

# Nytteverdi av øyesporing i brukbarhetstesting

En kvalitativ studie

**Elin Standal**

Master i datateknikk  
Oppgaven levert: Juni 2010  
Hovedveileder: Dag Svanæs, IDI



# Oppgavetekst

Oppgaven omhandler bruk av øyesporing i brukbarhetstesting av interaktive systemer og tjenester. Problemstillingen er knyttet til hvilken verdi data fra øyesporing har som del av en brukersentrert utviklingsprosess. Prosjektet gjøres i samarbeid med Amnesty Norge. Deres nye nettside skal brukbarhetstestes med samtidig bruk av TOBII øyesporing. Nytteverdien av øyesporingsdata skal evalueres både fra et forskningsperspektiv, og i praktisk dialog med utviklerne.

Oppgaven gitt: 15. januar 2010  
Hovedveileder: Dag Svanæs, IDI





---

## Forord

---

Denne masteroppgaven er utarbeidet ved institutt for datateknikk og informasjonsvitenskap ved Norges Teknisk-Naturvitenskapelige Universitet i Trondheim, som en avslutning til sivilingeniørstudiet i datateknikk med retning innefor systemutvikling og menneske-maskin interaksjon.

Jeg har lyst å takke min veileder Dag Svanæs for all god rettleiding når det gjelder rapportskrivning og gjennomføring av undersøkelsene. Spesielt takk til Terje Røsand også, som alltid har stilt opp når det gjelder utstyret på testlaboratoriet, hjulpet til under brukbarhetstesting og svart på alle de spørsmålene jeg har hatt som er relevante for oppgaven.

Takk også til Amnesty for et godt samarbeid og for at jeg fikk anledning til å benytte prosjektet deres som case i oppgaven, og spesielt takk til Kristin Rødland Buick som har svart på alle epostene jeg har sendt.

I tillegg vil jeg takke Thor Fredrik Eie fra NetLife Research som stilte opp til intervju. Det gav meg mye verdifull informasjon.

Jeg vil heller ikke glemme alle dere som frivillig stilte opp som testdeltakere under brukbarhetstesten, det er jeg veldig takknemlig for. Alle som bidro har vært viktige for resultatet i denne rapporten.

Takk også til Ole Andreas Alsos for tips til litteratur. Og til slutt takk til mine foreldre, Marta og Kjetil Standal, som ville lese korrektur på kort frist, og ellers alle som har støttet meg dette semesteret gjennom ord og tanker.

*Trondheim, 11. juni 2010*



---

*Elin Standal*



---

## Sammendrag

---

I denne rapporten retter vi søkelyset mot bruken av øyesporing som del av brukbarhetstesting. Mange lurer nok på om det er verdt å investere i slikt utstyr, og denne masteroppgaven er et forsøk på å finne ut mer om hvilken nytteverdi øyesporing kan ha når det gjelder brukbarhetstesting i systemutvikling. Jeg vil poengtere at jeg her ikke har prøvd å gjøre noen beregninger når det gjelder økonomiske aspekter, men konsentrert meg om å *utforske hvilke områder* man kan finne en nytteverdi i bruk av øyesporing.

For å få svar på dette har jeg gjennomført en brukbarhetstest med øyesporing og sett på hvilke problemer vi oppdaget uten å ta hensyn til øyesporingsdata, og deretter sammenlignet med hva slags tilleggsinformasjon vi fikk utfra øyesporingen. I forskningen benyttet jeg meg av et case, et systemutviklingsprosjekt iverksatt av Amnesty Norge for å fornye nettsidene deres, og testet brukbarheten på disse nettsidene. Jeg hadde så et møte med utviklingsgruppen der hvor jeg presenterte funnene fra brukbarhetstesten. Samtidig benyttet jeg anledningen til å se om resultatene fra øyesporingen hadde noen nytteverdi i forhold til kommunikasjonen mellom meg som brukbarhetstester og Amnesty som kunde.

I tillegg til dette har jeg gjort et intervju med Thor Fredrik Eie fra NetLife Research, som benytter øyesporing i brukbarhetstesting ofte, for å finne ut hvilken nytteverdi de som jobber med dette opplever. Til slutt har jeg også trukket fram noen av de praktiske og metodiske utfordringene vi har kommet over i løpet av denne studien når det gjelder bruken av øyesporing i brukbarhetstesting.

Når det gjelder hva øyesporingen bidro med for å bedømme brukbarheten av nettsiden vi testet, så fant vi hovedsaklig tre områder hvor øyesporing var til hjelp. For det første var det med på å avdekke et problem med designet hvor brukeren ikke oppdaget et viktig element, ved at vi fikk se hvor han rettet oppmerksomheten sin. Øyesporingen var også til hjelp for å bekrefte og forsterke teorier om problemer vi hadde gjort oss på forhånd, og vi fikk en bedre forståelse for hvor brukerne forventet at ting skulle være på nettsiden, ved at vi kunne se hvilke menyvalg de vurderte før de klikket.

Fra intervjuet mitt med Thor Fredrik Eie kom det fram at øyesporing kan ha nytteverdi som et hjelpeverktøy under gjennomføringen av testen også, på den måten at det er enklere for testleder å tilpasse spørsmål til brukeren når man kan se hvor brukeren fester blikket. I tillegg er data fra øyesporing et kraftig verktøy når det gjelder kommunikasjonen mellom kunder og brukbarhetstestere. Jeg opplevde at kunden fikk større forståelse for problemene når de fikk se med egne øyne hvor brukerne så hen. Det er også med på å gi argumentene man kommer med større tyngde og troverdighet når man kan vise bilder og videoer med øyesporing. Men her må man huske på at slike visuelle framstillinger av data kan gi inntrykk av å være viktigere enn de er, da det er lett å lage noe som ser stilig ut, men ikke egentlig forteller deg noe viktig.

Man må også være klar over at problemer med utstyr og programvare kan oppstå med jevne mellomrom. Heller ikke alle typer mennesker kan benyttes når man gjør brukbarhetstester med øyesporing, for eksempel kan bruk av briller være problematisk.



---

# Innhold

---

<b>Innhold</b>	<b>v</b>
<b>Figurer</b>	<b>ix</b>
<b>Tabeller</b>	<b>xi</b>
<b>1 Introduksjon</b>	<b>1</b>
1.1 Motivasjon . . . . .	1
1.2 Forskningsspørsmål . . . . .	1
1.3 Struktur av rapporten . . . . .	2
<b>I Teori og bakgrunnstoff</b>	<b>3</b>
<b>2 Brukbarhet og brukbarhetstesting</b>	<b>5</b>
2.1 Brukbarhet . . . . .	5
2.1.1 Definisjoner og begreper . . . . .	5
2.2 Brukbarhetstesting . . . . .	6
2.2.1 Utstyr og setting . . . . .	6
2.2.2 Prosedyre . . . . .	6
2.3 System Usability Scale (SUS) . . . . .	7
2.3.1 Beregning av poengsum . . . . .	8
2.3.2 Betydningen av individuelle poengsummer . . . . .	8
2.3.3 T-testen og statistiske begreper . . . . .	8
<b>3 Øyesporing</b>	<b>11</b>
3.1 Hva er øyesporing? . . . . .	11
3.1.1 Utvikling fram til i dag . . . . .	11
3.1.2 Hvordan fungerer dagens øyesporere? . . . . .	13
3.2 Hvor benyttes øyesporing? . . . . .	13
3.3 Hvordan fungerer synet vårt? . . . . .	14
3.4 Viktige begreper . . . . .	14
3.4.1 Fiksering (Fixation) . . . . .	15
3.4.2 Sakkader (Saccades) . . . . .	15
3.4.3 Blikkdiagram (Gaze plot) . . . . .	15
3.4.4 Varmekart (Heat maps) . . . . .	15
3.4.5 Interesseområde (Area Of Interest — AOI) . . . . .	15

