerd zijn met observaties en verbalisaties. Door de hoge datadichtheid van de oogbewegingen zouden er meer problemen moet worden gedetecteerd met observaties (Guan, 2006) maar dat blijkt niet het geval. Problemen met betrekking tot de relevantie van de inhoud, de structuur, het gebruiksgemak en overige problemen worden vaker geverbaliseerd dan dat ze worden geobserveerd. Problemen die vaker worden geobserveerd dan geverbaliseerd, zijn de navigatieproblemen. Er zijn wel unieke problemen gevonden met behulp van observaties in de dimensie bereikbaarheid, maar met verbalisaties worden op iedere subdimensie (inhoud, bereikbaarheid en vormgeving) meer unieke problemen gedetecteerd. Observaties zijn ook minder geschikt voor het signaleren van inhoudsproblemen. Proefpersonen kunnen met verbalisaties heel specifiek aangeven wat voor moeilijkheden ze met de inhoud hebben ervaren. Er zijn wel inhoudsproblemen gevonden met observaties maar deze zijn ook met verbalisaties te detecteren; met observaties komen geen unieke inhoudsprobleem aan het licht. In de RTA conditie zijn meer problemen geverbaliseerd met betrekking tot de vormgeving en de navigatie dan in de RTE conditie. Wanneer een proefpersoon de eigen oogbewegingen niet ziet, is deze in staat om meer navigatie en meer vormgevingsproblemen te verbaliseren. Ook het totaal aantal verbalisaties is hoger in de RTA conditie waardoor het wederom bevestigd lijkt dat oogbewegingen de proefpersoon afleidt tijdens het verbaliseren.

1. Inleiding

Bij het ontwikkelen van websites is het erg belangrijk om rekening te houden met de manier waarop bezoekers de website benaderen. Een website heeft immers vaak als doel de bezoekers van informatie, diensten, producten of amusement te voorzien. Zeker op een informatieve website is het belangrijk dat een bezoeker snel begrijpt hoe de website in elkaar zit en de benodigde informatie snel kan vinden. Gemeentewebsites zijn een bron van informatie, producten, actualiteiten en reglementen en een heldere structuur moet die informatie voor iedere bezoeker goed toegankelijk maken. In praktijk is dat alleen nog niet altijd het geval, het belang van internet voor overheidscommunicatie wordt niet door iedere gemeente erkend (De Jong, Lentz, Elling en Schellens, 2004). Volgens de onderzoekers schuilt het grootste probleem van gemeentewebsite in de gebrekkige toegankelijkheid van de informatie dat te wijten is aan de slechte vormgeving en structuur van de navigatie. Het is dus niet verwonderlijk dat een bezoek aan een gemeentewebsite door de meeste mensen geassocieerd worden met een eindeloos zoekproces vol ergernis en frustratie.

Om websites te verbeteren en problemen op het spoor te komen, kan gebruikersonderzoek worden ingezet. Zo kunnen bezoekers bijvoorbeeld een taak op de website uitvoeren en vlak na of tijdens de taak hun gedachten verbaliseren. De hardop denkmethode geeft de onderzoeker direct inzicht in de gedachten van de proefpersoon; waarom moeizame interpretaties proberen te maken wanneer er beschikking is over een toelichting uit de eerste hand? Er worden alleen nogal wat kanttekeningen geplaatst bij de betrouwbaarheid van de verbalisaties van proefpersonen. Zijn mensen bijvoorbeeld wel in staat om alle gedachten die ze hebben te verwoorden? Daarnaast bepalen mensen zelf wat ze relevant achten om te vermelden en kunnen ze dingen overslaan of vergeten te verbaliseren. Hetgeen ze wel verwoorden staat soms in geen enkel verband met de taakuitvoering, of representeert de taakuitvoering op een foutieve manier.

Om meer inzicht te krijgen in het proces van de taakuitvoering, kunnen de oogbewegingen van de proefpersoon tijdens de taakuitvoering worden geregistreerd. Volgens de *eye-mind hypothesis*, ontwikkeld door Just en Carpenter (1976), bestaat er een nauw verband tussen waar iemand kijkt en waar iemand aan denkt. Zo kan er op een kwantitatieve manier worden vastgelegd hoe een proefpersoon een taak uitvoert. Met behulp van fixaties op bepaalde elementen op een site kan er worden vastgesteld waar iemands aandacht naar uit is gegaan. Een fixatie alleen zegt feitelijk niets over het specifieke cognitieve proces. Als iemand fixeert op een tekstlink kan dit zijn omdat deze

inhoudelijk niet begrepen wordt, niet relevant geacht wordt, omdat de vormgeving opvallend is of omdat de proefpersoon nieuwsgierig is en overweegt om erop te klikken. Het analyseren van oogbewegingen blijft wat dat betreft deels een interpretatieproces. Door zowel de oog- als de muisbewegingen te observeren, ontstaat er een duidelijk beeld van de taakuitvoering; de oogbewegingen kunnen in context worden geplaatst waardoor het kijk- en klikgedrag kan worden geanalyseerd. Verbalisaties van de proefpersoon zelf zijn ook een belangrijke schakel bij het in kaart brengen van het cognitieve proces. Een manier om de toegevoegde waarde van de observaties te vergelijken met de toegevoegde waarde van de verbalisaties, is onderzoeken welk type problemen via de verschillende methoden zijn op te sporen. Zo kan er worden vastgesteld waar de sterktes en zwaktes van de methodes liggen en in welke mate de methodes elkaar afvallen dan wel aanvullen. De hoofdvraag van dit onderzoek is dan ook:

Wat dragen observaties en verbalisaties bij aan het detecteren van problemen? En in welke mate beïnvloeden persoonskenmerken en de conditie RTE en RTA de observaties en verbalisaties?

Om de hoofdvraag te beantwoorden worden in hoofdstuk twee belangrijke theorieën omtrent de hardop denkmethode en eye tracking besproken. Dit hoofdstuk wordt gevolgd door een bespreking van de deelvragen en hypothesen in hoofdstuk 3. Hoofdstuk 4 behelst vervolgens een verantwoording van de onderzoeksopzet en de operationalisering. In hoofdstuk 5 worden allereerst de algemene resultaten over de taakuitvoering gepresenteerd, waarna de specifiekere beantwoording van de hypothesen volgt. Dit alles leidt tot de conclusie die te vinden is in hoofdstuk 6 en wordt afgesloten met de discussie in hoofdstuk 7.

2. Theoretisch kader

2.1 Gebruikersonderzoek met behulp van de hardop denkmethode

Bij het uitvoeren van usability-onderzoek met behulp van gebruikers kan er worden gekozen om proefpersonen tijdens of vlak na de taakuitvoering hun gedachten te laten verbaliseren. Door proefpersonen tijdens de taakuitvoering hardop te laten nadenken, ook wel Concurrent Think-Aloud (CTA) genoemd, wordt de taakuitvoering echter beïnvloed; het tegelijkertijd verbaliseren leidt af van de taakuitvoering en kan ervoor zorgen dat er meer over de te nemen stappen wordt nagedacht dan de proefpersoon normaal gesproken zou doen (Eger, Ball & Dodd, 2007). Dit kan enerzijds leiden tot een betere prestatie dan wanneer de taak zonder tussentijdse overpeinzingen zou worden uitgevoerd, anderzijds kan het de taakuitvoering ook ophouden en leiden tot een slechtere prestatie. Het retrospectieve hardop nadenken, Retrospective Think-Aloud (RTA) wordt daarom door onder andere Guan e.a. (2006) geopperd om deze negatieve aspecten te omzeilen. Bij deze onderzoeksmethode voert de proefpersoon eerst de taak stilzwijgend uit. De muisbewegingen worden tijdens de taakuitvoering opgenomen en na afloop getoond waarbij de proefpersoon de gedachten verbaliseert. Omdat de gedachten pas na de taakuitvoering worden geverbaliseerd, wordt deze niet verstoord en heeft een onderzoeker een ecologisch valide beeld van de taakuitvoering. Het achteraf verbaliseren van de uitgevoerde taak wordt in deze studie onderzocht.

Een belangrijke rol bij het adequaat verbaliseren van gedachten na afloop van de taak, is weggelegd voor het korte termijn geheugen (Ericsson en Simon, 1984). Hoe korter de uitgevoerde taak duurt, hoe beter de proefpersoon in staat is om de zojuist ervaren gedachten te verbaliseren. Toch is er ondanks de voordelen ook veel kritiek op de hardop denkmethode. Omdat er bij het verbaliseren na de taakuitvoering niet alleen uit het korte termijngeheugen maar ook uit het lange termijn geheugen wordt geput, geeft de proefpersoon als het ware een reconstructie van de zojuist uitgevoerde taak. De gedachten die achteraf geverbaliseerd worden, zijn niet per definitie dezelfde gedachten zoals die zich tijdens de taakuitvoering profileerden. Zo kan het bij RTA voorkomen dat zaken worden vergeten of worden overgeslagen omdat de proefpersoon niet alles even relevant acht om te vermelden. Hierdoor kan belangrijke informatie verloren gaan.

Daarnaast zijn er nog praktische bezwaren aan te dragen over het feit dat mensen nu eenmaal minder snel kunnen praten dan ze kunnen denken; het verwerken van informatie gaat sneller dan ze kunnen zeggen. Proefpersonen laten bijna altijd wel een stilte vallen, zodat er altijd informatie

verloren zal gaan. En dan is er nog de vraag of proefpersonen eigenlijk wel in staat zijn om na te denken over hun eigen denken (Baker, 1989) en dit vervolgens te uiten. Individuele karaktertrekken van proefpersonen kunnen roet in het eten gooien, sommige mensen zijn minder goed in staat om een reflectie op het eigen gedrag te geven. Als iemand verbaal niet sterk is, kan het erg moeilijk zijn om te begrijpen waar een opmerking over gaat als het niet zichtbaar is waar deze betrekking op heeft.

2.2 Registratie van oogbewegingen

Een methode die enig houvast biedt in het cognitieve proces is het registreren van oogbewegingen. Oogbewegingen bestaan uit een aaneenschakeling van fixaties en saccades, dit zijn zeer snelle oogbewegingen tussen de fixaties. Door razendsnel op verschillende elementen te fixeren zijn we in staat om scherp te zien (Williams e.a., 2005). De *eye-mind* hypothese van Just en Carpenter (1980) suggereert dat er een sterke samenhang is tussen waar iemand kijkt en waar iemand aan denkt. Door de oogbewegingen vast te leggen, kan het kijkgedrag in kaart worden gebracht. Uit dit kijkgedrag kunnen vervolgens weer verschillende patronen worden afgeleid. Zo wordt het zichtbaar welke elementen mensen bekijken, wat ze niet bekijken, waar ze langer naar kijken, wat ze lezen, wat ze opnieuw lezen of wat ze vluchtig scannen. Op deze manier kan de taakuitvoering op een kwantitatieve manier worden geregistreerd.

Naast het inzichtelijk maken van het gevolgde pad en van de gemaakte keuzes, kunnen oogbewegingen ook worden vergeleken in het onderzoek naar evaluatiemethoden. De uitgevoerde taak wordt opgenomen en de proefpersoon verbaliseert de gedachten achteraf bij het terugzien van de opgenomen beelden van de taakuitvoering. Guan e.a. (2006) hebben met behulp van eye tracking onderzocht of de hardop denkmethode eigenlijk wel valide en betrouwbaar is. Dit hebben ze gedaan door proefpersonen een probleemoplossende taak te laten uitvoeren waarbij de oogbewegingen zijn geregistreerd. Achteraf moeten de proefpersonen bij het zien van de eigen taakuitvoering hun gedachten verbaliseren. Door de oogbewegingen met de verbalisaties na de taakuitvoering te vergelijken, hebben ze vastgesteld dat er maar liefst 80% overlap was tussen de verbalisaties en de observaties. De verbalisaties volgden dezelfde reeks van objecten, in dezelfde volgorde als tijdens de taakuitvoering. Ze stelden een verschil vast in dichtheid van de data tussen oogbewegingen en verbalisaties; met een korte beschrijving zijn vele oogbewegingen samen te vatten. Dit heeft als resultaat dat verbalisaties minder gedetailleerd zijn dan de observaties. Desalniettemin concludeerden de onderzoekers dat het retrospectieve hardop denken een

betrouwbare en valide onderzoeksmethode is om het cognitieve proces in kaart te brengen.

Van der Haak, de Jong en Schellens (2007) stellen dat mensen in een retrospectieve conditie door het gebrek aan visuele stimuli vaker stilvallen en niet kunnen terughalen wat ze ten tijde van de taakuitvoering dachten. In het onderzoek van Van der Haak e.a. zijn drie hardop-denkmethoden met elkaar vergeleken op de website van gemeente Haarlem. De Current Think Aloud methode (CTA); de Retrospective Think Aloud methode (RTA) en de Constructive Interaction methode (Team) waarbij twee personen samenwerken tijdens het uitvoeren van de taak. De onderzoekers vergeleken de geobserveerde problemen met de geverbaliseerde problemen en onderzochten de manier waarop de proefpersonen de methoden ervoeren. Grote verschillen waren er niet, met CTA werden meer usability problemen gevonden, de Team conditie werd als prettigst ervaren door de proefpersonen en de RTA conditie is aan te raden als het belangrijk is dat de taak zo natuurlijk mogelijk verloopt. De onderzoekers verklaren de magere opbrengst van RTA met het feit dat het gros van de handelingen op een informatieve website bestaat uit zoeken en lezen van informatie. Omdat er relatief weinig handelingen worden uitgevoerd, heeft een proefpersoon tijdens het retrospectieve hardop denken weinig stimuli om de taakuitvoering te reproduceren. Om proefpersonen na afloop van de taak meer te stimuleren om het cognitieve proces adequaat te verbaliseren, kan het helpen om ze hun eigen oogbewegingen van de taakuitvoering terug te laten zien. Dit zou proefpersonen moeten voorzien van visuele cues die meer houvast bieden, waardoor de verbalisaties de taakuitvoering meer zouden volgen. Dit wordt Retrospective Think-aloud with Eye-movement (RTE) genoemd.

2.3 Onderzoek naar de hardop denkmethode met eye tracking

Eger, Ball en Dodd (2005), vergelijken het hardop denken tijdens de taak (Concurrent Think Aloud) met retrospectief hardop denken waarbij de helft van de proefpersonen alleen de muisbewegingen ziet (RTA) en de andere helft ziet ook de eigen oogbewegingen terug (RTE). Als testobject gebruiken ze een zoekmachine met een eenvoudige interface (Google) en een zoekmachine die moeilijker te gebruiken is (Infomagnet). Het terugzien van de eigen oogbewegingen geeft volgens de onderzoekers meer inzicht in het moeilijke zoekproces en levert een grotere hoeveelheid gebruikersproblemen op dan in de CTA conditie het geval is. Eger e.a. concluderen daarom dat RTE en RTA voordelen bieden in gebruikersonderzoek, zeker ook vanwege de grotere samenhang in de verbalisaties ten opzichte van de CTA conditie. Of het wel of niet terugzien van de oogbewegingen voordelen biedt bij het opsporen van problemen, laten de onderzoekers in het midden.

In verschillende scriptieonderzoeken is aandacht besteed aan het verschil tussen RTE en RTA. Potappel (2007) heeft een onderzoek uitgevoerd waarin ze onderzocht of proefpersonen meer gedachten weten te verbaliseren wanneer ze de eigen oogbewegingen terugzien. Ze liet 32 proefpersonen twee opdrachten uitvoeren waarna deze hardop moesten denken. De helft van de proefpersonen zag bij het verbaliseren alleen de eigen schermbewegingen zoals de muisbewegingen en clicks (RTA), de andere helft zag ook de eigen oogbewegingen terug (RTE). Er bleken niet meer of andersoortige problemen aan het licht te komen bij de verschillende retrospectieve hardop denkmethoden. In 2009 zijn enkele scriptieonderzoeken (Gombert; Haas; Nell) gewijd aan het opsporen van verschillen in aantal en type verbalisaties die de verschillende methodes opleveren. Ook hier werden geen duidelijke resultaten gevonden, er kwamen slechts marginale verschillen boven water over het aspect op de website waar de verbalisaties betrekking op had. De meeste proefpersonen hebben hun eigen oogbewegingen nog nooit terug gezien en de snelle en misschien onverklaarbare oogbewegingen zouden de proefpersonen meer afleiden dan dat ze echt ondersteuning bieden aan de verbalisaties.

Uit de literatuur blijken dus enerzijds wel degelijk voordelen van het terugzien van de eigen oogbewegingen tijdens het verbaliseren, maar anderzijds zijn er ook nadelen aan te wijzen en het ontbreekt nog steeds aan harde resultaten die aanduiden dat RTE betere resultaten oplevert dan bijvoorbeeld RTA of CTA. Een groot minpunt van de meeste onderzoeken die gedaan zijn naar eye tracking en de hardop denkmethode is de samenstelling van de deelnemers aan het onderzoek. In het onderzoek van bijvoorbeeld Van der Haak e.a. (2007) zijn studenten communicatiekunde ingezet om taken op een website uit te voeren. Dit is een doelgroep die hoogopgeleid is en veelal is opgegroeid met internet. Daardoor heeft deze groep bovengemiddeld veel ervaring met het doorzoeken van een website. Een dergelijke doelgroep is geen correcte representatie van de bezoekers van een website. De manier waarop ouderen en lager opgeleiden een dergelijke opdracht zouden uitvoeren kan verschillen en daarom is het van belang dat dergelijk onderzoek ook gedaan wordt met een heterogene doelgroep. Elling heeft met haar promotieonderzoek daarom gekozen om mensen aan haar onderzoek te laten deelnemen die uit alle lagen van de bevolking komen en van verschillende leeftijdscategorieën zijn. Op die manier ontstaat er een betere afspiegeling van problemen die 'de gemeentewebsite-bezoeker' tegenkomt.

Het promotieonderzoek van Elling 'Evaluatie van gemeentelijke websites', bestaat uit het onderzoeken hoe verschillende evaluatiemethodes zich tot elkaar verhouden. Onderzoeksubject zijn

de gemeentewebsite van Amersfoort, Apeldoorn, Dordrecht en Nijmegen. Daarbij wordt zowel de expertgerichte als de doelgroepgerichte evaluatie onderzocht. Elling leidt het doelgroepgerichte project waar weer vijf evaluatiemethoden onder vallen. Deze methoden kunnen worden onderverdeeld in in-use en non-use en in laboratoriumonderzoek en online onderzoek, te zien in tabel 1.

Tabel 1. Overzicht doelgroepgerichte evaluatiemethoden (Elling e.a. 2005 -2011)

	In-use methoden	Non-use methoden
Laboratoriumonderzoek	Hardop denkmethode Oogbewegingsregistratie	Infocus
Online onderzoek	Feedbackbutton	Vragenlijst

Bij de *non-use* methoden beoordeelt de evaluator de website als testobject, bij de *in-use* methoden wordt de website gebruikt zoals bij dagelijks gebruik het geval is. Bij laboratoriumonderzoek is er sprake van een opgezette testomgeving en bij online onderzoek is de evaluator in de eigen omgeving (www.gemeentewebsites.nl, 2010). Elling gebruikt voor haar onderzoek zowel CTA, RTA als RTE. Naast het uitvoeren van drie taken op een van de gemeentewebsites zijn de scherm- en oogbewegingen geregistreerd en hebben de proefpersonen hardop hun gedachten verwoord. De onderzoeksmethoden, RTE en RTA, die in deze scriptie worden onderzocht komen rechtstreeks uit de dataset van Ellings onderzoek en vallen dus onder de in-use methoden in een opgezette testomgeving. De gebruikersproblemen zijn een gemeenschappelijke deler van de verschillende onderzoeksmethoden; hierdoor wordt het mogelijk om een vergelijking tussen de verschillende methoden te maken. Door te vergelijken welke gebruikersproblemen er met welke onderzoeksmethode boven water komen, kan worden vastgesteld in hoeverre de methoden met elkaar overeenkomen dan wel van elkaar verschillen. Op deze manier wordt in dit onderzoek de RTA met de RTE conditie vergeleken.

3. Onderzoeksvraag

3.1 Hoofd- en deelvragen

Uit de bestaande theorieën blijkt dat de combinatie van oogbewegingen en de hardop denkmethode bij kan dragen aan het opsporen van gebruikersproblemen (Eger e.a., 2005) en kan zorgen voor een specifiek inzicht (Haas; Gombert, 2009). Maar het zien van de eigen oogbewegingen kan ook afleiden tijdens het verbaliseren. In het licht van deze theorieën is het nog niet helemaal duidelijk hoe de gebruikersproblemen die worden gevonden met observaties en verbalisaties zich tot elkaar verhouden. Omdat oogbewegingen specifiek zijn dan verbalisaties, valt te verwachten dat de probleemdetectie bij oogbewegingen ook specifiek is. Oogbewegingen zorgen zoals in het theoretisch kader is uitgelegd voor een grotere datadichtheid en mensen kunnen hun gedachten minder snel verbaliseren dan ze hun ogen kunnen bewegen. Maar is dit beeld ook terug te zien wat betreft probleemdetectie? De hoofdvraag van dit onderzoek is daarom als volgt geformuleerd:

Wat dragen observaties en verbalisaties bij aan het detecteren van problemen? En in welke mate beïnvloeden persoonskenmerken en de conditie RTE en RTA de observaties en verbalisaties?

De hoofdvraag wordt met behulp van zes deelvragen beantwoord:

1. Wat voor type en hoeveelheid problemen wordt met behulp van observaties gedetecteerd?
2. Wat voor type en hoeveelheid problemen wordt aan de hand van verbalisaties gedetecteerd?
3. In hoeverre hebben persoonskenmerken als geslacht, leeftijd en opleidingsniveau invloed op de taakuitvoering?
4. In hoeverre hebben persoonskenmerken als geslacht, leeftijd en opleidingsniveau invloed op het type en de hoeveelheid problemen die wordt gedetecteerd?
5. In hoeverre komen de problemen die enerzijds met observaties en anderzijds met verbalisaties worden gedetecteerd met elkaar overeen?
6. In hoeverre beïnvloedt de conditie RTE of RTA het aantal en type problemen dat gevonden wordt?

3.2 Hypothesen bij deelvraag 1

1. Wat voor type en hoeveelheid problemen wordt met behulp van observaties gedetecteerd?

Door alleen naar oogbewegingen te kijken kan er een heel gedetailleerd beeld ontstaan van de taakuitvoering en de afwijkingen daarvan. Door fixaties, lees- en scanpatronen te interpreteren kan de taakuitvoering heel gespecificeerd in beeld worden gebracht. Zoals Guan e.a. (2006) in hun onderzoek aangaven is de datadichtheid van oogbewegingen groot, waardoor verwacht wordt dat het aantal problemen dat wordt gevonden met behulp van observaties groot is. Wanneer er wordt gekeken naar het type problemen, stelt Vereijken (2007) dat observaties vooral bijdragen aan het krijgen van een gedetailleerd beeld van problemen in de categorieën *bereikbaarheid* en *vormgeving*. Uit de kijkpatronen kan precies worden afgeleid waar proefpersonen moeilijkheden ervaren. Omdat de onderzoekster veel voorkennis had over de mogelijke problemen die zich op de website kunnen voordoen heeft ze problemen in de categorie *inhoud* te snel geïnterpreteerd in plaats van deze alleen te observeren. Ze stelt dat observaties nog wat minder geschikt zijn om specifieke inhoudsproblemen op te sporen, omdat er niet altijd met zekerheid vast te stellen is wat het probleem precies is. De Bode (2008) stelt dat vormgevingsproblemen niet per definitie vast te stellen zijn met observaties. Ze maakt daarbij het verschil tussen de klassieke esthetiek en de expressieve esthetiek; problemen met functionaliteit van vormgeving zijn wel te observeren, problemen met betrekking tot de beleving van de proefpersoon niet.

3.3 Hypothesen bij deelvraag 2

2. Wat voor type en hoeveelheid problemen wordt aan de hand van verbalisaties gedetecteerd?

Vereijken stelt dat verbalisaties beter dan observaties geschikt zijn om specifieke problemen binnen de categorie *inhoud* te detecteren. Bij verbalisaties kan met zekerheid worden vastgesteld wat het probleem is en hoeft er in mindere mate geïnterpreteerd te worden. De Bode (2008) stelt dat usability-problemen die gevonden worden met verbalisaties heel divers zijn. Net als Vereijken concludeert ze dat de meeste problemen met verbalisaties gevonden worden in de categorie *inhoud*. De problemen zijn minder specifiek te lokaliseren dan met eye tracking, maar ze bieden daarentegen een hele goede verklaring voor wat er mis ging. Daarnaast zijn problemen in de categorie *vormgeving* heel goed op te sporen met verbalisaties, volgens De Bode zijn ze door proefpersonen eenvoudig te uiten terwijl ze niet automatisch tot veranderd kijkgedrag leiden.

3.4 Hypothesen bij deelvraag 3 en 4

In hoeverre hebben persoonskenmerken als geslacht, leeftijd en opleidingsniveau invloed op de taakuitvoering en op het type en de hoeveelheid problemen dat wordt gedetecteerd?

Bij de meeste onderzoeken naar gebruiksvriendelijkheid van websites worden studenten als proefpersonen ingezet, waardoor er voornamelijk kennis is over de manier waarop deze groep een website benadert. Nell (2009) onderzocht in haar scriptie in hoeverre verbalisaties in de RTA conditie verschillen van verbalisaties in de CTA conditie. Daarbij heeft ze gekeken of er verschillen zijn aan te wijzen in de verbalisaties en de taakuitvoering waarbij onderscheid gemaakt werd in geslacht, leeftijd en opleidingsniveau. Ze kon geen verschil aanwijzen wat betreft de duur van de taakuitvoering voor geslacht en opleidingsniveau. Wel concludeerde ze dat vijftigplussers meer tijd nodig hadden om de taak uit te voeren dan proefpersonen tussen 18 en 29 jaar. Nell heeft in haar onderzoek alleen de vijf hoofdproblemen opgespoord en daarvoor werden ook geen verschillen gevonden voor de persoonskenmerken.

3.5 Hypothesen bij deelvraag 5

In hoeverre komen de problemen die enerzijds met observaties en anderzijds met verbalisaties worden gedetecteerd met elkaar overeen?

Van den Haak e.a (2007) hebben voor hun onderzoek vergeleken wat voor problemen er kunnen worden opgespoord met verbalisaties, met observaties en met een combinatie van observaties en verbalisaties. Ze concluderen dat er geen significante verschillen kunnen worden waargenomen tussen de hoeveelheid problemen die via de verschillende methoden gevonden worden; alle methoden zijn even vruchtbaar.

3.6 Hypothesen bij deelvraag 6

6. In hoeverre beïnvloedt de conditie RTE of RTA het aantal en type problemen dat gevonden wordt?

Over het verschil in probleemdetectie tussen RTA en RTE is nog weinig bekend. In haar scriptie stelt Gombert (2009) vast dat er geen verschil is aan te wijzen in het aantal hoofdproblemen dat genoemd wordt. Om een algemene verwachting over de retrospectieve conditie op te kunnen stellen is er gekeken naar het onderzoek van Van den Haak e.a. (2007). Daaruit blijkt dat in een retrospectieve conditie meer problemen worden geobserveerd dan geverbaliseerd, dan in een conditie waar tegelijkertijd hardop wordt nagedacht. Proefpersonen gaan volgens de onderzoekers in een retrospectieve conditie eerder scannen waardoor ze minder taken succesvol kunnen

volbrengen dan in een CTA of Team conditie. In de retrospectieve conditie werden significant meer problemen met betrekking tot de volledigheid van de inhoud gevonden dan bij CTA en Team. Te verwachten is dus dat er veel problemen met betrekking tot de volledigheid van de inhoud zullen worden gevonden en dat er meer geobserveerde dan geverbaliseerde problemen zullen worden aangetroffen.