<b>4</b>	<b>Forskningsmetode</b>	<b>17</b>
4.1	Kvantitativ og kvalitativ forskningsstrategi . . . . .	17
4.1.1	Kvantitativ forskningsstrategi . . . . .	17
4.1.2	Kvalitativ forskningsstrategi . . . . .	17
4.2	Ulike forskningsmetoder . . . . .	17
4.2.1	Ekspirimentelt studie . . . . .	18
4.2.2	Tverrsnittstudier og langsgående studier . . . . .	18
4.2.3	Casestudier . . . . .	18
4.2.4	Komparative studier . . . . .	18
4.3	Datainnsamlingsteknikker . . . . .	19
4.3.1	Spørreundersøkelser . . . . .	19
4.3.2	Intervjuer . . . . .	19
4.3.3	Observasjon . . . . .	20
4.4	Valg av metode i oppgaven . . . . .	21
4.4.1	Overordnet forskningsspørsmål . . . . .	21
4.4.2	Identifisering av problemer . . . . .	22
4.4.3	Nytteverdi med øyesporing for utviklere . . . . .	22
4.4.4	Nytteverdi i dialog med kunder og utviklere . . . . .	23
4.4.5	Praktiske og metodiske utfordringer . . . . .	23
<b>5</b>	<b>Case: Systemutviklingsprosjekt for Amnesty</b>	<b>25</b>
5.1	Hva er Amnesty? . . . . .	25
5.2	Systemutviklingsprosjektet . . . . .	25
5.2.1	Min rolle i prosjektet . . . . .	26
5.3	Struktur av nettsiden . . . . .	26
5.3.1	Forsiden . . . . .	26
5.3.2	Aksjon . . . . .	29
5.3.3	Aktuelt . . . . .	32
5.3.4	Støtt oss . . . . .	33
5.3.5	Om Amnesty . . . . .	34
5.3.6	Ressurser . . . . .	35
<b>II</b>	<b>Gjennomføring</b>	<b>37</b>
<b>6</b>	<b>Gjennomføring av brukbarhetstesten</b>	<b>39</b>
6.1	Forberedelser . . . . .	39
6.1.1	Testlokaler . . . . .	39
6.1.2	Utstyr . . . . .	40
6.1.3	Pilottest . . . . .	40
6.2	Testpersoner . . . . .	41
6.2.1	Rekruttering . . . . .	41
6.2.2	Testdeltakere . . . . .	41
6.3	Testprosedyre . . . . .	42
6.4	Testoppgaver . . . . .	43

<b>III</b>	<b>Resultater</b>	<b>45</b>
<b>7</b>	<b>Problemidentifisering</b>	<b>47</b>
7.1	Resultater fra brukbarhetstesting . . . . .	47
7.1.1	Satisfaction . . . . .	47
7.1.2	Effectiveness . . . . .	48
7.2	Problemområder . . . . .	51
7.2.1	Signering av aksjon . . . . .	51
7.2.2	Fremmedord . . . . .	54
7.2.3	Navigasjon . . . . .	55
7.2.4	Hva pengene som samles inn går til . . . . .	56
<b>8</b>	<b>Intervju med NetLife Research</b>	<b>59</b>
8.1	NetLife Research . . . . .	59
8.1.1	Noen kunden vil ha . . . . .	59
8.1.2	Hvordan gjennomfører de en test med øyesporing? . . . . .	59
8.2	Øyesporing som analyseringsverktøy . . . . .	60
8.2.1	Understøtting av konklusjoner . . . . .	60
8.2.2	Nye oppdagelser . . . . .	60
8.2.3	Kombinasjonen med testresultater . . . . .	60
8.3	Presentasjon av resultater . . . . .	60
8.3.1	Slagkraftig kommunikasjonsverktøy . . . . .	60
8.3.2	Varmekart og blikkdiagram . . . . .	61
8.4	Som hjelpeverktøy under selve testen . . . . .	61
8.5	Ulemper . . . . .	61
8.5.1	Testpersoner . . . . .	61
8.5.2	Tekniske utfordringer . . . . .	61
<b>9</b>	<b>Presentasjon for Amnesty</b>	<b>63</b>
9.1	Setting og gjennomføring av møtet . . . . .	63
9.2	Opplysende og overbevisende for kunden . . . . .	64
9.3	Informasjon fra varmekart . . . . .	65
<b>10</b>	<b>Praktiske og metodiske utfordringer</b>	<b>67</b>
10.1	Programvaren . . . . .	67
10.2	Sittestilling og avstand til øyesporeren . . . . .	68
10.3	Nøyaktighet av målingene . . . . .	68
10.4	Signal fra øyesporeren faller ut . . . . .	68
<b>IV</b>	<b>Drøfting og konklusjon</b>	<b>71</b>
<b>11</b>	<b>Drøfting</b>	<b>73</b>
11.1	Problemidentifisering . . . . .	73
11.1.1	Visuelle vollgraver . . . . .	73
11.1.2	Bekrefting eller forsterking av konklusjoner . . . . .	73
11.1.3	Forstå brukernes mentale modell av systemet . . . . .	75
11.1.4	Oppsummering . . . . .	75

11.2	Nytteverdi for brukbarhetstestere . . . . .	75
11.2.1	Tilpassing av oppfølgingsspørsmål under testen . . . . .	76
11.2.2	Nyttig i dialog med kunden . . . . .	76
11.2.3	Må ha effektiv analyse av data . . . . .	76
11.2.4	Oppsummering . . . . .	77
11.3	Som kommunikasjonsverktøy . . . . .	77
11.3.1	Bruk av varmekart (heat maps) . . . . .	77
11.3.2	Oppsummering . . . . .	78
11.4	Praktiske og metodiske utfordringer . . . . .	78
11.4.1	Utstyret kan være ustabil . . . . .	79
11.4.2	Nøyaktighet av målinger . . . . .	79
11.4.3	Ikke alle egner seg som testdeltakere . . . . .	79
11.5	Metodediskusjon . . . . .	80
11.5.1	Forskningsspørsmål 1 . . . . .	80
11.5.2	Forskningsspørsmål 2 . . . . .	81
11.5.3	Forskningsspørsmål 3 . . . . .	82
11.5.4	Forskningsspørsmål 4 . . . . .	82
<b>12</b>	<b>Konklusjon</b>	<b>85</b>
12.1	Konklusjon . . . . .	85
12.2	Videre forskning . . . . .	86
	 <b>Vedlegg</b>	 <b>87</b>
<b>A</b>	<b>System Usability Scale</b>	<b>89</b>
<b>B</b>	<b>Pilotoppgaver</b>	<b>91</b>
<b>C</b>	<b>Intervjurunde etter test</b>	<b>93</b>
<b>D</b>	<b>SUS-resultater</b>	<b>95</b>
<b>E</b>	<b>NetLife Research - intervju</b>	<b>97</b>
<b>F</b>	<b>Presentasjon for Amnesty</b>	<b>101</b>
<b>G</b>	<b>Episode fra møtet med Amnesty</b>	<b>113</b>
<b>H</b>	<b>Intervju med Terje Røsand</b>	<b>117</b>
	<b>Referanser</b>	<b>121</b>



---

## Figurer

---

2.1	Adjektiver koblet til SUS-verdier . . . . .	8
3.1	Eksempel på øyesporer som festes på hodet . . . . .	12
3.2	Skjerm med integrert øyesporer fra Tobii . . . . .	12
3.3	Frittstående øyesporer fra Tobii . . . . .	12
3.4	Hvordan øyesporeren fungerer . . . . .	13
3.5	Eksempel på blikkdiagram . . . . .	15
3.6	Eksempel på varmekart . . . . .	16
3.7	Eksempel på interesseområder (AOIer) . . . . .	16
5.1	Sidekart . . . . .	27
5.2	Forsiden til Amnesty . . . . .	28
5.3	Aksjonssiden . . . . .	29
5.4	Flere aksjoner . . . . .	30
5.5	Aksjonsartikkel . . . . .	31
5.6	Aktueltsiden . . . . .	32
5.7	Støtt oss . . . . .	33
5.8	Om Amnesty . . . . .	34
5.9	Ressurser . . . . .	35
6.1	Bilde av testområdet . . . . .	40
6.2	Bilde av en Tobii Eye Tracker . . . . .	40
7.1	Konfidensintervaller av SUS for ny/gammel nettside . . . . .	47
7.2	Blikkdiagram(1) . . . . .	52
7.3	Blikkdiagram(2) . . . . .	53
7.4	<i>Audiovisuelt arkiv</i> : Tid fra første fiksering til klikk . . . . .	54
7.5	Varmekart for aksjonssiden . . . . .	55
7.6	Blikkdiagram A . . . . .	57
7.7	Blikkdiagram B . . . . .	57
7.8	Blikkdiagram C . . . . .	57
7.9	Blikkdiagram D . . . . .	57
9.1	Bilde fra presentasjonen for Amnesty . . . . .	64
9.2	Hvor de skulle signere . . . . .	65
9.3	Hvor de faktisk signerte . . . . .	65
9.4	Varmekart av forside (kvinner) . . . . .	66
9.5	Varmekart av forside (menn) . . . . .	66
10.1	Aksjonskalender . . . . .	69
11.1	Visuell vollgrav . . . . .	74



---

## Tabeller

---

6.1	Oversikt over testdeltakere . . . . .	42
7.1	Effectiveness . . . . .	48
7.2	Problemoppgaver . . . . .	49
7.3	Problemidentifisering . . . . .	50
D.1	SUS-resultat . . . . .	95



## Introduksjon

---

Denne rapporten er en masteroppgave skrevet ved Norges Teknisk-Naturvitenskapelige Universitet (NTNU) ved institutt for datateknikk og informasjonsvitenskap. Vi skal her se litt nærmere på bruken av øyesporing innenfor brukbarhetstesting. Her i introduksjonen vil jeg fortelle litt om motivasjonen for oppgaven (1.1), og definere hvilke forskningsspørsmål vi skal prøve å besvare i denne rapporten (1.2). Til slutt vil jeg også gi en oversikt over innholdet i rapporten (1.3).