4. Onderzoeksmethode

4.1 Materiaal

Zoals in het theoretisch kader is aangegeven, sluit dit onderzoek aan bij het promotieonderzoek van Elling. Voor het onderzoek naar de in-use-evaluatiemethoden in laboratoriumsetting liet Elling in totaal 180 proefpersonen op een van de websites van de gemeente Dordrecht, Apeldoorn of Nijmegen taken uitvoeren. Op iedere website zijn de taken door twintig proefpersonen voor de drie condities CTA, RTA en RTE uitgevoerd. In dit scriptieonderzoek zijn de data van 60 proefpersonen die meewerkten aan het onderzoek geanalyseerd. Daarvoor zijn 17 RTA en 15 RTE-filmpjes van Dordrecht en 13 RTA en 15 RTE-filmpjes van Nijmegen onderzocht. De verhouding tussen de twee websites is niet precies gelijk vanwege de bruikbaarheid van het materiaal wat betreft de kwaliteit van de oogbewegingen. Op iedere website zijn drie taken uitgevoerd, waarvan er één per website is geanalyseerd, deze taken zijn terug te vinden in de bijlage.

De proefpersonen voeren de taak eerst stilzwijgend uit en geven antwoord wanneer ze de betreffende informatie op de website hebben gevonden. Het zoekproces met alle muis- en oogbewegingen wordt vastgelegd en na afloop van het uitvoeren van alle taken krijgen de proefpersonen de eigen schermbewegingen terug te zien. In de RTA-conditie zien ze de eigen taakuitvoering zonder de eigen oogbewegingen, in de RTE-conditie zien ze de taakuitvoering met de eigen oogbewegingen.

4.2 Proefpersonen

Uniek aan het gebruikte materiaal is de verscheidenheid aan proefpersonen. Waar in de meeste wetenschappelijke onderzoeken naar gebruikersevaluatie studenten worden ingezet, koos Elling ervoor om haar onderzoek uit te voeren met proefpersonen van verschillende leeftijdsgroepen en opleidingsniveaus. Met deze verscheidenheid aan proefpersonen wordt de echte doelgroep van een gemeentewebsite enerzijds meer benaderd en anderzijds wordt er op deze manier waarschijnlijk een grotere verscheidenheid aan gebruikersproblemen opgespoord. De gemaakte verdeling is gelijk over alle evaluatiecondities en ziet er als volgt uit:

- 20 proefpersonen
- leeftijd: 5 personen 18-29 / 5 x 30-39 / 5 x 40-54 / 5 x 55+
- geslacht: 10 x man/ 10 x vrouw

- opleidingsniveau: 6 x laag / 7 x midden/ 7 x hoog

Voor dit onderzoek is een willekeurige selectie van 60 proefpersonen uit het materiaal van Elling genomen waarbij een gelijke verdeling is aangehouden. Binnen de RTA conditie is de gemiddelde leeftijd 39,7 en bij de RTE conditie is dit 41,1 ($t(60) = -.353, p = .726$). Ook het opleidingsniveau en het geslacht zijn gelijk verdeeld over de condities ($t(60) = -.455, p = .651$); ($t(60) = -.776, p = .441$). Voor hun bijdrage aan het onderzoek kregen de proefpersonen een financiële bijdrage.

4.3 Procedure

In 2009 hebben de sessies gedurende drie maanden in het onderzoekslab op het Janskerkhof plaatsgevonden. De retrospectieve sessies duurden ieder een uur. Elling heeft bij de afname van het onderzoek veel aandacht besteed aan de instructie bij het hardop denken. Uit eerder onderzoek bleek dat proefpersonen voornamelijk vertelden wat ze deden en niet zozeer waarom ze dit deden. Om het hardop denkproces te bevorderen werden de proefpersonen van een uitgebreide instructie voorzien met voorbeelden van goede en minder goede uitingen.

Onderzoekers als Boren en Ramey (2000) hebben hun kritiek geuit over onderzoeken waar de hardop denkmethode wordt toegepast. Zij stellen dat iedere onderzoeker zijn eigen manier hanteert om hardop denkonderzoek uit te voeren waardoor het moeilijk wordt om resultaten te vergelijken. Om vergelijking van de verschillende evaluatiemethoden mogelijk te maken, is het belangrijk om als testleider alle evaluaties volgend eenzelfde regime te laten verlopen. Om dit te kunnen bewerkstelligen heeft Elling voor afname conventies opgesteld om proefpersonen te stimuleren verder te praten wanneer ze vijf seconden of langer stil vielen. Ze deed dit door enkel neutrale vragen te stellen als: 'Wat dacht u hier?', 'Probeer u zo veel mogelijk hardop te blijven denken,' 'Weet u nog wat u hier dacht?' en 'En wat dacht u verder?' Wanneer het zoekproces langer dan vijf minuten duurt zonder dat de proefpersoon de benodigde informatie heeft gevonden, wijst de testleider de plek waar de informatie te vinden is. Vanaf die plek krijgt de proefpersoon dan een kans om de taak alsnog te beantwoorden. Na afloop van de drie taken werd er een vragenlijst ingevuld met een aantal evaluatieve vragen en een aantal vragen over persoonlijke gegevens en ervaringen.

4.4 Analyse

4.4.1 Observatie-analyse

Allereerst zijn er analyses van de oog- en klikbewegingen gemaakt zoals deze tijdens de taakuitvoering zijn vastgelegd. Hierbij is gekeken naar de patronen die uit de oogbewegingen kunnen worden afgeleid en op welk element op de website ze betrekking hebben. Om de oogbewegingen te kunnen duiden is er stap voor stap beschreven welke elementen op de pagina gezien worden en op welke manier ze gezien worden. Daarbij is vastgelegd of de oogbewegingen slechts vluchtig of helemaal niet over een bepaald element gaan, op welk element er door de proefpersoon wordt gefixeerd en of bepaalde informatie gescand, gelezen of herlezen wordt. Zo is er een uitgeschreven observatie-route ontstaan. Aan de hand van de observatie-route is een overkoepelende beschrijving gemaakt van de oogbewegingen en van deze overkoepelende beschrijvingen is vervolgens vastgesteld of er sprake is van een probleem. Het aanmerken van problemen is gebaseerd op een persoonlijke interpretatie van gedrag (Lentz en De Jong, 2009). Ook Molic en Dumas (2008) laten in hun onderzoek zien dat verschillende experts tot een heel andere dataset kunnen komen. In de volgende paragraaf wordt daarom precies uitgelegd welke criteria zijn aangehouden.

4.4.2 Probleemdetectie bij observaties

Om aan de hand van observaties vast te kunnen stellen of er een probleem in de taakuitvoering gedetecteerd kan worden, is er gebruik gemaakt van het ideale taakmodel. De taakmodellen van Dordrecht en Nijmegen zijn in de bijlagen na te lezen. In een ideaal taakmodel wordt vastgelegd wat de ideale route is om de benodigde informatie te bereiken. De handelingen die achtereenvolgens moeten worden uitgevoerd om de taak te kunnen beantwoorden als het zien, het lezen en het aanklikken van een bepaald element staan er in uitgeschreven. Op deze manier wordt het precies zichtbaar of iemand het goede pad kiest, of en wanneer iemand van de route afwijkt en wanneer het langer duurt voor er een volgende stap in het model wordt genomen. Er wordt een probleem vastgesteld wanneer aan een van de volgende twee criteria is voldaan:

- De proefpersoon wijkt af van de route van het ideale taakmodel;
- Het duurt langer voor er een volgende stap in het taakmodel wordt genomen.

Aan de hand van het patroon van de oogbewegingen is er vervolgens geanalyseerd waarom er wordt afgeweken van het ideale taakmodel of waarom het zo lang duurt. Deze interpretatie van het probleem wordt ten slotte gecategoriseerd. Hoe de categorisering precies tot stand is gekomen

wordt in paragraaf 4.4.5 uitgelegd.

4.4.3 Verbalisatie-analyse

Naast het documenteren en analyseren van de observaties zijn ook de verbalisaties van de proefpersonen onder de loep genomen. In voorgaande scriptieonderzoeken van Gombert en van Nell (2009) is een groot deel van de verbalisaties al geanalyseerd op samenstelling binnen de RTE en RTA condities. Hierdoor was er beschikking over 38 reeds uitgeschreven transcripten en zijn de verbalisaties van de andere 22 proefpersonen zelf getranscribeerd.

4.4.4 Probleemdetectie bij verbalisaties

Vervolgens zijn de problemen uit de verbalisaties geïdentificeerd. Dit is gedaan door te kijken naar de manier waarop een proefpersoon bepaalde moeilijkheden verbaliseert, zoals het uiten van twijfel, irritatie of ongemak (Van der Haak e.a. 2007). Hier volgen enkele voorbeelden:

- Problemen met de vormgeving:

‘En dan zoek je iets met inschrijven natuurlijk enzo. Maar dat stond dan weer net nergens, of ik had het compleet gemist.’ RTA8, Nijmegen.

‘Dat schuine van de letters vond ik een beetje dwarrelig voor m’n ogen. Beetje vermoeiend lezen, ik moest heel erg zoeken.’ RTA18, Dordrecht.

- Problemen met de navigatie:

‘Nou die woorden, dat digitale, dat was voor mij, ik kijk wel wat ze daar mee bedoelen.’ RTA12, Nijmegen.

‘Vond ik een beetje, ja moeilijk. Ja voortijdig schoolverlater is het niet hè, het is zuiver van de leerplicht.’ RTA1, Dordrecht

- Problemen met volledigheid van de inhoud:

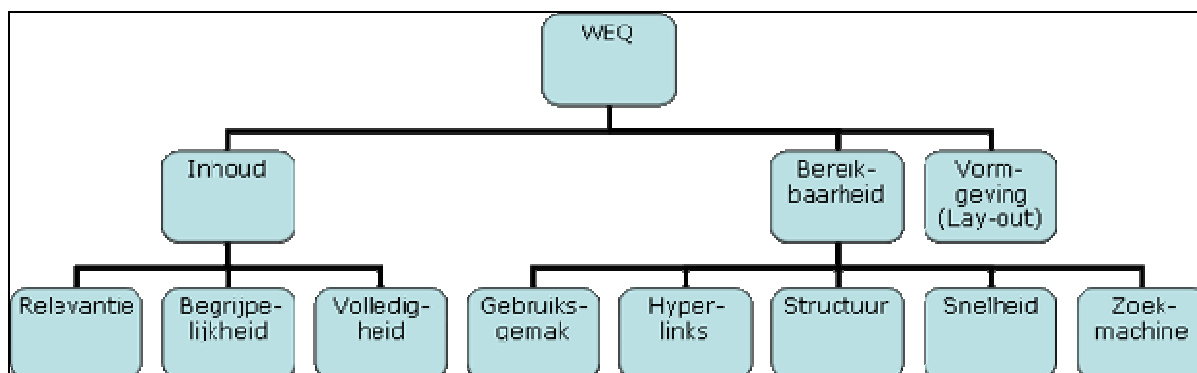
‘Of er dan een link stond voor leerplicht. Of ze vrij konden of niet. Toen dacht ik, is dat het enige wat er staat? Want dan heb ik er niks aan.’ RTE1, Dordrecht.

‘Schoolvakanties, kijken of er nog extra toelichting bij staat toevallig. Over primair onderwijs of zo iets dergelijks. Dat ze daarbuiten ook nog zoveel dagen mogen missen. Maar daar kan ik volgens mij niet echt iets over vinden.’ RTE11, Dordrecht.

4.5 Probleemcategorieën

Om problemen met elkaar te kunnen vergelijken is het belangrijk om deze volgens eenzelfde stramien te categoriseren. De dimensies die bij de evaluatieve vragenlijst voor informatieve websites zijn opgesteld (de WEQ van Elling, Lentz & de Jong, 2007), zijn gebruikt om zowel de observatie- als de verbalisatieproblemen te kunnen categoriseren. Er is gebruik gemaakt van dit bestaande model omdat het aan de hand van literatuuronderzoek is opgesteld en al een aantal keer in voorgaande onderzoeken van andere scriptiestudenten is gebruikt. Op deze manier kunnen de resultaten van dit onderzoek in een groter geheel geplaatst worden en worden vergeleken met de eerdere studies.

De drie hoofddimensies waar de problemen in kunnen worden gecategoriseerd zijn: vormgeving, bereikbaarheid en inhoud. Problemen die te maken hebben met het niet zien van bepaalde elementen als links, tekst of beeld vallen in de dimensie vormgeving. Problemen met betrekking tot de weg naar de informatie toe, vallen binnen de dimensie bereikbaarheid. Problemen die zich voordoen met de reeds gevonden informatie vallen binnen de dimensie inhoud. De dimensies met bijbehorende subdimensies zijn in afbeelding 1 weergegeven.



Afbeelding 1: Dimensies en subdimensies van de WEQ (Website evaluation Questionnaire)

De subdimensies zoekmachine en snelheid zijn niet opgenomen als probleemcategorieën omdat er tijdens het onderzoek niet met de zoekmachine gewerkt is en de website op een adequate snelheid werkt. De categorie *Overig* is toegevoegd aan de subdimensies voor problemen die proefpersonen uiten over andere zaken als bijvoorbeeld de taakuitvoering zelf of wanneer er een probleem was met de functionering van de website. In de volgende paragraaf wordt besproken hoe de problemen zijn gecategoriseerd. In onderstaande tabel is aangegeven welke regels zijn gehanteerd bij het

categoriseren van de problemen.