### 1.1 Motivasjon

Øyesporing er en teknologi som har vært av interesse for forskere siden slutten av 1800-tallet, men opp igjennom 1900-tallet har den tekniske delen stort sett vært for stor til at man kan benytte øyesporing i industrien. Nå har plutselig den tekniske delen blitt god nok og nøyaktig nok til å kunne bruke øyesporing som del av brukbarhetstesting i virkelige systemutviklingsprosjekter, og man står i den situasjonen at man ikke helt vet om man tør å satse på dette eller om man skal vente litt til for å se om det virkelig vil lønne seg å investere i slikt utstyr. Det er derfor et stort behov for studier som sammenligner resultater fra tradisjonell brukbarhetstesting og tester med bruk av øyesporing, som Goldberg med flere nevner i [8]. Denne rapporten er et forsøk på å gjøre nettopp dette og det bringer oss over på forskningsspørsmålet jeg skriver om under.

### 1.2 Forskningsspørsmål

Tradisjonelt har brukbarhetstesting i systemutvikling brukt teknikker som video- og lydopptak, høyttenkning (think-aloud) og intervjuer. Konseptet med bruk av øyesporing i slike tester har blitt sett på som svært lovende en stund, men det er forholdsvis nylig at det tekniske utstyret har blitt godt nok til å kunne brukes til et slikt formål. Men selv om metoden har blitt karakterisert som “lovende”, betyr ikke det nødvendigvis at det er verdt kostnaden med å investere i slikt utstyr. Derfor skal vi i denne rapporten se litt nærmere på hvor nyttig det egentlig er å bruke øyesporing som del av brukbarhetstesting i systemutviklingsprosessen. Det overordnede forskningsspørsmålet vårt blir derfor:

*Har øyesporing en nytteverdi som del av brukbarhetstesting i systemutvikling?*

Etterhvert som vi jobbet med stoffet var det 4 mer spesifikke spørsmål som utkrystalliserte seg. Disse var altså ikke klart definerte på forhånd, men kom til underveis i prosessen, gjennom det som kalles et induktivt forskningsdesign. Jeg har valgt å ikke ta med et eget kapittel om hvordan vi kom fram til disse, men vil integrere det i metodekapittelet under 4.4. Spørsmålene har jeg listet opp her:

1. I hvilken grad bidrar øyesporing til å identifisere problemer?
2. Hvilken nytteverdi ser utviklere som til daglig bruker øyesporing i brukbarhetstesting?
3. Hva er nytteverdien av øyesporing som del av systemutvikling i dialogen med kunder og utviklere?
4. Hvilke praktiske og metodiske utfordringer møter man på?

Jeg vil også trekke fram at denne oppgaven ikke er et forsøk på å kvantifisere økonomiske betraktninger angående bruken av øyesporing i systemutvikling, men heller en kvalitativ undersøkelse hvor vi ønsker å gjøre interessante oppdagelser.

### 1.3 Struktur av rapporten

Her vil jeg gi en presentasjon av hvordan rapporten er oppbygd og fortelle litt om hva hvert kapittel omhandler.

Første del av rapporten inneholder en del bakgrunnstoff og teori om det vi skal se nærmere på senere i oppgaven. Kapittel 2 tar for seg begreper innenfor brukbarhet og vi ser litt nærmere på hvordan brukbarhetstesting tradisjonelt gjøres. I kapittel 3 forteller jeg mer om hva øyesporing er og introduserer noen begreper som brukes innenfor dette. Her sier jeg også litt om hvordan synet vårt fungerer. Videre gir jeg en oversikt over hvilke forskningsmetoder som finnes og beskriver hvilke metoder jeg har valgt for denne oppgaven (kapittel 4). Til slutt i denne delen presenterer jeg caset jeg benytter meg av i studien, det finner du i kapittel 5.

Neste bolk tar for seg hvordan jeg gjennomførte brukbarhetstesten (kapittel 6). Her forklarer jeg forberedelsene mine i forkant av testen, og gir også en beskrivelse av testdeltakerne og testopplegget.

Del III tar for seg resultatene jeg fikk etter datainnsamling for hvert av de fire forskningsspørsmålene. Kapittel 7 onhandler resultatet etter brukbarhetstesting, og fokuserer på å utheve de områdene hvor øyesporingen hadde noe å si. I kapittel 8 beskriver jeg resultatet fra intervjuet jeg gjorde med Thor Fredrik Eie fra NetLife Research. Videre forteller jeg om møtet jeg hadde med Amnesty for å presentere resultatene fra brukbarhetstesten i kapittel 9. Siste del av denne bolken (kapittel 10) tar for seg resultatet jeg samlet inn for å svare på det fjerde forskningsspørsmålet.

Kapittel 11 tar for seg drøfting rundt resultatene jeg har samlet inn i de fire foregående kapitlene, og har også med en diskusjon angående metodene jeg valgte for å besvare forskningsspørsmålene.

Til slutt i rapporten (kapittel 12) forsøker jeg å formulere en konklusjon for det vi har funnet ut i denne studien, og nevner også litt om muligheter for videre forskning innenfor dette feltet.

Et sammendrag av hele rapporten finner du rett etter forordet, på side iii.

# **Del I**

## **Teori og bakgrunnstoff**

---



### Brukbarhet og brukbarhetstesting

---

Dette kapitlet vil gi en introduksjon til hva brukbarhet er og hvordan vi kan gi en definisjon på det (2.1). Jeg vil også forklare hvorfor og hvordan man gjør brukbarhetstesting slik de fleste kjenner det i 2.2. Videre vil jeg i 2.3 presentere System Usability Scale-metoden som brukes for å måle hvor tilfredse testdeltakere er med brukbarheten på produktet man tester.

#### 2.1 Brukbarhet

Brukbarhet er et område det blir satt mer og mer fokus på rundt om i IT-verdenen. Interessen kan ha med at flere og flere oppdager hvor dårlig design mange brukergrensesnitt faktisk har, og hvilke fordeler gode brukergrensesnitt gir brukerne [24]. Dårlig brukbarhet kan gi irriterte og frustrerte brukere, som igjen kan føre til at de gir opp å bruke systemet. Dette er dårlig nytt for både utviklerne og de som er avhengige av at systemet blir brukt. I noen tilfeller kan det også føre til at brukerne gjør kritiske og uopprettelige feil, hvilket kan resultere i meget alvorlige situasjoner.

##### 2.1.1 Definisjoner og begreper

Så hva er egentlig brukbarhet og hvordan kan man måle det? Det finnes flere definisjoner på hva brukbarhet er, men den mest kjente er fra ISO 9241-11 hvor definisjonen lyder:

“[Usability is defined as] the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.”[19]

Videre defineres begrepene *effectiveness*, *efficiency* og *satisfaction* slik:

**Effectiveness** er hvor nøyaktig og fullstendig brukerne klarer å oppnå bestemte mål.

**Efficiency** handler om hvor mye ressurser brukerne trenger for å gjøre det de skal, f.eks. hvor mye tid de bruker.

**Satisfaction** beskriver hvor fornøyd brukerne er med produktet, altså deres subjektive mening om brukeropplevelsen.

Brukbarhet handler altså om at bestemte brukere (brukere i målgruppen for produktet) skal kunne utføre bestemte oppgaver riktig, effektivt og på en tilfredstillende måte i de omgivelsene produktet er ment å brukes.

## 2.2 Brukbarhetstesting

Målet med brukbarhetstesting er (som navnet logisk nok indikerer) å finne ut om systemet som utvikles er *brukbart* i forhold til målgruppen det lages for og de oppgavene som skal utføres [18, s. 646]. I prinsippet handler brukbarhetstesting om å identifisere problemområder i brukergrensesnittet, som deretter vil lede til nødvendig redesign.

Tid er penger som det heter, og til tidligere i utviklingsprosessen man oppdager problemer i brukergrensesnittet, til enklere er det å gjøre forandringer i designet. Samtidig må man da være klar over at testing på et tidlig stadium ikke vil være like realistisk som testing på et ferdig produkt, og dermed ha sine begrensninger når det gjelder hvilke problemområder det er mulig å oppdage. Brukbarhetstesting i sin enkleste form, og ofte en god måte å få tidlig tilbakemelding på, er rett og slett å få en kollega, kunde eller lignende til å si hva de synes om designet [24, s. 141]. Skal man derimot ha en mer grundig gjennomgang av brukbarheten til produktet bør man gjennomføre en mer formell brukbarhetstest, gjerne i et brukbarhetslaboratorium som har nødvendig utstyr.

### 2.2.1 Utstyr og setting

En brukbarhetstest (eller brukertest) gjennomføres ofte i kontrollerte omgivelser, slik som et brukbarhetslaboratorium [18, s. 646]. Et brukbarhetslaboratorium består som regel av to rom:

- Et **testrom** hvor selve testpersonen sitter og gjennomfører testen av produktet, mens det han gjør blir tatt opp ved hjelp av mikrofoner og kameraer. Et slikt testrom kan av og til gjøres om for å etterligne en virkelig setting hvor produktet skal bli brukt, slik at man skaper den riktige stemningen for testpersonen.
- I **observasjonsrommet**, også kalt **teknisk rom**, kontrollerer man kameraene og opptakene som gjøres i testrommet, og de som skal evaluere testen kan sitte her å følge med på hva som foregår via skjerm og høyttaler. De som observerer kan for eksempel være utviklere eller kunder.

Kameraene i testrommet skal fange opp testpersonens ansiktsuttrykk og kroppsbevegelser ([18, s. 647]). Man har også kameraer rettet mot mus og tastatur slik at man kan følge med på hvordan testpersonen bruker disse, i tillegg til at man avbilder selve skjermen.

### 2.2.2 Prosedyre

En brukbarhetstest går vanligvis ut på at testpersonen skal gjøre noen konkrete oppgaver på brukergrensesnittet man tester. I etterkant av testen har man så en spørreundersøkelse eller et intervju for å måle hvor tilfreds testpersonen er med brukergrensesnittet (Satisfaction-faktoren), i tillegg til en avsluttende intervjurunde om temaer man ikke har kommet inn på tidligere i testen, eller for å utdype problemområder man støtte på under testen. Dette kalles på engelsk for en *debriefing-session*.

I forkant av testen må man ha laget klar oppgavene testpersonen skal i gjennom og hvilke spørsmål man skal stille i etterkant. Dette kalles for en detaljert testplan [24, s. 146]. Oppgavene må være formulert på en enkel og konkret måte slik at testpersonen umiddelbart forstår oppgavens mål.

Antall personer som skal testes blir også bestemt i denne testplanen, i tillegg til hvilke typer testpersoner man trenger (alder, kjønn, utdanning, osv.) og hvordan man skal få tak i disse. I de fleste brukbarhetstester er det kvaliteten på testdataene som betyr noe ikke kvantiteten, man er altså ikke interessert i statistikk basert på testdataene, men hva slags informasjon man kan få ut av de. Derfor er det vanligvis nok med 5-12 personer i en slik brukertest [18, s. 647]. De fleste som jobber med brukbarhetstesting må holde seg innenfor et stramt budsjett, og da er det også en fordel å ha færrest mulig testpersoner. Jacob Nielsen påstår i sin Alertbox på nett at man trenger bare 5 testpersoner for å avdekke de fleste feil i et brukergrensesnitt, og begrunner det med at “*As you add more and more users, you learn less and less because you will keep seeing the same things again and again*” [13].

Før man gjennomfører selve testen, bør man kjøre en *pilottest* slik at man vet at utstyr og testopplegg fungerer godt. Denne pilottesten blir vanligvis gjennomført med 1-3 testpersoner [24, s. 146]. Målet her er først og fremst å sjekke at tekniske detaljer er i orden, at oppgavene er godt formulert og at man holder tidsplanen noenlunde. Denne testen blir gjerne utført uken før selve testen slik at man fremdeles har tid til å gjøre endringer i testopplegget.

På selve testdagen, når deltakerne kommer til testlokalet, får de en beskrivelse av testprosedyren, hvorfor man gjennomfører en slik test og hvilke oppgaver de skal i gjennom. Man gjør det også klart for deltakerne at det er *systemet* som blir testet, og ikke de selv, slik at de ikke skal føle seg dumme om de ikke får til oppgavene. De blir også gjort oppmerksomme på at de vil bli filmet og tatt lydopptak av. Noen føler seg litt ukomfortable med dette i starten, men det viser seg at de fleste glemmer dette helt etter et par minutter inn i testen og de konsentrerer seg derimot fullt om oppgavene [24, s. 147].

Under testen blir deltakerne spurt om å *tenke høyt* hele veien, det som på engelsk blir kalt “the think-aloud technique”, slik at man enklere kan forstå hva slag mental modell de danner seg av systemet, hva de forventer seg, hva de misforstår, og hvorfor de gjør det de gjør. Et eksempel kan være at de klikket på et ord på en nettside for så å utbryte “*Å nei! Jeg trodde det var en link, for det er jo sånn strek under...*”. Denne høyttenkningsteknikken kan også lede til spontane forslag til redesign underveis [24, s. 147], der noen også kan være bedre enn de utviklerne selv hadde kommet opp med.

## 2.3 System Usability Scale (SUS)

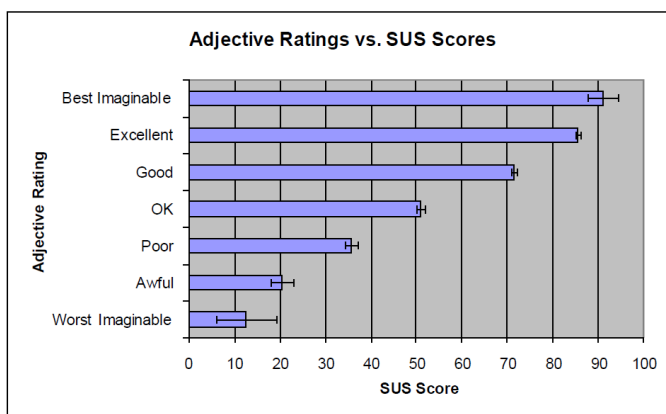
System Usability Scale (ofte forkortet til SUS) ble utviklet for å imøtekomme et økende behov for en enkel og rask måte å måle brukernes subjektive tilfredshetsfølelse etter at de hadde brukt et system [4], det som vi kaller *satisfaction*-faktoren. Det fungerer slik at en bruker får presentert 10 påstander som de skal ta stilling til hvorvidt de er enig eller uenig (se vedlegg A). De svarer på en skala fra 1 til 5, hvor 1 er sterkt uenig og 5 er sterkt enig. Halvparten av påstandene sier noe negativt om systemet, mens den andre halvparten påstår noe positivt om systemet. Dette er for å sikre at brukeren svarer bevisst og ikke bare velger et tilfeldig svar. Brukeren bør gjøre denne delen rett etter at han/hun har gjennomført testen, mens de fremdeles har litt av den samme følelsen for systemet som de hadde under testen. De blir også spurt om å ikke grunne på svaret, men velge det første som faller dem inn. Hvis man skulle gjort de samme oppgavene hjemme i stedet for på laboratoriet, er det jo den umiddelbare reaksjonen på systemet som avgjør om man likte eller ikke likte å bruke det.

### 2.3.1 Beregning av poengsum

Når brukeren har tatt stilling til alle påstandene, kan man beregne den totale poengsummen. Dette vil være et tall mellom 0 og 100, hvor 0 er dårligst tenkelige resultat, mens 100 er det beste. Poengberegningen foregår slik: Verdiene på svarene fra påstand 1, 3, 5, 7 og 9 er lik posisjonen til svaret minus 1 (altså, hvis brukeren har svart 5 - *Sterkt enig*, tilsvarer det 4 poeng). Verdiene på svarene fra påstandene 2, 4, 6, 8 og 10 er lik 5 minus posisjonen til svaret (altså vil svaret 5 - *Sterkt enig* her tilsvare 0 poeng). Til slutt legger man sammen alle poengene fra hvert svar og ganger det hele med 2.5 slik at svaret havner på skalaen 0-100.

### 2.3.2 Betydningen av individuelle poengsummer

SUS gir egentlig bare en tallverdi som for det utrente øyet ikke sier så veldig mye, med mindre man sammenligner det med noe. Aaron Bangor med flere har i [2] prøvd å gjøre noe med dette ved å koble adjektiver til områder på SUS-skalaen. Figur 2.1 viser resultatet disse kom fram til, og gir oss en pekepinn på hvilket inntrykk folk har av systemet på en mer forståelig måte.