Tabel 2. Verantwoording categorisering dimensies

Dimensie	Geobserveerd probleem	Geverbaliseerd probleem
Inhoud		
- Relevantie inhoud	Informatie wordt vluchtig gelezen of gescand waarna meteen verder wordt geklikt	Informatie voldoet niet aan de behoeften van de proefpersoon
- Begrijpelijkheid inhoud	Informatie wordt steeds opnieuw bekeken en herlezen	De proefpersoon kan de informatie niet eenvoudig begrijpen
- Volledigheid inhoud	Informatie wordt een langere tijd bekeken waarna toch besloten wordt verder te klikken	De proefpersoon ervaart de informatie niet als volledig, als niet voldoende
Bereikbaarheid		
- Structuur	De proefpersoon blijft maar heen en weer klikken en kijken, is duidelijk aan het zoeken	De proefpersoon ervaart geen structurele samenhang, weet niet waar hij of zij zich op de website bevindt
- Navigatie (hyperlinks in model)	Er wordt op een label geklikt die niets te maken heeft met de route, of de link wordt wel bekeken maar er wordt niet op geklikt	De proefpersoon is niet in staat de betekenis van de links eenvoudig te begrijpen en correct te interpreteren.
- Gebruiksgemak	De website reageert niet zoals de proefpersoon zou verwachten	De site wordt niet als gebruiksvriendelijk ervaren
Vormgeving	Een link, beeld of tekst wordt niet gezien	De vormgeving wordt niet op de bedoelde manier waargenomen
Overig	Gebeurtenissen die niet direct met de taakuitvoering te maken hebben en over de functionering van de website	Gebeurtenissen die niet direct met de taakuitvoering te maken hebben als het aangeven niet meer te weten wat hij of zij dacht

4.6 Opzetten probleemsets

Na het detecteren van alle problemen uit de data van de 60 proefpersonen kan er een probleemset worden opgeleverd zodat de problemen die door observaties en verbalisaties aan het licht komen met elkaar kunnen worden vergeleken. Voor het opzetten van de probleemset is alleen gekeken naar problemen die voor zijn gekomen op pagina's die binnen het ideale taakmodel vallen. Zo blijft er een kern van belangrijkste problemen over die echt betrekking hebben op de taak. Er waren bijvoorbeeld proefpersonen die de pagina van de natuuragenda of van het WMO hebben bekeken. Dit zijn pagina's die niets met een van de beide onderzoekstaken te maken hebben, waardoor ze niet relevant zijn voor de evaluatie van de gebruikte taken. Om een grens te kunnen trekken welke

pagina's relevant genoeg zijn om de taak mee te kunnen evalueren, is daarom gekozen om in de probleemset alleen die problemen op te nemen die op de pagina's binnen het taakmodel voorkomen. Voor Dordrecht zijn daarom de problemen van de volgende pagina's opgenomen in de probleemset:

1. Homepage
2. Onderwijs, werk & welzijn
3. Schoolvakanties
4. Leerplicht

Voor Nijmegen zijn de problemen van de volgende pagina's opgenomen in de probleemset:

1. Homepage
2. Digitale Balie
3. Leven, reizen, papieren/ Bouwen&wonen/ Verhuizen
4. verhuizing

Per website en per pagina is er een probleemset gemaakt die is onderverdeeld in de condities RTE en RTA. Zo kan er worden vastgesteld of het terugzien van de oogbewegingen invloed heeft op de hoeveelheid en het type probleem dat boven water komt. Problemen die identiek zijn maar vaker voorkomen zijn er nauwkeurig uit gefilterd; problemen die over hetzelfde gaan vallen samen onder één probleem. Uiteindelijk zijn de problemen uit de losse sets bij elkaar gezet, waardoor het inzichtelijk wordt hoeveel en wat voor problemen met observaties en verbalisaties kunnen worden gedetecteerd en in welke mate er sprake is van overlap. Hieronder is een voorbeeld weergegeven van de probleemset die is ontstaan voor de homepage van de gemeente Dordrecht, in de RTA conditie.

Voorbeeld 1: Probleemset homepage Dordrecht, RTA

Probleemset Homepage

Dordrecht RTA		
Home-Obs	Home-Verb	Beschrijving probleem
Navigatie1		Ziet benodigde link wel, klikt niet meteen
Vormgeving1	Vormgeving1	Ziet benodigde link niet
	Navigatie2	Opties links en boven hetzelfde
Overig1		Snapt de bedoeling niet
	Vormgeving2	Staan veel verschillende dingen

Uit het voorbeeld blijkt dat er zowel met observaties als met verbalisaties drie problemen zijn gedetecteerd. Er is slechts voor één probleem overlap gevonden, het niet zien van de benodigde link wordt zowel geobserveerd als geverbaliseerd. De andere problemen als het zien van de benodigde link maar er niet op klikken, wordt alleen geobserveerd en niet geverbaliseerd en is daardoor een uniek probleem dat alleen geobserveerd is. In het volgende hoofdstuk worden de resultaten van de analyses gepresenteerd.

5. Resultaten

In dit hoofdstuk worden allereerst de algemene bevindingen omtrent de taakuitvoering besproken. Vervolgens wordt er specifiek gekeken wat voor problemen er zijn geobserveerd en geverbaliseerd. Hierbij wordt gekeken of de conditie en de persoonskenmerken invloed hebben op de problemen die enerzijds met observaties en anderzijds met verbalisaties zijn gedetecteerd. De resultatensectie wordt afgesloten met het analyseren van de probleemsets, waarmee kan worden vastgesteld wat observaties en verbalisaties bijdragen aan de probleemdetectie.

5.1 Algemene resultaten van de taakuitvoering

De algemene analyse van de taakuitvoering bestaat uit het correct kunnen beantwoorden van de vragen, het zelfstandig kunnen vinden van de benodigde informatie en de duur van de taak. Deze drie verschillende aspecten van de taakuitvoering zijn vergeleken per conditie en per website. Ook wordt er gekeken of persoonlijkheidskenmerken als leeftijd, geslacht en opleidingsniveau invloed hebben op de drie aspecten van de taakuitvoering.

5.1.1 Duur van de taakuitvoering

Allereerst is er gekeken hoe lang het gemiddeld duurt om de taak te volbrengen, te zien in tabel 3. Met behulp van een ANOVA-toets is vastgesteld dat er geen verschil is in de taak in zowel de RTA als de RTE conditie even lang duurde; gemiddeld tussen de 4,5 en 5,5 minuut.

Tabel 3. Duur van de taakuitvoering (in seconden) per conditie

Conditie	Gemiddelde duur	St.dev
RTA	320.46	144.8
RTE	280.21	96.2

Omdat er twee verschillende websites zijn geanalyseerd is er ook gekeken naar de tijdsduur van de twee verschillende taken, tabel 4. Er blijkt een significant verschil ($F(46) = 38.9 ; p < .001$) tussen de twee websites. Dat betekent dat de taak op de website van de gemeente Nijmegen langer duurde dan die van de gemeente Dordrecht.

Tabel 4. Duur van de taakuitvoering (in seconden) per website

Website	Gemiddelde duur	St.dev
Dordrecht	239.4	77.6
Nijmegen	418.6	119.4

5.1.2 Zelfstandigheid beantwoorden vraag

Er is ook onderzocht in welke mate de proefpersonen in staat zijn om de taak zelfstandig uit te voeren. Wanneer proefpersonen meer dan vijf minuten op de website rondzwierven zonder de goede route te hebben gevonden kregen ze hulp van de testleider. Deze wees dan waar ze het goede antwoord konden vinden. In tabel 5 is de mate van zelfstandigheid per conditie terug te zien, met behulp van een ANOVA-toets blijkt deze per conditie niet significant te verschillen.

Tabel 5. Mate van zelfstandigheid in beantwoorden vraag per conditie (0= niet zelfstandig, 1= zelfstandig)

Conditie	Zelfstandigheid	St.dev
RTA	.67	.5
RTE	.83	.4

Er blijkt wel een significant verschil te zijn in de mate van zelfstandigheid tussen de twee websites. De taak op de website van de gemeente Dordrecht werd vaker zelfstandig uitgevoerd ($F(59) = 10.14 ; p < .01$), te zien in tabel 6.

Tabel 6. Mate van zelfstandigheid in beantwoorden vraag per website (0= niet zelfstandig, 1= zelfstandig)

Website	Zelfstandigheid	St.dev
Dordrecht	.91	.3
Nijmegen	.57	.5

5.1.3 Correctheid beantwoorden vraag

Verder is er nog gekeken of er een verschil is in de mate waarin de vragen correct zijn beantwoord. In tabel 7 is te zien dat de vragen in zowel de RTA als de RTE-conditie even goed worden beantwoord. Uit tabel 8 blijkt verder dat er geen verschil is waar te nemen in de mate waarin de vragen op de twee websites correct zijn uitgevoerd.

Tabel 7. Mate van correctheid in beantwoorden vraag per conditie (0= beide vragen fout, 1= deels goed 2=goed)

Conditie	Correctheid	St.dev
RTA	1.57	.6
RTE	1.57	.6

Tabel 8. Mate van correctheid in beantwoorde vraag per website (0= beide vragen fout, 1= deels goed 2= goed)

Website	Correctheid	St.dev
Dordrecht	1.69	.6
Nijmegen	1.43	.5

5.1.4 Invloed persoonlijkheidskenmerken op taakuitvoering

Ook is er gekeken of persoonlijkheidskenmerken als leeftijd, geslacht en opleidingsniveau invloed hebben op de drie aspecten van de taakuitvoering. Achtereenvolgens wordt besproken of er verschillen zijn voor de persoonskenmerken wat betreft de duur van de taakuitvoering, gevolgd door de mate van zelfstandigheid en de mate van correctheid. Uit tabel 9, 10 en 11 blijkt achtereenvolgens dat de persoonskenmerken niet zorgen voor verschillen wat betreft duur van de taakuitvoering.

Tabel 9. Duur van de taakuitvoering, gekeken naar geslacht

Geslacht	Gemiddelde duur	St.dev
Man	327.27	125.3
Vrouw	263.57	125.4

Tabel 10. Duur van de taakuitvoering, gekeken naar opleidingsniveau

Opleidingsniveau	Gemiddelde duur	St.dev
Laag	302.69	83.5
Midden	293.55	163.4
Hoog	310.13	134.6

Tabel 11. Duur van de taakuitvoering, gekeken naar leeftijdscategorie

Leeftijdscategorie	Gemiddelde duur	St.dev
18-29	269.77	114.9
30-39	315.9	130.1
40-54	285.5	98.2
55+	350.42	127.8

Ook wat betreft zelfstandigheid zijn er geen verschillen gevonden voor de persoonskenmerken. Uit de analyses blijken alle groepen dezelfde mate van zelfstandigheid te vertonen. Dit blijkt uit de tabellen 12, 13 en 14.

Tabel 12. Mate van zelfstandigheid bij uitvoeren taak, gekeken naar geslacht, (0= niet zelfstandig, 1= zelfstandig)

Geslacht	Zelfstandigheid	St.dev
Man	.74	.4
Vrouw	.76	.4

Tabel 13. Mate van zelfstandigheid, gekeken naar opleidingsniveau (0= niet zelfstandig, 1= zelfstandig)

Opleidingsniveau	Zelfstandigheid	St.dev
Laag	.69	.5
Midden	.73	.5
Hoog	.79	.4

Tabel 14. Mate van zelfstandigheid, gekeken naar leeftijdscategorie (0= niet zelfstandig, 1= zelfstandig)

Leeftijdscategorie	Zelfstandigheid	St.dev
18-29	.82	.4
30-39	.75	.5
40-54	.65	.5
55+	.79	.4

Ten slotte is er voor de verschillende persoonskenmerken gekeken of er verschil is in de mate van correctheid van de gegeven antwoorden, de gemiddeldes zijn te zien in de tabellen 15, 16 en 17.

Voor geslacht en opleidingsniveau werden geen verschillen gevonden in de mate waarin een correct antwoord werd gegeven, maar er is wel een verschil gevonden tussen de leeftijdscondities, ($F(59) = 3.63$; $p < .05$). Proefpersonen tussen de 40 en 54 jaar hebben de vragen significant slechter beantwoord dan de proefpersonen uit de andere categorieën.

Tabel 15. Mate van correctheid van uitvoeren taak, gekeken naar geslacht, (0= beide vragen fout, 1= deels goed 2= goed)

Geslacht	Correctheid	St.dev
Man	1.46	.6
Vrouw	1.72	.5

Tabel 16. Mate van correctheid, gekeken naar opleidingsniveau (0= beide vragen fout, 1= deels goed 2= goed)

Opleidingsniveau	Correctheid	St.dev
Laag	1.56	.6
Midden	1.8	.5
Hoog	1.6	.6

Tabel 17. Mate van correctheid antwoord, gekeken naar leeftijdscategorie (0= beide vragen fout, 1= deels goed 2= goed)

Leeftijdscategorie	Correctheid	St.dev
18-29	1.71	.5
30-39	1.83	.4
40-54	1.24	.7
55+	1.57	.5

5.2 Probleemdetectie met behulp van observaties en verbalisaties

In deze paragraaf worden eerst de algemene resultaten besproken; wat voor type problemen kunnen met behulp van observaties worden gedetecteerd en wat voor type problemen met behulp van verbalisaties. Hierbij is nog niet gekeken of er overlap is tussen de problemen die met observaties en verbalisaties worden opgespoord, een specifieke analyse hiervan volgt in paragraaf 5.3. Vervolgens is er gekeken in welke mate de conditie, RTA en RTE, en de persoonskenmerken invloed hebben op de detectie van deze problemen. De problemen zijn ingedeeld in de categorieën: relevantie inhoud, volledigheid inhoud, begrijpelijkheid inhoud, structuur, navigatie, gebruiksgemak, vormgeving en overig. Met behulp van een ANOVA-toets is gekeken of problemen in de ene categorie vaker voorkomen dan anderen voor de observaties en de verbalisaties, in tabel 18 is per categorie aangegeven hoe vaak de categorieën gemiddeld per proefpersoon voorkomen.

Tabel 18. Gemiddeld aantal observaties en verbalisaties per proefpersoon

Categorie	Observaties	Verbalisaties
Relevantie inhoud	2.23 (2.5)	.97 (1.1)
Begrijpelijkheid inhoud	.73 (.8)	.77 (1)
Volledigheid inhoud	1.88 (1.7)	1.62 (6)
Structuur	.82 (1)	1.23 (1.2)
Navigatie	4.22 (3)	1.08 (1.5)
Gebruiksgemak	.07 (.3)	.62 (1)
Vormgeving	2.03 (1.4)	2.68 (2.7)
Overig	.43 (.67)	.93 (1.3)
Totaal (gemiddeld)	12.48 (7.4)	9.92 (6.1)

In de tabel is te zien dat de meeste geobserveerde problemen, gemiddeld 4.22, gevonden worden in de categorie navigatie. De meeste geverbaliseerde problemen worden gevonden in de categorie vormgeving, gemiddeld 2.68 per proefpersoon. Met behulp van de gepaarde t-toets zijn een aantal significante verschillen aan te wijzen tussen de twee evaluatiemethoden. Zo wordt er in de dimensie *inhoud* significant vaker geobserveerd dat er problemen zijn met de relevantie van de inhoud dan dat deze geverbaliseerd worden, ($t(59) = -3.475, p < .001$). In de dimensie *bereikbaarheid* wordt er significant vaker geverbaliseerd over structuurproblemen dan dat deze worden geobserveerd, ($t(59) = -2.472, p < .05$). Navigatieproblemen worden vaker geobserveerd dan dat ze geverbaliseerd worden en problemen met het gebruiksgemak worden weer vaker geverbaliseerd. Problemen die niet direct met de taakuitvoering te maken hebben, de categorie *overig* wordt vaker geverbaliseerd dan geobserveerd, ($t(59) = -3.305, p < .01$).

5.2.1 Invloed van RTA en RTE op probleemdetectie

Allereerst is er gekeken of er een verschil aan te tonen is tussen de condities RTA en RTE wat betreft de observaties en verbalisaties. De gemiddeldes zijn weergegeven in tabel 19. Algemeen beschouwd is er op één categorie na geen verschil aan te wijzen tussen de problemen die via observaties gedetecteerd worden in de RTA en in de RTE condities. Alleen problemen met betrekking tot het gebruiksgemak zijn vaker aangetroffen in RTA dan in RTE, ($t(58) = 2.112, p < .001$).

Tabel 19. Gemiddeldes observaties en verbalisaties per categorie per conditie

Conditie Categorie	Observaties		Verbalisaties	
	RTA	RTE	RTA	RTE
Relevantie inhoud	2.6 (2.8)	1.87 (2)	1.07 (1.2)	.87 (1)
Begrijpelijkheid inhoud	.63 (.9)	.83 (.7)	.73 (1)	.8 (1)
Volledigheid inhoud	1.63 (1.5)	2.13 (1.9)	1.57 (1.8)	1.67 (1.5)
Structuur	.83 (1.1)	.8 (.8)	1.5 (1.3)	.97 (1.2)
Navigatie	4.3 (3.1)	4.13 (3)	1.37 (1.9)*	.8 (.9)*
Gebruiksgemak	.13 (.3)*	0 (0)*	.63 (1)	.6 (1)
Vormgeving	1.77 (1.2)	2.3 (1.5)	3.1 (3.4)*	2.27 (1.7)*
Overig	.5 (.7)	.37 (.6)	.93 (1.5)	.93 (1.1)
Totaal (gemiddeld)	12.37 (7.9)	12.6 (7)	10.93 (7.5)**	8.9 (4.2)**

* $p < .001$ ** $p < .05$

Dat wil zeggen dat er bij de proefpersonen in de RTA en de RTE conditie via observaties vergelijkbare problemen zijn gedetecteerd tijdens het uitvoeren van de taak. Er zijn iets meer verschillen gevonden tussen de condities bij de geverbaliseerde problemen. Zo worden er significant meer navigatieproblemen en problemen met betrekking tot de vormgeving geverbaliseerd in de RTA conditie dan in de RTE conditie, ($t(58) = 1.467, p < .05$), ($t(58) = 1.199, p < .001$). Ook is er een verschil gevonden in het totale aantal verbalisaties, bij RTA wordt significant meer geverbaliseerd dan in RTE, ($t(58) = 1.294, p < .01$).

5.2.2 Invloed van persoonskenmerken op probleemdetectie

Dan is er nog gekeken of persoonskenmerken als geslacht, opleiding en leeftijd invloed hebben op het aantal en type problemen dat wordt geverbaliseerd dan wel geobserveerd. Eerst worden de resultaten van de observaties gepresenteerd in tabel 20, waarna de resultaten van de verbalisaties volgen in tabel 21.

Tabel 20. Gemiddeldes observaties per categorie voor de persoonskenmerken

Persoonsken- merken Categorie	Observaties								
	Geslacht		Leeftijdscategorie				Opleidingsniveau		
	Man	Vrouw	18-29	30-39	40-54	55+	Laag	Midden	Hoog
Relevantie inhoud	2.71 (2.8)	1.56 (1.7)	2.76 (2.7)	1.75 (2)	3 (3.1)	1.07 (1)	3.06 (2.9)	2.47 (2.7)	1.66 (1.9)
Begrijpelijkheid inhoud	.71 (.8)	.76 (.8)	.41 (.5)	1.17 (.8)	.71 (.7)	1.07 (1)	.81 (.8)	.4 (.5)	.86 (.9)
Volledigheid inhoud	2.11 (1.8)	1.56 (1.4)	1.82 (2)	1.75 (1.7)	1.71 (1.5)	2.29 (1.6)	1.69 (1.4)	1.8 (2.3)	2.03 (1.5)
Structuur	1.06* (1.1)	.48* (1)	.82 (1)	1.08 (1.4)	.82 (.9)	.57 (.8)	.81 (.9)	.87 (.9)	.79 (1.1)
Navigatie	4.63 (3.2)	3.64 (2.8)	4.59 (2.7)	3.33 (2.3)	5.12 (4)	3.43 (2.2)	5.25 (3.6)	4.27 (3.3)	3.62 (2.4)
Gebruiksgemak	.09 (.3)	.04 (.2)	.24* (.4)	0* (0)	0* (0)	0* (0)	0 (0)	.13 (.4)	.07 (.3)
Vormgeving	2.26 (1.4)	1.72 (1.3)	1.88 (1.1)	2.25 (1.1)	2.59 (1.2)	1.36 (1.7)	2.5 (1)	2.33 (1.2)	1.62 (1.5)
Overig	.43 (.6)	.44 (.8)	.24 (.4)	.5 (.8)	.47 (.8)	.57 (.6)	.38 (.5)	.33 (.5)	.52 (.8)
Totaal (gemiddeld)	14.11* (7.9)	10.2* (6)	12.82 (8.1)	11.83 (5.7)	14.41 (9)	10.29 (5.7)	14.38 (8)	12.67 (9.1)	11.34 (6)

* $p < .05$

In de tabel is te zien dat er geen enorme verschillen zijn gevonden tussen de geobserveerde problemen voor de verschillende persoonskenmerken. Voor het persoonskenmerk geslacht is er voor één probleemcategorie een verschil gevonden. Bij mannen is er significant vaker een structuurprobleem geobserveerd dan bij vrouwen het geval was, ($F(59) = 5.205$; $p < .05$). Gekeken naar het gemiddelde van het totaal aantal geobserveerde problemen blijkt er ook een significant verschil te zitten tussen mannen en vrouwen, bij mannen worden er significant meer problemen geobserveerd dan bij vrouwen, ($F(59) = 4.295$; $p < .05$). Als er gekeken wordt naar het verschil in leeftijd wordt er enkel voor de categorie gebruiksgemak een significant verschil gevonden. Alleen bij de proefpersonen in de leeftijdscategorie van 18-29 jaar zijn problemen met betrekking tot gebruiksgemak geobserveerd en bij de andere categorieën was dit niet het geval, ($F(59) = 4.116$; $p < .05$). Voor het persoonskenmerk opleidingsniveau werden geen significante verschillen gevonden.

Vervolgens is er gekeken naar de invloed die de persoonskenmerken hebben op de problemen die gedetecteerd worden met behulp van verbalisaties, de gemiddeldes zijn te zien in tabel 21.

Tabel 21. Gemiddeldes observaties per categorie voor de persoonskenmerken

Persoonsken- merken Categorie	Verbalisaties								
	Geslacht		Leeftijdscategorie				Opleidingsniveau		
	Man	Vrouw	18-29	30-39	40-54	55+	Laag	Midden	Hoog
Relevantie inhoud	1.03 (1.2)	.88 (.8)	1.12 (1.2)	.67 (.8)	1.29 (1.4)	.64 (.7)	1.19 (.8)	1.2 (1.5)	.72 (.96)
Begrijpelijkheid inhoud	.8 (1.1)	.72 (.8)	.65 (1)	1.17 (1.2)	.76 (.9)	.57 (.9)	.94 (1.1)	.73 (.8)	.69 (1)
Volledigheid inhoud	1.63 (1.5)	1.6 (1.8)	1.53 (1.60)	2.17 (1.6)	1.53 (1.8)	1.36 (1.4)	1.81 (1.5)	1.27 (1.6)	1.69 (1.7)
Structuur	1.43 (1.4)	.96 (1)	1.24 (1.2)	1.08 (1)	1.35 (1.4)	1.21 (1.4)	1.25 (1.1)	1.53 (1.6)	1.07 (1.2)
Navigatie	1.06 (1.2)	1.12 (1.9)	1.12 (1.2)	.5 (.8)	1.12 (1.3)	1.5 (2.3)	1.13 (1.2)	1.2 (2.4)	1 (1.1)
Gebruiksgemak	.69 (.9)	.52 (1)	.71 (1)	.58 (.5)	.41 (.9)	.79 (.6)	.44 (.6)	.6 (1.3)	.72 (1)
Vormgeving	2.91 (2.8)	2.36 (2.6)	3.24 (2.7)	2.92 (2.9)	2.47 (2.6)	2.07 (2.7)	2.63 (2.7)	2.33 (2.8)	2.9 (2.7)
Overig	1 (1.4)	.84 (1.2)	1.12 (1.4)	1.08 (1.8)	1.06 (1.3)	.43 (.6)	.69 (.9)	1.2 (1.1)	.93 (1.6)
Totaal (gemiddeld)	10.54 (5.8)	9.04 (6.5)	10.71 (5.2)	10.17 (5.6)	10 (6,5)	8.64 (7.5)	10 (4.5)	10 (7.7)	9.83 (6.2)

Met behulp van de ANOVA-toets is er voor geen van de variabelen een verschil gevonden. Voor geen van de persoonskenmerken kan worden gesteld dat er een bepaalde categorie probleem vaker wordt geformuleerd dan een andere. Het maakt voor het geformuleerde aantal en type probleem dus niet uit welk geslacht, welke leeftijd of welk opleidingsniveau iemand heeft.

5.3 Bijdrage van observaties en verbalisaties aan probleemdetectie

Om vast te kunnen stellen wat observaties en verbalisaties specifiek bijdragen aan de probleemdetectie is er van elke pagina die voorkomt in het ideale taakmodel vastgesteld wat voor problemen er gedetecteerd worden door observaties en door verbalisaties. Zo zijn er probleemsets ontstaan waarin voor ieder probleem is aangegeven of het via observaties of verbalisaties aan het licht is gekomen. Totaal zijn er 117 problemen gedetecteerd, waarvan 19 alleen door observaties, 64 alleen door verbalisaties en 34 problemen zijn met zowel observaties als verbalisaties te detecteren. In tabel 22 is te zien hoe de problemen uit de verschillende categorieën verdeeld zijn over observaties, verbalisaties en hoeveel er dubbel voorkomen.

Tabel 22. Verdeling probleemcategorieën over observaties en verbalisaties

Categorie	Observatie	Verbalisatie	Dubbel	Totaal
Inhoud	0 (0%)	9 (47.4%)	10 (52.6%)	19 (16.23%)
Volledigheid	0 (0%)	1 (16.7%)	5 (83.3%)	6 (5.1%)
Relevantie	0 (0%)	3 (60%)	2 (40%)	5 (4.3%)
Begrijpelijkheid	0 (0%)	5 (62.5%)	3 (37.5%)	8 (6.8%)
Bereikbaarheid	9 (18.4%)	30 (61.2%)	10 (20.4%)	49 (41.9%)
Navigatie	7 (33.3%)	8 (38.1%)	6 (28.6%)	21 (17.9%)
Structuur	2 (13.3%)	10 (66.7%)	3 (20%)	15 (12.8%)
Gebruiksgemak	0 (0%)	12 (92.3%)	1 (7.7%)	13 (11.1%)
Overig	5 (27.8%)	9 (50%)	4 (22.2%)	18 (15.4%)
Vormgeving	5 (16.1%)	16 (51.6%)	10 (32.3%)	31 (26.5%)
Totaal	19 (16.2%)	64 (54.7%)	34 (29.1%)	117 (100%)

Als er naar de verdeling van het totaal aantal problemen wordt gekeken valt uit de tabel af te lezen dat 29.1% van alle problemen zowel met observaties als met verbalisaties is te detecteren. Daarnaast valt op dat de meeste unieke problemen (54.7%) aan het licht komen met verbalisaties en 16.2% van alle problemen worden alleen met observaties opgespoord. Gekeken naar de dimensie inhoud, valt het op dat er met observaties geen problemen zijn te detecteren die niet met verbalisaties kunnen worden opgespoord. Daarnaast leveren verbalisaties bijna de helft (47.4%) van het totaal aantal inhoudsproblemen op, de meerderheid kan met beiden gevonden worden (52.6%). Vooral wat betreft het detecteren van problemen met betrekking tot de volledigheid van de inhoud zijn de beide methoden erg vruchtbaar, 83.3% wordt zowel met observaties als verbalisaties gedetecteerd.

Voor de dimensie bereikbaarheid is te zien dat ook hier de meeste problemen met verbalisaties aan het licht komen, 56.6%. Bij de categorie navigatie is te zien dat er zowel met observaties als met verbalisaties een flink aantal unieke problemen te vinden zijn (33.3% en 38.1%), ook wordt een groot deel van de navigatieproblemen met de beide methoden gevonden, 28.6%. Zowel de structuurproblemen als problemen met het gebruiksgemak worden het meest met verbalisaties gedetecteerd. Ook problemen in de categorieën vormgeving en overig komen het meest met behulp van verbalisaties aan het licht. Verbalisaties blijken dus overall gezien meer en een breder scala aan unieke problemen op te leveren dan observaties.