Figur 2.1: Figuren viser en oversikt over adjektiver koblet til SUS-verdier. Hentet fra [2].

### 2.3.3 T-testen og statistiske begreper

Skal man sammenligne SUS-verdier, er det fornuftig å benytte seg av det som kalles for *t-testen*. Kort fortalt så er t-testen er en metode innenfor statistikken for å avgjøre hvor sannsynlig det er at en hypotese er riktig eller ikke riktig. Man bruker den gjerne for å teste om en gjennomsnittsverdi er lik en bestemt verdi, eller om to gjennomsnittsverdier er like [20, kap. 16]. En vanlig måte å teste det siste på er å sette opp en hypotese  $H_0$  som påstår at gjennomsnittsverdiene i to utvalg er like, og man bestemmer at hvis det er mindre enn 5% sannsynlighet for at hypotesen er riktig, så forkaster man den. Det vil si at når det er under 5% sannsynlig at gjennomsnittsverdiene er like, så sier vi at de er ulike, at forskjellen er *signifikant*.

5% er det som kalles for *signifikansnivå*, og bestemmes på forhånd. Man beregner så sannsynligheten for at gjennomsnittsverdiene faktisk er like, og dette kalles *p-verdien* [23]. Hvis altså  $p < 0,05$ , så forkaster vi hypotesen.

Når man skal presentere resultatet, bruker man gjerne begrepet *konfidensintervall* som definerer ytre grenser for hvor det er sannsynlig at den virkelige gjennomsnittsverdien for et utvalg befinner seg. Et 95%-konfidensintervall som går fra 40 til 60 for eksempel, sier oss at den virkelige verdien befinner seg innen for 40 og 60 med 95% sannsynlighet.



## Øyesporing

---

Her skal vi se litt nærmere på hva øyesporing egentlig er (3.1) og hvilke i felt øyesporing blir benyttet (3.2). Vi skal også se på hvordan synet vårt fungerer i 3.3, og til slutt vil jeg introdusere noen begreper som blir mye brukt innenfor øyesporing (3.4).

### 3.1 Hva er øyesporing?

Øyesporing eller *eyetracking* (som det heter på engelsk) handler kort sagt om å følge øyets bevegelser, eller som Nielsen og Pernice sier: “*Eyetracking is simply following the trail of where a person is looking*” [15, s. 3]. Dette vil si at man registrer hvor, i hvilken rekkefølge og hvor lenge øyet fikserer på en flate eller et objekt. Eller man kan også registrere øyets posisjon i forhold til hodet uten at øyet nødvendigvis er fokusert på noe spesielt.

#### 3.1.1 Utvikling fram til i dag

De første spede forsøkene på å spore øyebevegelser startet allerede på slutten av 1800-tallet og konsentrerte seg først og fremst om forskning på lesing [17]. Teknologien som ble brukt var imidlertid meget ubehagelig for testpersonene og innebar for eksempel elektroder festet til huden rundt øyet for å kunne måle forskjeller i elektrisk potensial og dermed registrere øyebevegelser. En annen metode gikk ut på å bruke store kontaktlinser som dekket både hornhinnen og sclera (det hvite rundt hornhinnen) med en metalltråd rundt kanten, som da registrerte variasjoner i et elektromagnetisk felt.

Innenfor brukbarhet ble øyesporing først benyttet av Paul Fitts med kolleger i 1947. De brukte da filmkamera til å se på hvordan en pilots øyne beveget seg over knapper og instrumenter i en flycockpit under landing av flyet [10, s. 574]. Selv om metodene da var ganske primitive i forhold til dagens teknologi, greide de å oppdage noen viktige prinsipper som fremdeles er brukbare i dag. Et eksempel er at fikseringsfrekvensen, altså hvor ofte personen ser på et element, forteller noe om hvor viktig elementet er. Eller hvis man ser på fikseringsvarigheten, altså hvor lenge personen ser på et visst punkt, så kan en lang varighet bety at elementet er vanskelig å forstå hva er eller hva det gjør [10, s. 576].

Likevel, i siste halvdel av 1900-tallet ble øyesporingmetoden innenfor brukbarhet aldri klassifisert som noe mer enn *lovene*. Robert J. K. Jacob og Keith S. Karn mener at grunnen til dette kan være fordi de tekniske problemene med øyesporing har hittil vært for store, samtidig som det krever mye ressurser å hente ut nyttig øyesporingsdata, og dataene kan også ofte være vanskelige å tolke [10, s. 577].



Figur 3.1: Et eksempel på en gammeldags øyesporer som man fester på hodet.



Figur 3.2: Skjerm med integrert øyesporer fra Tobii. Kilde: [http://www.acuity-ets.com/images/products\\_t-series.jpg](http://www.acuity-ets.com/images/products_t-series.jpg)



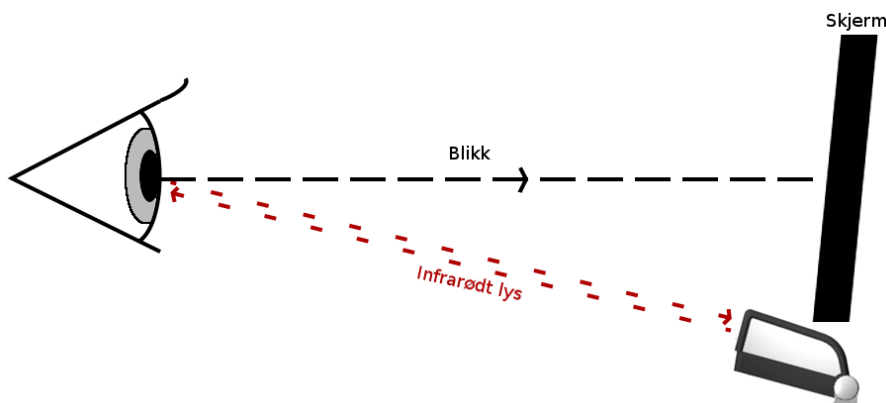
Figur 3.3: Frittstående øyesporer fra Tobii, som kobles til en datamaskin Kilde: [http://www.pstnet.com/images/hardware/TobiiX60\\_medium.jpg](http://www.pstnet.com/images/hardware/TobiiX60_medium.jpg)

Bilde 3.1 viser et eksempel på en øyesporer som man fester på hodet. Det er klart at å ha på seg noe slikt vil være ubehagelig for brukeren, og dette kan gjøre det vanskeligere å konsentrere seg om det han skal gjøre. Heldigvis har teknologien blitt mye forbedret, og i dag bruker man helst øyesporere som enten er integrert i dataskjermen (se fig. 3.2) eller en frittstående øyesporer som kan brukes til flere ting enn bare til dataskjerm (se fig. 3.3). I forhold til utstyret som Fitts brukte rundt 1950, så er dagens moderne øyesporere utrolig enkle å både installere og bruke.



### 3.1.2 Hvordan fungerer dagens øyesporere?

Bruker man en øyesporer i dag, kan man sitte helt vanlig framfor skjermen, som om man var hjemme. Brukerne kan bevege hodet fritt innenfor et gitt område og slipper å ha på seg spesialutstyr. Ideen bak denne typen øyesporere baserer seg på infrarød teknologi. Infrarødt lys sendes fra øyesporeren og blir reflektert i øynene til brukeren (se illustrasjon i figur 3.4). Det reflekterte lyset blir så registrert av en sensor i øyesporeren og øyets posisjon blir beregnet [27]. Bruker man en frittstående øyesporer (som i figur 3.3), må man også ta hensyn til dennes posisjon i forhold til skjermen og skjermens størrelse når man konfigurerer programvaren.



Figur 3.4: Figuren viser hvordan infrarød stråling sendes mot øyet og blir reflektert tilbake til øyesporeren.

## 3.2 Hvor benyttes øyesporing?

Det finnes flere ulike felt hvor øyesporing er et tema. Man var for eksempel tidlig ute med å bruke øyesporing til å forske på hvordan mennesker leser (som nevnt i 3.1.1). Øyesporing er blitt mye brukt innenfor psykologien for forskning på hvordan øyet reagerer på ulike visuelle stimuli og hvordan kognitive prosesser kan kobles til øyets bevegelser [10, s. 575]. Å bruke øyesporing som input-enhet (i stedet for mus eller tastatur for eksempel) er også et interessant felt, kanskje spesielt fordi det kan åpne uante muligheter for handikappede.

I følge Duchowski blir øyesporing innenfor brukbarhet benyttet på to ulike måter [5, kap. 14]:

- *Forskningsmessig* for å kunne dokumentere hvilke visuelle elementer til et brukergrensesnitt som fungerer best. For eksempel hvilke type knapper er best, hvordan man skal gruppere elementer, osv. Det blir også brukt til å utvikle retningslinjer innenfor design av brukergrensesnitt.
- *Industrimessig* for å kunne evaluere brukbarhet av for eksempel websider eller programvarer og bruke resultatet til å gjøre nødvendige forandringer.