5.4 Bijdrage van RTA en RTE aan probleemdetectie

Ook is er in de probleemsets bijgehouden welke problemen via RTA en via RTE kunnen worden gedetecteerd. Van de 117 gevonden problemen zijn er 34 (29.1%) met de RTA conditie en 32 (27.4%) met de RTE conditie gevonden. 51 van de problemen (43.6%) is met beide condities op te sporen, dit is te zien bij de totaal aantallen in tabel 22.

Tabel 22. Verdeling probleemcategorieën over observaties en verbalisaties

Categorie	RTA	RTE	Beide	Totaal
Inhoud	7 (36.8%)	4 (21.1%)	8 (42.1%)	19 (16.23%)
Volledigheid	0 (0%)	1 (16.7%)	5 (83.3%)	6 (5.1%)
Relevantie	4 (80%)	1 (20%)	0 (0%)	5 (4.3%)
Begrijpelijkheid	3 (37.5%)	2 (25%)	3 (37.5%)	8 (6.8%)
Bereikbaarheid	15 (30.6%)	11 (22.4%)	23 (46.9%)	49 (41.9%)
Navigatie	5 (28.3%)	2 (9.5%)	14 (66.7%)	21 (17.9%)
Structuur	4 (26.7%)	4 (26.7%)	7 (46.7%)	15 (12.8%)
Gebruiksgemak	6 (46.2%)	5 (38.5%)	2 (15.4%)	13 (11.1%)
Overig	5 (27.8%)	9 (50%)	4 (22.2%)	18 (15.4%)
Vormgeving	7 (22.6%)	8 (25.8%)	16 (51.6%)	31 (26.5%)
Totaal	34 (29.1%)	32 (27.4%)	51 (43.6%)	117 (100%)

Wanneer er gekeken wordt naar problemen met betrekking tot de inhoud valt op dat de meeste problemen met de beide condities kunnen worden opgespoord. Door iets specifieker te kijken naar de categorieën is te zien dat er geen overlap is tussen de condities wat betreft de relevantie van de inhoud. In de RTA conditie worden veruit de meeste unieke problemen (80%) met betrekking tot de relevantie van de inhoud gedetecteerd. Problemen met betrekking tot de volledigheid van de inhoud worden met beide condities even vaak opgespoord (83.3%).

In de dimensie bereikbaarheid kunnen de meeste problemen met de beide condities worden gedetecteerd, 46.9% van de problemen komt in beide condities aan het licht. Wat betreft de problemen met betrekking tot het gebruiksgemak valt het op dat deze het meest in de RTA conditie tot uiting komen, 46.2%. De enige categorie waarin de RTE conditie de meeste unieke problemen opspoorde is de conditie overig. 50% van de problemen die niet direct met de taakuitvoering te maken heeft, wordt gedetecteerd in de RTE conditie. De meeste problemen met betrekking tot de vormgeving, 43.6%, worden met beide condities gevonden. Wanneer er wordt gekeken welke conditie de meeste unieke problemen oplevert, is te zien dat er in de RTA conditie de meeste unieke

problemen in de dimensies inhoud en bereikbaarheid zijn op te sporen. Met de RTE conditie worden de meeste unieke problemen in de categorie vormgeving en overig gevonden.

6. Conclusie

Na het presenteren van de resultaten wordt in dit hoofdstuk een antwoord gegeven op de hoofdvraag:

Wat dragen observaties en verbalisaties bij aan het detecteren van problemen? En in welke mate beïnvloeden persoonskenmerken en de conditie RTE en RTA de observaties en verbalisaties?

Er is allereerst gekeken naar de algemene kenmerken van de taakuitvoering als tijdsduur, mate van zelfstandigheid en de mate waarin de vraag correct is beantwoord. De taak op de website van de gemeente Nijmegen duurde langer dan die op de website van de gemeente Dordrecht. Voor de tijdsduur maakt het niet uit in welke conditie de proefpersoon zit, bij RTA en RTE deden ze er even lang over. De proefpersonen op de website van Nijmegen zijn minder vaak in staat om de antwoorden zelfstandig te vinden dan op de website van Dordrecht het geval is. Ook hier heeft de conditie waar de proefpersoon in zit geen invloed, de taak werd in beide condities even zelfstandig uitgevoerd. Er is geen verschil gevonden voor de mate waarin er een correct antwoord op de vragen is gevonden voor de website of de conditie. Er kan dus geconcludeerd worden dat de conditie, RTA of RTE geen invloed heeft gehad op de taakuitvoering zelf. Wel kan er worden gesteld dat de taak op de website van de gemeente Nijmegen langer duurde en de proefpersonen minder vaak in staat waren de antwoorden zelfstandig te vinden dan op de website van de gemeente Dordrecht het geval was.

Dan is er ook gekeken of persoonskenmerken als geslacht, leeftijdscategorie en opleidingsniveau invloed hebben gehad op de drie aspecten van de taakuitvoering. Er is vastgesteld dat het zowel voor de duur van de taak als de mate van zelfstandigheid waarin de taak is uitgevoerd niet uitmaakt wat het geslacht, de leeftijd en het opleidingsniveau van de proefpersoon is. De conclusie die Nell (2009) in haar scriptie stelt, dat oudere proefpersonen langer over de taak doen is daarmee, voor de condities en websites die in dit onderzoek zijn bestudeerd, niet bevestigd. Voor de mate van correctheid is geconstateerd dat de groep in de leeftijd van 40-54 jaar de vragen slechter heeft beantwoord dan de andere leeftijdsgroepen.

Vervolgens is er gekeken naar de algemene verschillen tussen observaties en verbalisaties. Het blijkt dat problemen met de relevantie van de inhoud, de structuur, het gebruiksgemak en overige problemen meer worden geverbaliseerd dan dat ze worden geobserveerd. Problemen die vaker

worden geobserveerd dan geverbaliseerd, zijn de navigatieproblemen. Er kan op basis van deze bevindingen gesteld worden dat er met verbalisaties een groter scala aan problemen gedetecteerd wordt dan met observaties het geval is. Dit geldt niet voor navigatieproblemen, die worden met observaties het meest opgespoord.

Ook is er gekeken of het voor de observaties en verbalisaties verschil maakt in welke conditie een proefpersoon zit. Bij de observaties is vastgesteld dat er in de RTA conditie meer problemen met betrekking tot gebruiksgemak worden gevonden dan in de RTE conditie. De andere problemen worden in zowel de RTA als de RTE conditie even vaak geobserveerd. De proefpersonen hebben zich dus over het algemeen op dezelfde manier gedragen op de websites en dit is op een consequente manier gerapporteerd. Voor de verbalisaties blijkt de conditie waar de proefpersoon in zit iets meer invloed uit te oefenen op de gevonden problemen. Zo worden er in de RTA conditie meer problemen geverbaliseerd met betrekking tot de vormgeving en de navigatie dan in de RTE conditie. Wanneer een proefpersoon de eigen oogbewegingen niet ziet, is deze in staat om meer navigatie en meer vormgevingsproblemen te verbaliseren. Ook het totaal aantal verbalisaties is hoger in de RTA conditie waardoor de bevinding van Gombert (2009) dat het zien van de oogbewegingen de proefpersoon afleidt van het verbaliseren, bevestigd wordt.

Daarnaast is er nog gekeken of de persoonskenmerken geslacht, leeftijd en opleidingsniveau invloed hebben op de mate waarin er problemen worden geobserveerd en geverbaliseerd. Bij de mannen zijn er totaal significant meer problemen geobserveerd dan bij vrouwen en specifiek zijn er meer structuurproblemen bij de mannen geobserveerd. Een structuurprobleem is geconstateerd wanneer er heen en terug wordt geklikt tussen pagina's en wanneer de oogbewegingen ongestructureerd over de pagina zijn gegaan, waardoor je kunt zien dat een proefpersoon op zoek is. Het is opvallend dat het kijk- en klikpatroon van mannen tot meer geobserveerde problemen leidt dan bij vrouwen het geval is. Bij de leeftijdscategorieën kon slechts één verschil gevonden worden, bij proefpersonen in de leeftijd van 18 tot 29 jaar werden de meeste problemen met betrekking tot gebruiksgemak geobserveerd. Bij proefpersonen uit deze leeftijdsgroep is vaker 'experimenteel' gedrag geobserveerd, waarbij bijvoorbeeld op data geklikt werd in de hoop daardoor meer informatie te bereiken. Voor de verbalisaties is er geen enkel verschil gevonden voor de persoonskenmerken en de gedetecteerde problemen. Er kan dus worden geconcludeerd dat het opleidingsniveau geen invloed heeft op de mate waarin er problemen zijn geobserveerd en geverbaliseerd. Wanneer er veel interesse is in het gebruiksgemak van de site, is het aan te raden om personen van 18 tot 29 jaar aan het onderzoek te laten deelnemen omdat dit een meer experimentele groep lijkt te zijn,

waardoor meer problemen met betrekking tot gebruiksgemak aan het licht kunnen komen via de observaties. Het lijkt erop dat mannen een iets ander kijk- en klikgedrag vertonen dan dat vrouwen dit doen, bij mannen zijn significant meer problemen geobserveerd dan bij de vrouwen.

Totaal zijn er met de probleemsets van de unieke problemen die met observaties, verbalisaties en met beiden zijn gevonden, 117 problemen vastgesteld op de pagina's binnen het ideale taakmodel van de twee websites. Het is opvallend dat de meeste problemen met verbalisaties zijn op te sporen. Guan (2006) concludeerde dat de hoge datadichtheid van de oogbewegingen juist meer problemen zou moeten opleveren, maar dat blijkt niet het geval te zijn. Vereijken (2007) vond in haar onderzoek dat observaties vooral bijdragen aan een gedetailleerd beeld in de categorieën bereikbaarheid en vormgeving. Er worden wel degelijk unieke problemen gevonden met behulp van observaties in de dimensie bereikbaarheid, maar met verbalisaties worden op iedere subdimensie meer problemen gedetecteerd. Ook voor vormgeving geldt dat er meer unieke problemen met verbalisaties gevonden zijn dan met observaties. Deze bevindingen van Vereijken kunnen daarom niet worden bevestigd. Wat wel overeenkomt is het feit dat Vereijken stelt dat observaties minder geschikt zijn voor het signaleren van inhoudsproblemen. Er werd geen enkel uniek inhoudsprobleem met observaties gevonden, er werden wel inhoudsproblemen gevonden met observaties maar deze zijn ook met verbalisaties te detecteren. Daarnaast wordt er met verbalisaties bijna 50% van de unieke problemen gedetecteerd. De proefpersonen kunnen verbaal heel specifiek aangeven wat voor moeilijkheden ze met de inhoud hebben ervaren. De bevindingen van De Bode (2008) kunnen worden bevestigd, de meeste unieke vormgevingsproblemen worden niet met observaties maar met verbalisaties vastgesteld. Dit komt omdat persoonlijke smaak een groot deel uitmaakt van vormgevingsproblemen en het met oogbewegingen alleen niet is vast te stellen wat iemands mening over de vormgeving van een bepaald element is.

7. Discussie

Met dit onderzoek is duidelijk geworden dat persoonskenmerken als leeftijd en opleidingsniveau nagenoeg geen invloed hebben op de taakuitvoering. Het enige verschil in de taakuitvoering is dat mensen tussen de 40 en 54 meer foute antwoorden hebben gegeven dan de andere leeftijdsgroepen. Het kan niet gesteld worden dat mensen met een hoger opleidingsniveau de taken per definitie beter of sneller uitvoeren dan mensen met een lager opleidingsniveau. Of dat oudere mensen meer moeite hebben met het zelfstandig bereiken van de informatie dan jongere mensen. Voor alle soorten proefpersonen blijkt het even moeilijk (of makkelijk) om informatie op een website zelfstandig, snel en foutloos te kunnen vinden.

Bij het detecteren van de problemen is wel een verschil gevonden tussen mannen en vrouwen wat betreft de geobserveerde problemen. Bij mannen zijn totaal significant meer problemen geobserveerd en specifiek zijn er meer structuurproblemen dan bij vrouwen het geval is. Structuurproblemen zijn gedetecteerd wanneer een proefpersoon steeds heen en weer blijft klikken en als uit de oogbewegingen duidelijk een zoekend patroon is af te lezen. Mannen blijken dit gedrag significant vaker te vertonen dan vrouwen, waardoor er meer problemen kunnen worden geobserveerd. Het zou voor vervolgonderzoek erg interessant zijn om verder uit te zoeken op welke manier het kijkpatroon en het klikgedrag verschilt tussen mannen en vrouwen.

Met dit onderzoek is het belang van verbalisaties onderstreept, er is een breed scala en een grote hoeveelheid problemen die aan het licht komt met verbalisaties. Een verklaring hiervoor kan zijn dat de kwaliteit van de oogbewegingen niet altijd even goed is, waardoor er niet altijd kan worden geobserveerd of er een probleem is en wat het probleem dan veroorzaakt en of de benodigde informatie wel of niet gezien is. Maar liefst 11,7% van de filmpjes is geheel onbruikbaar omdat de oogbewegingen ontbreken. Daarnaast zijn er filmpjes waarbij de oogbewegingen deels ontbreken en komt het voor dat de oogbewegingen niet goed gekalibreerd zijn waardoor het niet heel nauwkeurig vast te stellen is waar de proefpersoon op fixeert of welke alinea hij of zij leest. Wanneer de apparatuur preciezer zou zijn, zouden er waarschijnlijk meer en specifiekere problemen kunnen worden geobserveerd.