I denne rapporten fokuserer vi på det siste punktet.

### 3.3 Hvordan fungerer synet vårt?

Øynene våre bombarderes med enorme mengder informasjon hvert øyeblikk. Det er umulig for hjernen vår å benytte seg av alt, derfor er vi laget slik at vi filtrerer ut alt som ikke virker interessant. Hvis vi for eksempel leter etter noe spesielt, kaster vi ikke bort tid på å se der vi ikke tror det kan være. En annen ting er at folk har en tendens til å se på det de tenker på. Alex Poole og Linden J. Ball hevder i [17] at “*what a person is looking at is assumed to indicate the thought “on top of the stack” of cognitive processes*”. Altså er det naturlig å anta at hvis vi kan spore hvor de ser, kan vi si noe om hva de tenker på.

Synet vårt deles gjerne inn i det som kalles *skarpsyn* og *sidesyn*. Skarpsynet vårt utgjør bare en liten del av netthinnen vår, men gir til gjengjeld veldig detaljrikt inntrykk av det man ser på, man kan si at det har høy oppløsning. Sidesynet har mye mindre kapasitet når det gjelder detaljer, men det dekker derimot et mye større området. Duchowski forklarer i [5, kap. 1] litt mer om hvordan øyet beveger seg. Han sier at synet vårt følger en syklisk prosess i tre steg:

1. Hele “bildet” sees parallelt gjennom sidesynet vårt, hvilket har veldig dårlig oppløsning, og man registrerer noe interessant som man vil inspisere nærmere.
2. Oppmerksomheten vår skruser av på skarpsynet vårt mens øynene raskt flytter seg til området som man oppfattet som interessant i punkt 1. Dette kalles gjerne en sakkade (se 3.4). Man tar altså ikke inn informasjon i synsfeltet under en sakkade.
3. Når øynene har flyttet seg dit de skal, er skarpsynet vårt rettet mot det interessante området, og oppmerksomheten kobles inn igjen for å tolke inntrykket i høy oppløsning. Dette er det som kalles for en fiksering (se 3.4).

Så hva blir oppfattet som interessant nok til å flytte skarpsynet for å undersøke nærmere? I følge Goldberg og Kotval fokuserer øynene naturlig på områder som er overraskende, fremtredende eller viktig utfra tidligere erfaringer [7].

Siden vi som regel ikke har kapasitet eller tid til å fokusere på alt (på for eksempel en nettside), *skanner* vi ofte over innholdet heller enn å lese det. Nielsen og Pernice skriver om noen forsøk de gjorde med øyesporing på ulike nettstedet. De oppdaget at noen sider var enklere å bruke enn andre til tross for mye tekst og informasjon på sidene. Grunnen til dette var at teksten var delt opp i seksjoner med overskrifter i en annen farge og fetere skrift, noe som gjorde at brukerne enkelt kunne skanne nedover siden til de fant informasjonen de var ute etter [15, s. 78].

### 3.4 Viktige begreper

Innenfor området øyesporing er det noen begreper som blir mye brukt når man analyserer resultatene. Her kommer en nærmere forklaring på de viktigste begrepene jeg vil benytte meg av i denne rapporten. Siden den engelske versjonen er mest kjent vil denne stå i parentes på alle sammen.

### 3.4.1 Fiksering (Fixation)

Når øynene fester blikket på et spesielt punkt kalles det for en fiksering. Det er her øynene fanger opp informasjon og tolker den. En fiksering varer normalt mellom 0,25 til 0,5 sekund. [12].

### 3.4.2 Sakkader (Saccades)

Sakkader er betegnelsen på de raske, frivillige bevegelsene øyet gjør for å flytte blikket fra et punkt til ett nytt, altså mellom to fikseringer. Varigheten av en sakkade er vanligvis mellom 0,01 til 0,1 sekund, altså en del mindre enn en fiksering.

### 3.4.3 Blikkdiagram (Gaze plot)



Figur 3.5: Eksempel på blikkdiagram. Hentet fra [26].

Et blikkdiagram viser hvor lenge øynene har fiksert på hvert punkt og i hvilken rekkefølge det er gjort. Et eksempel på et blikkdiagram er vist i figur 3.5. En sirkel representerer en fiksering der større sirkel indikerer lenger fiksering på det punktet. Sakkader er representert ved streker mellom fikseringene. Ved å nummerere sirklene får man fram i hvilken rekkefølge fikseringene har funnet sted. Ulike brukere blir skilt ved å bruke ulike farger.

### 3.4.4 Varmekart (Heat maps)

Et varmekart sier ikke noe om rekkefølgen av fikseringer og lite om enkeltfikseringer. Derimot viser det hvor konsentrasjoner av fikseringer oppstår, det vil si hvor øynene har fiksert flest ganger. Figur 3.6 viser et eksempel på varmekart. Til rødere et område er til flere fikseringer har det vært på det punktet. Et varmekart skiller ikke mellom ulike brukere, men gir en samlet oversikt for alle som har fiksert på de ulike punktene.

### 3.4.5 Interesseområde (Area Of Interest — AOI)

Ved å markere områder av spesiell interesse kan man ut fra dette hente ut ulik statistisk informasjon om disse områdene. Slike områder blir ofte referert til som AOIer (fra engelsk Area Of Interest). Figur 3.7 viser hvordan logoen i nedre venstre hjørne er markert



Figur 3.6: Eksempel på varmekart. Hentet fra [26].

som et interesseområde. Typiske interesseområder er for eksempel et bilde, en overskrift, annonser for reklame, menyelementer, osv. Man kan så hente ut informasjon om f.eks. antall ganger det valgte interesseområdet ble besøkt, hvor ofte det ble besøkt (frekvensen), hvor lenge det ble besøkt per gang, hvor lang tid det gikk før området ble besøkt første gang, hvor mange av brukerne som besøkte området, osv.



Figur 3.7: Eksempel på markering av interesseområder (AOIer). Hentet fra [26].

## Forskningsmetode

---

Her vil jeg gi en oversikt over hvilke metoder og strategier som er vanlig innenfor forskning (4.1-4.2), og også fortelle litt om hvilke teknikker det er vanlig å bruke for å samle inn data til forskningen (4.3). I 4.4 vil jeg fortelle hvilke metoder jeg har valgt å benytte meg av i denne masteroppgaven, og hvorfor jeg valgte disse.

### 4.1 Kvantitativ og kvalitativ forskningsstrategi

En forskningsstrategi handler om hvilken metode man skal bruke for å oppnå det resultatet man ønsker. Det finnes hovedsaklig to ulike strategier man kan velge mellom, kvantitativ eller kvalitativ forskningsstrategi [20, s. 91]. Hvilken man velger avhenger av problemstillingen man har og hvilken praktisk nytte det gir å bruke hver av de. Det er i dag vanlig å se på disse to strategiene som komplementære, altså at de utfyller hverandre. Vi skal se litt nærmere på hva kvantitativ og kvalitativ forskningsstrategi går ut på under.

#### 4.1.1 Kvantitativ forskningsstrategi

Kvantitativ forskningsstrategi handler om å samle inn data eller tall for å bekrefte eller avkrefte en hypotese eller et spørsmål som er framsatt på bakgrunn av en teori. Man har et stort utvalg for at svarene skal være representative og statistisk gyldige for det området man gjør undersøkelsen i [20, s. 92].

#### 4.1.2 Kvalitativ forskningsstrategi

Kvalitativ forskningsstrategi går ut på sette seg dypere inn i en problemstilling. Det handler om forstå situasjoner og handlinger, og å oppdage nye informasjon. Dette krever en stor nærhet til studieobjektet og gjør dermed at store utvalg blir vanskelig. Til forskjell fra kvantitativ strategi vil svarene her ofte være i tekstform og analyseres dermed med mer uformelle teknikker [20, s. 92].

### 4.2 Ulike forskningsmetoder

Et forskningsmetode forteller noe om hvordan forskeren har tenkt å gjennomføre undersøkelsen. I følge Kristen Ringdal finnes det generelt 5 typer forskningsmetoder: Eksper-

imentelt, tverrsnitt, langsgående, casestudie og komparativt studie [20, s. 93]. For helhetens del vil jeg gi en kort beskrivelse av alle disse typene studier:

### 4.2.1 Eksperimentelt studie

Et eksperimentelt studie handler om å finne årsakssammenhenger og blir gjerne utført i strengt kontrollerte omgivelser. Man har gjerne to variabler X og Y, og når forskeren manipulerer X slik at han får Y, kan han konkludere med at X forårsaker Y. Undersøkelsen gjøres gjerne med to grupper som i utgangspunktet er like, så utføres eksperimentet på den ene gruppen, mens den andre beholdes slik den er som en kontrollgruppe. Hvis forskeren så ser at resultatet han var ute etter, altså Y, finnes i den første gruppen, men ikke i den andre, kan han konkludere med at faktoren X, som han tilførte eksperimentgruppen, var årsaken til forandringen [20, s. 93]. Eksperimentelt design benyttes vanligvis i kvantitativ forskningsstrategi.