Een beoordelaar analyseert alle filmpjes op een zelfde manier volgens vast regels, waardoor kijkgedrag consequent op dezelfde manier wordt beoordeeld. Een groot pluspunt aan verbalisaties

is daarbij dat het een heel verfrissende kijk biedt op de problemen die voorbij komen. Een beoordelaar die de taakuitvoering observeert ziet proefpersonen afwisselend steeds hetzelfde gedrag vertonen. Iets wordt wel of niet gezien, vluchtig of nauwkeurig bekeken en daaraan worden volgens eenzelfde regime een conclusie verbonden. Verbalisaties daarentegen laten iedere keer opnieuw een ander licht op de gebeurtenissen schijnen. Verbalisaties zijn daardoor een stuk interessanter om te analyseren, het is iedere keer weer een verrassing hoe de proefpersoon het proces heeft ervaren. Het is ook aan te raden om de problemen met meer beoordelaars te categoriseren om zo de betrouwbaarheid van de beoordelingen meer te kunnen garanderen en over eventuele probleemgevallen te kunnen discussiëren.

Net als in de voorgaande scripties onder leiding van Elling is ook in deze scriptie wederom niet aangetoond dat het terugzien van de eigen oogbewegingen de proefpersoon extra stimuleert bij het verbaliseren. Gekeken naar de hoofddimensies inhoud, bereikbaarheid en vormgeving worden de meeste unieke problemen in de beide condities gevonden. Meer uitgesplitst blijken de meeste unieke problemen met relevantie en begrijpelijkheid van de inhoud en het gebruiksgemak in de RTA conditie gevonden te worden. De meeste unieke problemen in de RTE conditie komen uit de categorie overig, dat wil zeggen dat dit problemen zijn die niet direct met de taakuitvoering te maken hebben; proefpersonen lijken echt afgeleid te worden door hun eigen oogbewegingen. Vooralsnog zijn proefpersonen nog niet genoeg bekend met oogbewegingen om er voordelen van te kunnen ondervinden. Om te onderzoeken of onbekendheid met oogbewegingen er voor zorgt dat de proefpersoon worden afgeleid in plaats van ondersteund, zouden er kleinschalige studies kunnen worden gedaan waarin experts die veel met het analyseren van oogbewegingen werken, aan RTE-onderzoek worden onderworpen. Als daaruit blijkt dat personen met meer ervaring met het zien van oogbewegingen wel profijt hebben van de methode, kan worden vastgesteld of er toekomst zit in het Retrospective Think-aloud with Eyemovement.

8. Literatuur

Baker, L. (1989) Metacognition, comprehension monitoring, and the adult reader. *Educational Psychology Review*, 1, 3-38.

De Bode, M. (2008). Weet ik wat jouw ogen zien? Een onderzoek naar de mogelijkheden van eyetracking als zelfstandige probleemopsporende methode in usabilityonderzoek. *Masterscriptie Communicatiestudies*, Universiteit Utrecht.

Boren, M. T., & Ramey, J. (2000). Thinking aloud: Reconciling theory and practice. *IEEE Transactions on Professional Communication*, 43(3), 261-278.

Cooke, L. & E. Cuddihy, 2005. Using eye tracking to address limitations in think aloud protocol. *IEEE International Professional Communication Conference Proceedings*: 653-658.

Eger, N., L. Ball, et al., 2007. Cueing retrospective verbal reports in usability testing through eye-movement replay. *People and Computers XXI – HCI ... but not as we know it: Proceedings of HCI 2007*.

Evaluatie van gemeentelijke websites, http://www.gemeentewebsites.nl/waarom_1.html (geraadpleegd in juni/juli 2010).

Ericsson, K.A., & Simon, H.A. (1984). Protocol Analysis: Verbal reports as Data. *Cambridge, MA: The MIT Press*.

Elling, S., Lentz, L., Jong, de, M. (2007) Website Evaluation Questionnaire: Development of a research-based tool for evaluating informational websites. In *M.A. Wimmer, H.J. Scholl & A. Gröhlund (Eds.) Lecture Notes in Computer Science*, 4656, Berlin Heidelberg: Springer- Verlag. 293-304.

Van Gog et al. (2005): 'Uncovering the problem-solving process: cued retrospective reporting vs. concurrent and retrospective reporting.' *In: Applied Journal of Experimental Psychology*, 11(4), 237 – 244.

Gombert, J. (2009). Je eigen ogen als afleidingsmanoeuvre?! Een kwalitatief onderzoek naar het effect van het terugzien van de eigen oogbewegingen bij participanten tijdens Retrospective Thinking-Aloud. *Masterscriptie Communicatiestudies*, Universiteit Utrecht.

Guan, Z., S. Lee, et al., (2006). The validity of the stimulated retrospective think-aloud method as measured by eye tracking. *CHI 2006 Proceedings*. Usability Methods, Montréal, Québec, Canada.

Van den Haak, M., De Jong, M. & Schellens, P.J. (2003). 'Retrospective versus concurrent think-aloud protocols: Testing the usability of an online library catalogue.' In: *Behaviour and Information Technology*, 22, 339 - 351.

Van den Haak, M., De Jong, M. & Schellens, P.J. (2007). 'Evaluation of an informational Web site: Three variants of the think-aloud method compared.' In: *Technical Communication*, 54(1), 58 – 71.

Van den Haak, M., De Jong, M. & Schellens, P.J. (2009). 'Evaluating municipal websites: A methodological comparison of three think-aloud variants.' In: *Government Information Quarterly*, 26, 193 – 202.

Haas, E. (2009). 'Ik moet even wennen aan die stippeltjes'. Een onderzoek naar de toegevoegde waarde van het terugzien van de oogbewegingen op verbalisaties bij retrospectief hardopdenken, uitgevoerd op een gemeentelijke website. *Masterscriptie Communicatiestudies*, Universiteit Utrecht.

Hertzum, M., Hansen, K.D. en Andersen, H.H.K. (2009). 'Scrutinising usability evaluation: does thinking aloud effect behaviour and mental workload?' In: *Behaviour and Information Technology*, 28(2), 165 - 181.

De Jong, M., L. Lentz, S. Elling & P.J. Schellens (2004) Scenario-evaluatie van gemeentelijke websites. De ontwikkeling en toepassing van een expertgericht evaluatie-instrument. *Aspect 69*. Enschede: Universiteit Twente, Afdeling Communicatiewetenschap.

Just, M. en P. Carpenter (1980). A theory of reading: from eye fixations to comprehension. Geciteerd door: Hertzum e.a. (2009)

Lentz, L., M. de Jong (2009). How do experts assess usability problems? An empirical analysis of cognitive shortcuts. *Technical Communications* 56 (2), 111-121.

Molich, R., J. Dumas (2008). Comparative Usability Evaluations (CUE-4). *Behavior and Information Technology* 27 (3), 263-281.

Nell, L. (2009). Ik verbaliseer dus ik evalueer. Een studie naar de invloed van de gebruikersgerichte webevaluatiemethoden synchroon hardopdenken en retrospectief hardopdenken en demografische kenmerken op de taakuitvoering en verbalisaties van proefpersonen. Masterscriptie Communicatiestudies, Universiteit Utrecht.

Potappel, E. (2007). Kun jij gedachten lezen? Een onderzoek naar de toegevoegde waarde van oogbewegingen bij retrospectieve hardopdenkprotocollen in probleemopsporend usabilityonderzoek. *Masterscriptie Communicatiestudies*, Universiteit van Utrecht.

Vereijken, E. (2007). 'Ik zie, ik zie wat jij (niet) ziet': een verkenning van de mogelijkheden van eye tracking voor probleemopsporend usability-onderzoek naar websites. *Masterscriptie Communicatiestudies*, Universiteit van Utrecht.

Williams, T.R., Mulligan, C., Koprowicz, K., Miller, J., Reimann, C., Wang, D.S. (2005). Does isolating a visual element call attention to it? Results of an eye-tracking investigation of the effects of isolation on emphasis. *Technical Communication*, 52 (1), 21-26.

9. Bijlagen

I Taak 1 Dordrecht & taakmodel Dordrecht

II Taak 1 Nijmegen & taakmodel Nijmegen

III Probleemsets

IV Voorbeelden uit data proefpersonen

Bijlage I

Taak 1 Dordrecht:

U hebt twee kinderen: een van vier jaar die op de basisschool zit en een van dertien jaar op de middelbare school. In de komende zomervakantie wilt u drie weken naar Turkije gaan. Voorlopig denkt u eraan te vertrekken op 9 augustus en terug te komen op 30 augustus. U wilt op de website van Dordrecht nagaan wanneer de schoolvakanties in de zomer van 2009 vallen.

1a. Hebben beide kinderen zomervakantie in de periode van 9 augustus tot 30 augustus?

1b. U vermoedt dat u met uw jongste kind van vier buiten de officiële zomervakantie weg mag gaan, maar u wilt dit voor de zekerheid nog checken op de website. Is het volgens de informatie op de website inderdaad toegestaan om met een kind van vier buiten de schoolvakantie op reis te gaan?

Ideale taakmodel Dordrecht:

1. Homepage	
	Zien topnavigatie/zien linklijst
1.1 Navigatie	Zien tekstlink/tabje Onderwijs, werk en welzijn
	Aanklikken Onderwijs, werk en welzijn
2. Onderwijs, werk en welzijn	Zien tekstlink Schoolvakanties
	Aanklikken tekstlink Schoolvakanties
3. Schoolvakanties	Zien data zomervakantie 2008-2009
	Lezen data zomervakantie 2008-2009
4. Leerplicht	Zien link Leerplicht
	Aanklikken link Leerplicht
	Lezen tekst Leerplicht bovenaan

Bijlage II

Taak 1 Nijmegen:

U verhuist uit een andere gemeente naar Nijmegen. Een dag na uw verhuizing, hoort u van een vriend dat u zich als nieuwe bewoner hoort te laten registreren bij de gemeente waar u bent gaan wonen. Op welke manieren kunt u zich bij de gemeente Nijmegen laten registreren?

1a. Kunt u zich laten registreren door iets op te sturen?

Ja /nee

Zo ja, wat moet u daar precies voor doen?

1b. Kunt u zich laten registreren door zelf naar het stadhuis te gaan?

Ja / nee

Zo ja, wat moet u daar precies voor doen?

1c. Kunt u zich laten registreren op de website zelf?

Ja / nee

Zo ja, wat moet u daar precies voor doen?

Ideale taakmodel Nijmegen

1. Homepage	
	Zien topnavigatie
	Zien dat de benodigde info niet via de navigatie gevonden kan worden
	Zien Digitale Balie
	Aanklikken Digitale Balie
2. Digitale Balie	Zien Rubrieken
	Aanklikken Leven, reizen & papieren (of Bouwen & wonen)
3. Leven, reizen & papieren (of Bouwen & wonen)	Zien subrubrieken
	Zien link Verhuizen
	Aanklikken Verhuizen
4. Verhuizen	Zien producten
	Zien link verhuizing
	Aanklikken verhuizing

5. verhuizing	Zien Voorwaarden
	Lezen Voorwaarden 1b+1c
	Zien verschillende printformulieren
	Lezen derde printformulier 1a

Bijlage III

Probleemsets

Startpagina

Probleemset Homepage

Dordrecht	RTA	
Home-Obs	Home-Verb	Beschrijving
Navigatie1		Ziet benodigde link wel, klikt niet meteen
Vormgeving1	Vormgeving1	Ziet benodigde link niet
	Navigatie2	Opties links en boven hetzelfde
Overig1		Snapt de bedoeling niet
	Vormgeving2	Staan veel verschillende dingen

Dordrecht	RTE	
Home-Obs	Home-Verb	Beschrijving
Vormgeving1		Ziet benodigde link niet
	Vormgeving3	Hoe beginnen?
Navigatie1		Ziet benodigde link wel, klikt niet meteen

Startpagina

Probleemset Homepage

Nijmegen	RTA	
Home-Obs	Home-Verb	Beschrijving
Vormgeving1	Vormgeving1	Ziet benodigde link niet
Navigatie1		Kiest voor foutieve link
	Navigatie2	Twijfelen over navigatieopties
Overig1	Overig1	Functionaliteit mouseovers
Structuur1	Structuur1	Geen overzicht hebben
	Vormgeving2	Om bewegend beeld heen kijken
	Gebruiksgemak1	Kan plaatjes en tekst van bewegend beeld niet tegelijk bekijken

	Vormgeving3	Vind de homepage erg onduidelijk
--	-------------	----------------------------------

Nijmegen	RTE	
Home-Obs	Home-Verb	Beschrijving
Vormgeving1		Ziet benodigde link niet
Navigatie1		Kiest voor foutieve link
	Vormgeving3	Moet goed zoeken, onduidelijk
	Navigatie2	Twijfelen over navigatieopties
	Overig2	Wat was de opdracht
Overig1		Functionaliteit mouseover

Overzichtspagina

Probleemset Onderwijs, werk & welzijn

Dordrecht	RTA	
Oww-Obs	Oww-Verb	Beschrijving
Vormgeving1	Vormgeving1	Ziet benodigde link niet
Overig1	Overig1	Aanklikken tekst tab werkt niet
Navigatie1		Kiest voor foutieve link
	Structuur1	Hoe moet ik verder?