### 4.2.2 Tverrsnittstudier og langsgående studier

Tverrsnittstudier handler om å gjøre undersøkelsen på et bestemt tidspunkt. Det benyttes både innenfor kvantitativ og kvalitativ forskningsstrategi. Et eksempel på det første er spørreundersøkelser på et stort utvalg, mens for eksempel intervjuer vanligvis hører til kvalitativ forskningsstrategi. Det som er spesielt med tverrsnittstudier er at data samles bare inn én gang for hver analyseenhet, og man søker å beskrive analyseenhetene slik de er da undersøkelsen blir gjennomført [20, s. 94].

I langsgående studier bruker man de samme teknikkene som i tverrsnittstudier, men forskjellen ligger i at man gjør datainnsamlingen flere ganger over tid. Hensikten med en slik metode er å studere eventuelle endringer over tid eller å følge utviklingen i en prosess [20, s. 94].

### 4.2.3 Casestudier

Casestudier handler om å konsentrere forskningen rundt et case eller et analyseobjekt. Man setter seg inn i en spesiell situasjon eller tilværelse, for eksempel en bedrift, en organisasjon eller et individ. Casestudier blir brukt både i kvalitativ og kvantitativ forskningsstrategi, men skal man bruke kvantitativ strategi må man gjøre dette på underenheter i caset [20, s. 96].

Bruker man bare et case er målet ofte å tolke eller forstå hva som er spesielt med akkurat dette caset. Man kan ikke trekke gyldige generelle slutninger basert på resultatet fra ett case, men man kan se på det som et bidrag til generell teori [20, s. 151].

### 4.2.4 Komparative studier

Komparative studier handler kort sagt om å sammenligne to eller flere case. Målet er å finne en teoretisk interessant sammenligning mellom de valgte analyseenhetene [20, s. 96]

## 4.3 Datainnsamlingsteknikker

Når man har både strategi og metode på plass, må man velge hvordan man vil gå fram for å samle inn de dataene man trenger for å kunne svare på problemstillingen (forskningsspørsmålene). Det finnes hovedsaklig 3 typer datainnsamlingsteknikker: *Spørreundersøkelser*, *intervjuer* og *observasjoner* [20, s. 24]. Jeg gir en beskrivelse av hver av disse her:

### 4.3.1 Spørreundersøkelser

I følge Ringdal kan en spørreundersøkelse defineres som “*en systematisk metode for å samle inn data fra et utvalg personer (bedrifter, organisasjoner) for å gi en statistisk beskrivelse av den populasjonen utvalget er hentet fra*” [20, s. 167]. En spørreundersøkelse er standardisert, det vil si at alle får de samme spørsmålene, og blir hovedsaklig brukt for kvantitative formål. Man kan velge å samle inn svarene ved hjelp av et selvutfyllingsskjema, over telefon eller ved et personlig besøk hos deltakeren.

Hver av disse har sine egne fordeler og ulemper. Et selvutfyllingsskjema er for eksempel billig, kan distribueres over et stort og geografisk spredt utvalg, og garanterer anonymitet for respondenten. Men samtidig har man lite kontroll over riktigheten til svarene og man har ofte veldig stort frafall, altså at mange ikke svarer på undersøkelsen.

Telefonundersøkelser eller personlige besøk gir mye større kontrollmuligheter i forhold til kvaliteten på svarene, men samtidig er det mer kostbart og tidkrevende siden man må ha noen til å gjennomføre disse undersøkelsene. Det begrenser også anonymiteten til respondentene.

Når man skal utføre en spørreundersøkelse er det to viktige steg man må igjennom før man kan samle inn dataene. Det ene er å bestemme *hvem* man skal spørre, altså å velge utvalget. Hvilken populasjon skal man hente de fra, hvem har man mulighet til å spørre, og hvor mange trenger man?

Det andre er å lage og teste spørreskjemaet. Det finnes ulike måter å stille et spørsmål på, man kan ha åpne spørsmål, lukkede spørsmål eller vurderingsspørsmål [20, s. 178-179]. Et åpent spørsmål lar respondenten få formulere et eget svar på spørsmålet, mens et lukket spørsmål kommer med alternativer som han må velge et eller flere av. Vurderingsspørsmål presenterer en påstand og en skala av et eller annet slag som han skal plassere seg på i forhold til påstanden. Et eksempel på en spørreundersøkelse med vurderingsspørsmål er SUS (se kapittel 2.3 og vedlegg A). Hvordan man velger å stille spørsmålene, avhenger av hvilken informasjon man er ute etter.

### 4.3.2 Intervjuer

Intervju som datainnsamlingsteknikk er en mye brukt metode innenfor kvalitativ forskning og casestudier. Et intervju er i prinsippet en samtale mellom mennesker, men har noen forhåndsdefinerte antagelser om personene det involverer, som ikke finnes i normale samtaler [16, s. 186]. Samtalen er gjerne arrangert, og den som intervjuer har et spesielt formål med samtalen: Å få informasjon om et spesielt tema.

I følge Oates [16, s. 187] er det hensiktsmessig å bruke intervjuteknikken når man vil ha detaljert informasjon om et emne eller spørre spørsmål som enten er for komplekse til å

uttrykkes i spørreskjemaer, krever åpne svar eller kanskje må tilpasses hvert av intervjuobjektene. Intervjuer passer også å bruke når man vil utforske følelser eller opplevelser som ikke fanges opp så lett gjennom spørreskjema eller ved observasjon, eller hvis temaet kanskje er så sensitivt at respondenten ikke føler seg komfortabel med å sende svaret skriftlig til noen man ikke kjenner.

Det finnes generelt 3 typer intervjuer [16, s. 187-188]:

- *Strukturerte intervjuer* hvor intervjueren har laget alle spørsmålene på forhånd, og holder seg til disse under hele intervjuet. Denne metoden minner mye om spørreundersøkelser, og intervjueren involverer seg ikke videre i samtalen enn for å få svar på de spørsmålene han har klare.
- *Semi-strukturerte intervjuer* fungerer slik at man har en agenda angående hva man skal snakke om, en liste med temaer man skal ta opp under intervjuet. Men samtalen kan utvikle seg til en viss grad dersom man skulle komme over interessante ting underveis. Intervjuobjektet kan nevne ting som han tror er relevant for temaet, men som intervjueren ikke spør direkte om. Intervjueren på sin side står da fritt til å følge opp dette temaet videre.
- *Ustrukturerte intervjuer* starter gjerne med at intervjueren introduserer et tema, men lar så samtalen utvikle seg selv uten at han prøver å styre den i en spesiell retning.

Både semi-strukturerte og ustrukturerte intervjuteknikker handler om å ville oppdage nye ting, noe som forskeren ikke visste på forhånd, mens strukturerte intervjuer fungerer mer som en "sjekklister", forskeren vil bare ha svar på det han lurte på i utgangspunktet [16, s. 188]

### 4.3.3 Observasjon

Observasjon som datainnsamlingsteknikk brukes både innenfor kvantitativ og kvalitativ forskning. Et problem med intervjuer og spørreundersøkelser kan av og til være at folk ikke rapporterer *korrekt* det de faktisk gjør. Det kan være flere årsaker til dette, temaet man tar opp kan være sensitivt slik at den man spør ikke har lyst å si hva det er han faktisk gjør, eller det kan være at han faktisk ikke husker det man er interessert i å finne ut om, eller til og med husker feil. Mennesker har gjerne en tendens til å huske det de *vil* huske eller store hendelser de har vært med på, mens alt i mellom flyter mer sammen eller går i glemmeboken. Observasjoner unngår dette ved at man ser på hva folk faktisk *gjør* heller enn hva de sier at de gjør [16, s. 202].

Observasjonsteknikker deles gjerne inn i *åpne* (overt) eller *skjulte* (covert) observasjoner [16, s. 203]. I en åpen observasjon er den eller de som blir observert klar over at det skjer. Dette kan ha noen ulemper fordi den det gjelder blir da gjerne mer bevisst sine egne handlinger og det kan være man mister noe av genuiniteten i det man observerer. På den andre siden, når man gjør en skjult observasjon, er ikke den som blir observert klar over at det skjer. Da vil alle handlingene være naturlige, men her kommer det etiske aspektet inn i bildet. Er det *riktig* å "spionere" på folk uten at de vet om det? Dette er avveininger man må ta hensyn til ut fra hva man vil finne ut.