Dordrecht	RTE	
Oww-Obs	Oww-Verb	Beschrijving
Vormgeving1	Vormgeving1	Ziet benodigde link niet
Overig2		Website reageert niet
	Structuur1	Hoe moet ik verder?
Structuur2		Terugklikken
	Navigatie2	Op zoek naar goede link
Navigatie1		Kiest voor foutieve link

Overzichtspagina

Probleemset Digitale Balie

Nijmegen	RTA	
DigiB-Obs	DigiB-Verb	Beschrijving
Navigatie1		Kiest voor foutieve link
Vormgeving1	Vormgeving1	Ziet benodigde link niet
Structuur2		Terugklikken
	Structuur1	Hoe moet ik verder?
	Vormgeving2	Onoverzichtelijke pagina
Relevantie1	Relevantie1	Pp denkt verkeerd te zitten
	Volledigheid1	Vindt informatie niet helemaal terug
	Structuur3	Waar bevind je je op de website

	Gebruiksgemak1	Krijgt nu geen submenu in mouseover
Navigatie2		Ziet benodigde link wel, klikt niet meteen
Nijmegen	RTE	
DigiB-Obs	DigiB-Verb	Beschrijving
Navigatie2	Navigatie2	Ziet benodigde link wel, klikt niet meteen
	Gebruiksgemak2	Kan doel niet snel genoeg bereiken
Vormgeving1	Vormgeving1	Ziet benodigde link niet
Navigatie1		Kiest voor foutieve link
Volledigheid 1	Volledigheid1	Moet verder zoeken
Structuur2	Structuur2	Terugklikken
	Gebruiksgemak3	Het is niet makkelijk aangegeven
	Overig1	Wil langsgaan bij gemeente
Overig2	Overig2	Gebruikt zoekbox
	Overig3	Weet niet waarom iets zo gegaan is
	Vormgeving3	Titel zoeken in digitale balie wekte andere verwachting op
Overig4		Kan venster niet vergroten

Productpagina

Probleemset Schoolvakanties 1a

Dordrecht	RTA	
Sv-Obs	Sv-Verb	Beschrijving
Begrijpelijkheid1	Begrijpelijkheid1	Begrijpelijkheid van de data
Overig1	Overig1	Probleem met vraagstelling
Volledigheid1	Volledigheid1	Info over leerplicht mist
Vormgeving1	Vormgeving1	Ziet benodigde link niet
Navigatie1		Kiest voor foutieve link
Vormgeving2	Vormgeving2	Ziet data niet
	Structuur1	Hoe kom je bij leerplicht
Gebruiksgemak1	Gebruiksgemak1	Data zijn niet aanklikbaar
	Relevantie1	Wat gevonden is, wordt niet gezocht
	Overig2	Thuis zou het beter zijn gegaan
	Structuur2	Verwarrend dat vorig jaar erbij staat
	Volledigheid2	Data jongste kind wordt gemist
	Relevantie2	Data van ander jaar niet nodig
	Navigatie2	Twijfelen over navigatieopties
Navigatie3		Ziet goede link klikt er niet op
	Gebruiksgemak2	Info niet snel te vinden

Dordrecht	RTE	
Sv-Obs	Sv-Verb	Beschrijving
Volledigheid1	Volledigheid1	Info over leerplicht mist
Begrijpelijkheid1	Begrijpelijkheid1	Begrijpelijkheid van de data
Vormgeving1		Ziet benodigde link niet

	Gebruiksgemak3	Data worden op andere plek gewenst
Navigatie1		Kiest voor foutieve link
	Overig1	Probleem met vraagstelling
	Navigatie2	Twijfelen over navigatieopties
	Vormgeving3	Hoe zit de data in elkaar
	Structuur1	Hoe kom je bij leerplicht
	Vormgeving2	Ziet data niet
Navigatie3		Ziet goede link klikt er niet op
	Overig2	Thuis zou het beter zijn gegaan
	Structuur3	Terugklikken
	Vormgeving4	Problemen met blauwe kleur
	Overig3	Weet niet wat er gedacht werd

Productpagina

verhuizing

Nijmegen	RTA	
verh-Obs	verh-Verb	Beschrijving
Begrijpelijkheid1	Begrijpelijkheid1	Voorwaarden niet begrepen
Vormgeving1		Informatie wordt niet gezien
	Vormgeving2	Erg veel tekst
Volledigheid1	Volledigheid1	Meer informatie nodig
Vormgeving3	Vormgeving3	Slechte vormgeving van tekst
	Relevantie1	Andere informatie nodig
	Navigatie1	Andere naam voor label in gedachte
Vormgeving4	Vormgeving4	Link niet zichtbaar
	Gebruiksgemak1	Weg naar antwoord toe te moeilijk
	Navigatie2	Rechtstreekse link wordt gemist
	Structuur1	Informatie onder andere navigatietab plaatsen
	Structuur2	Heen en weer klikken
	Navigatie3	Je wordt de goede richting niet op gestuurd
Vormgeving5		Voorwaarden niet gelezen
	Overig1	Thuis zou het beter gaan
	Navigatie4	Begrijpelijkheid link Digitale balie
	Begrijpelijkheid2	Terminologie niet begrepen
Begrijpelijkheid3		Begrijpelijkheid formulieren
Overig2		Functionaliteit website
	Gebruiksgemak2	Formulieren overzichtelijker met icoontje

Nijmegen	RTE	
verh-Obs	verh-Verb	Beschrijving
Begrijpelijkheid1	Begrijpelijkheid1	Voorwaarden niet begrepen
Vormgeving6		Onopvallende vormgeving formulieren
Vormgeving3	Vormgeving3	Slechte vormgeving van tekst
	Gebruiksgemak3	Ziet informatie niet snel genoeg
	Volledigheid1	Meer informatie nodig

Vormgeving1	Vormgeving1	Informatie wordt niet gezien
Begrijpelijkheid3	Begrijpelijkheid3	Begrijpelijkheid formulieren
	Navigatie1	Andere naam voor label in gedachte
	Vormgeving7	Plaats van formulieren onderaan pagina
	Overig3	Weet niet waarom het zo gegaan is
	Structuur2	Heen en weer klikken
	Vormgeving2	Staat erg veel tekst op pagina
Vormgeving4	Vormgeving4	Link niet zichtbaar
	Structuur1	Informatie onder andere navigatietab plaatsen
	Gebruiksgemak1	Weg naar antwoord toe te moeilijk
	Vormgeving8	Plaats van navigatietab niet optimaal

Productpagina II

Probleemset Leerplicht 1b

Dordrecht	RTA	
Lp-Obs	Lp-Verb	Beschrijving
Vormgeving1		Tekst niet gelezen
Navigatie1	Navigatie1	Kiest voor foutieve link
Vormgeving2		Ziet link naar schoolvakantie niet
	Vormgeving3	Opbouw van linklijst
Volledigheid1	Volledigheid1	Waar is informatie schoolvakanties
	Structuur1	Hoe kom je bij schoolvakanties
Volledigheid2	Volledigheid2	Meer informatie nodig
Overig1		Website reageert niet
	Relevantie1	Komt niet overeen met zoekvraag
	Vormgeving4	Kan informatie niet vinden
Relevantie2		Op zoek naar andersoortige informatie
	Gebruiksgemak1	Moeilijk te vinden
	Vormgeving5	Schuine tekst moeilijk te lezen
	Structuur2	Waar is de informatie te vinden

Dordrecht	RTE	
Lp-Obs	Lp-Verb	Beschrijving
Overig1		Website reageert niet
Volledigheid1	Volledigheid1	Waar is informatie schoolvakanties
Navigatie1		Kiest voor foutieve link
	Overig2	Begrip taakuitvoering
	Structuur2	Waar is de informatie te vinden
Volledigheid2	Volledigheid2	Meer informatie nodig
Vormgeving1		Tekst niet gelezen
Vormgeving2	Vormgeving2	Ziet link naar schoolvakantie niet
	Structuur3	Informatie onder andere navigatietab plaatsen
	Vormgeving6	Had link op schoolvakanties verwacht
	Vormgeving5	Schuine tekst moeilijk te lezen
	Begrijpelijkheid1	Formulering tekst leerplicht
	Relevantie2	Geen informatie over schoolvakantie

Structuur4	Structuur4	Hoe verder?
	Gebruiksgemak1	Moeilijk te vinden
	Vormgeving4	Kan informatie niet vinden

Overzichtspagina II

Probleemset Leven, reizen, papieren/ Bouwen&wonen/ Verhuizen

Nijmegen	RTA	
Lrp-Obs	Lrp-Verb	Beschrijving
Vormgeving1		Ziet benodigde link niet
	Vormgeving2	Links onoverzichtelijk weergegeven
Navigatie1		Verkeerde link kiezen
Navigatie2		Ziet benodigde link maar klikt er niet op
	Volledigheid1	Informatie wordt gemist
Navigatie3		Link wordt niet meteen gezien
	Overig1	Weet niet meer wat er werd gedacht
	Vormgeving3	Vormgeving tekst
Structuur1		Heen en weer klikken
	Begrijpelijkheid1	Formulering verhuizen ->alsof je weggaat
	Begrijpelijkheid2	Wat is stadswinkel
	Gebruiksgemak1	Info moeilijk te bereiken

Nijmegen	RTE	
Lrp-Obs	Lrp-Verb	Beschrijving
Navigatie3		Link wordt niet meteen gezien
	Gebruiksgemak2	Info kan niet snel genoeg worden bereikt
Navigatie1		Verkeerde link kiezen
Vormgeving1		Ziet benodigde link niet
Structuur1		Heen en weer klikken
	Navigatie4	Naam link niet ladingdekkend
	Volledigheid1	Informatie wordt gemist
	Navigatie5	Terminologie link niet die van bezoeker
	Begrijpelijkheid3	Formulering informatie

Bijlage IV

Voorbeeld uit data proefpersonen

Website	Dordrecht	Problemen Observaties: 3	Opmerkingen:	Problemen Verbalisaties: 3
Conditie	RTA	2 x Volledigheid inhoud	Geen hulp, goed antwoord op 1b,	2 x Navigatie
Proefpersoon	10-31	1 x Navigatie	bijna goed 1a.	1 x Volledigheid inhoud
Geslacht	Man			
Leeftijd	62			
Opleidingsniveau	Hoog			

Ideale taakmodel	Gevolgd klikpad	Observaties	Beschrijving
1. Homepage	Zien topnavigatie/zien linklijst	Scannen linklijst	Pp scant de linklijst en ziet de benodigde link meteen.
1.1 Navigatie	Zien tekstlink/tabje Onderwijs, werk en welzijn	Fixatie op Onderwijs, werk & welzijn	
	Aanklikken Onderwijs, werk en welzijn	<u>Onderwijs, werk & welzijn</u>	
2. Onderwijs, werk en welzijn		Lezen tekst	Het wordt voor de pp meteen duidelijk welke tekstlink hij moet hebben.
		Fixeren op hyperlinks in tekst	
	Zien tekstlink Schoolvakanties	Fixeren op schoolvakantie	
	Aanklikken tekstlink Schoolvakanties	<u>Schoolvakantie</u>	
3. Schoolvakanties		Scannen tekst schoolvakanties	Hij ziet de benodigde data snel, en zoekt dan door naar 1b.
	Zien data zomervakantie 2008-2009	Fixatie op herfstvakantie	
		Scannen datalist	
		Fixatie op hemelvaartsvakantie	

Het onderste gedeelte volgt op pagina 50, het rechter deel van de pagina op pagina 49

Opmerkingen:
Kan redelijk goed na vertellen hoe het ging.

Probleem	Verbalisaties	Probleem
	<p><u>Homepage</u> Nou, website van Dordrecht ziet er cool uit. Watergemeente, dus blauw. Dat doet me daaraan denken.</p> <p><u>Onderwijs, werk en welzijn</u> En ik moet natuurlijk bij onderwijs gaan kijken</p> <p><u>Schoolvakanties</u> En schoolvakanties. Nou, dat staat hier al heel mooi op. Herfstvakantie, nou schoolvakantie zal aan, zomervakantie zal aan het einde staan. Dat klopt ook. En oh, van juli tot augustus vakantie ja. Ja, dan valt het basisschoolkind er 1 week buiten. En dan moet ik nu even op gaan zoeken wat de leerplichtwet zegt over...</p>	

		Fixatie zomervakantiedata 2009	
		Lezen zomervakantiedata 2009	Pp blijft huidige pagina afstruinen maar vindt geen informatie over de leerplicht
		Fixatie zomervakantiedata 2009	
		Scannen tekst bovenaan pagina	Kijkt goed rond voor hij verder klikt.
		Scannen datalijst	
		Fixatie zomervakantiedata 2009	
		Scannen toelichting midden op pagina	
		Lezen van breadcrumb	
		Fixatie op onderwijs, werk en welzijn in breadcrumb	
		Fixatie scholen in kaart in breadcrumb	
	Onderwijs, werk & welzijn (in breadcrumb)	Scannen hyperlinks in tekst	Pp gaat terug naar vorige pagina en zoekt daar naar aanknopingspunt. Hij kiest uiteindelijk voor een verkeerde link. Zoekt een poosje rond en klikt dan verder.
	<u>Onderwijs</u>		
	Zien link Leerplicht	Buiten ideale pad	
	Aanklikken link Leerplicht	<u>Leerplicht</u>	
4. Leerplicht	Lezen tekst Leerplicht	Lezen tekst van leerplichtwet	
		Fixeren op hyperlink Regionaal Bureau Leerplicht	
		Fixatie op hyperlink Leerlingbegeleiding	Kiest vervolgens voor de goede link
		Herlezen tekst leerplichtwet	kijkt rond en leest de goede tekst.
		Scannen tekst Leerplichtwet	
			<i>Geeft goed antwoord op 1b</i>

Rechter deel volgt op de volgende pagina

	zoeken wat de leerplichtwet zegt over...	
<i>Volledigheid inhoud</i>	Ja, toch nog even kijken. Om het te onthouden. Nee, hier staat het nog niet van de leerplichtwet.	<i>Volledigheid inhoud</i>
	<u>Onderwijs, werk en welzijn</u> Dus dan moet ik gewoon terug naar het begin van deze, dit onderwerp. En dan moet ik daar iets kunnen vinden over de leerplichtwet.	
	<u>Onderwijs</u> Werk en inkomen, onderwijs. Ja, daar zou daar zou, daar moet het bij staan zou ik zeggen. Leerplicht, ik zie het al in de linkerbalk, eh staan. Oh nee, ik ben nu eerst naar die blauwe balk kijken.	<i>Navigatie</i>
<i>Navigatie</i>	Dat is toch eigenlijk wel de volgorde, dat ik daar eerst moet gaan zoeken. Nou, leerplicht, voortijdig schoolverlaten, daar moet het staan. Even kijken, maar welke link moet ik dan hebben. Ministerie, leerling begeleiding.	<i>Navigatie</i>
<i>Volledigheid inhoud</i>	<u>Onderwijs</u> <u>Onderwijs, werk en welzijn</u> Oh kijk, ik moet weer naar die linkerbalk. Daar staat leerplicht waarschijnlijk.	
	<u>Schoolvakanties</u>	