En annen måte å skille ulike observasjonsteknikker på er mellom *systematisk observasjon* og *deltakende observasjon* [16, s. 204]. I systematisk observasjon har forskeren



bestemt hva han skal se etter på forhånd, og handler vanligvis om å notere ned hvor mange ganger en hendelse inntreffer og hvor lenge den pågår.

Deltakende observasjon handler om at forskeren er *tilstede* der det skjer på en eller annen måte. Formålet hans er å oppdage og notere ned *alt* som kan synes å være interessant. Han har altså ikke definert hva han skal se etter på forhånd, men skal prøve å være åpen for alle inntrykk. Han kan også være deltakende på ulike nivåer, som beskrevet i [16, s. 208-209]:

- *Ren observatør* hvor han er tilstede, men ikke gjør noe. Han bare ser og observerer, som en flue på veggen, noe som kan gjøres både skjult og åpent.
- *Ren deltaker* der han gjør alt som de andre også gjør. Han prøver å bli som et fullverdig medlem av den gruppen han observerer. Dette gjøres som regel skjult.
- *Deltaker-observatør* er en som følger i fotsporene til den han observerer og kikker over skulderen. Hvis man klarer å vinne tilliten til den man observerer, slik at denne kanskje til og med glemmer at han blir observert, kan forskeren oppdage mye spennende.
- En *praktiker-forsker* er en som allerede er en del av gruppen man er interessert i å forske på, og som får i oppgave å observere gruppen fra innsiden. Man slipper da ofte å sette seg inn i nye og ukjente situasjoner, og man har naturlig tilgang til det man skal observere. Men man må gjerne gjøre dette med tillatelse fra ledelsen, og det er også en fare for at man ikke oppdager ting som ville vært opplagt for en utenforstående, siden man selv anser det som helt naturlig.

## 4.4 Valg av metode i oppgaven

Her vil jeg forklare hvilke metoder og teknikker jeg har valg å bruke i oppgaven min. Jeg har brukt forskjellige teknikker for å få svar på de ulike forskningsspørsmålene alt etter hva som var mest hensiktsmessig. Under vil jeg koble valget opp mot hver av forskningsspørsmålene (se kapittel 1.2) og begrunne hvorfor jeg har valgt å gjøre det på denne måten. Her har jeg også beskrevet kort hvordan vi kom fram til hvert av underspørsmålene.

### 4.4.1 Overordnet forskningsspørsmål

Hovedmålet med oppgaven er å finne ut om øyesporing har en nytteverdi som del av brukbarhetstesting i systemutvikling. For å få svar på dette valgte vi å bruke et reelt case, siden vi ønsket å få resultatet så nært en virkelig situasjon som mulig. Vi hadde også tilgang til et case via Terje Røsand som hadde hatt kontakt med Amnesty, som holdt på med et systemutviklingsprosjekt og gjerne ville ha en brukbarhetsvurdering av den nye nettsiden de utvikler. Siden vi ikke visste helt hva slags informasjon vi så etter, var det også naturlig å hovedsaklig ha en kvalitativ forskningsstrategi fordi da har man som mål å oppdage nye ting og å prøve å forstå situasjoner og handlinger (se 4.1.2).

## 4.4.2 Identifisering av problemer

I første forskningsspørsmål vil vi vite: *I hvilken grad bidrar øyesporing til å identifisere problemer i brukbarhetstesting?* Dette forskningsspørsmålet kom tidlig i prosessen i og med at det var naturlig å finne ut hvilken informasjon man får ut fra selve øyesporingen i en brukbarhetstest. Vi valgte her å gjennomføre en tradisjonell *brukbarhetstest* med samtidig bruk av øyesporing, hvilket er et type eksperimentert studie (se 4.2.1), men med hovedvekt på det kvalitative resultatet. Vi kunne så gjøre en objektiv sammenligning av hva slags problemer vi fant uten å ta hensyn til øyesporingen, og deretter se på hvilken ekstra informasjon vi fikk fra øyesporingen. Under brukbarhetstesten ville vi bruke både observasjon, intervju og spørreundersøkelse som datainnsamlingsteknikk, slik at resultatene kunne utfylle hverandre.

Observasjon ville vi bruke under testen ved at jeg skulle sitte ved siden av testpersonen og følge med på det han/hun gjorde, og vi ville også ta det på video for å kunne se på det senere. Dette er det som kalles for *åpen* observasjon (se 4.3.3), siden brukeren er klar over at vi gjør dette, og selv om noen kanskje føler det litt ukomfortabelt til å begynne med, så glemmer de det forholdsvis fort når de blir engasjerte i oppgavene.

Etter testen ville vi la deltakeren svare på en spørreundersøkelse for å finne ut hvor tilfredse de var med brukervennligheten på nettsiden de testet. Dette gjorde vi fordi vi ville bruke en standard metode (System Usability Scale, se 2.3) for å måle *Satisfaction*-faktoren (se kapittel 2.1.1) slik at vi kunne sammenligne det med andre nettsider.

Til slutt ville vi også ha et intervju med testdeltakeren som skulle fange opp ting vi ikke hadde oppdaget tidligere i testen. Vi ville bruke *semi-strukturert intervju* (se 4.3.2) fordi det var noen tema vi ville være sikker på at vi gikk igjennom, men samtidig ville vi være åpne for ny informasjon.

I denne delen trengte vi både kvalitativ og kvantitativ informasjon. Kvalitativ fordi vi var interesserte i å finne ut hva slags problemer man kan finne med eller uten øyesporing. Samtidig ville vi ha kvantitativ informasjon for å kunne gi de kvalitative konklusjonene våre litt mer tyngde, og trengte derfor flere testpersoner enn en normal brukbarhetstest.

Hvordan vi gjennomførte denne brukbarhetstesten rent praktisk har jeg skrevet om i kapittel 6 og resultatet fra testen finner du i kapittel 7.

## 4.4.3 Nytteverdi med øyesporing for utviklere

Det andre forskningsspørsmålet vårt spør: *Hvilken nytteverdi ser utviklere som til daglig bruker øyesporing i brukbarhetstesting?* Dette var et spørsmål som ble mer og mer klart etterhvert som jeg jobbet med resultatet fra brukbarhetstesten. Jeg følte at jeg hadde liten erfaring med dette, og lurte på hvordan andre brukte resultatet fra brukbarhetstesting med øyesporing. Det ville derfor være interessant å få litt input utenfra.

For å få svar på dette valgte vi å utføre et *semi-strukturert intervju* med NetLife Research. Det er ikke så mange i Norge som driver med øyesporing i brukbarhetstesting, så derfor er det naturlig å gå litt i dybden på temaet gjennom et intervju heller enn å fange bredden av synspunkter fra mange gjennom en spørreundersøkelse. Resultatet fra dette intervjuet har jeg skrevet om i kapittel 8. Intervjuguiden jeg brukte vil framgå av transkripsjonen av intervjuet i vedlegg E.

#### 4.4.4 Nytteverdi i dialog med kunder og utviklere

I forskningsspørsmål 3 lurer vi på: *Hva er nytteverdien av øyesporing som del av systemutvikling i dialogen med kunder og utviklere?* Det går altså ut på å finne ut hvor nyttig øyesporing er som kommunikasjonsmiddel mellom de som utfører brukbarhetstesten og kunden eller andre utviklere.

Dette spørsmålet ble klart definert etter at vi hadde gjennomført brukbarhetstesten og bestemt at vi skulle intervju noen fra NetLife Research. Vi visste hele tiden at vi skulle presentere problemområdene vi hadde funnet for Amnesty, men jeg hadde ikke tenkt over muligheten for at øyesporingen kunne ha noen effekt overfor kunden utover det at det hadde gitt forbedret problemidentifisering. Men Terje Røsand, som av og til er i kontakt med NetLife Research, hadde forstått at de ofte brukte det som hjelp i kommunikasjonen med kunden. Dette syntes vi var interessant, og ville se nærmere på det.

I vårt tilfelle ville vi bruke oss selv som case og presentere resultatet vi fikk fra testen vi gjorde i første del av oppgaven for Amnesty og de som utvikler sidene deres. Vi ville så bruke deltakende observasjon med meg selv som *praktiker-forsker* i og med at jeg allerede er i rollen som brukbarhetskonsulent, og samtidig ta møtet opp på video slik at vi kan se på det senere. Observasjonen ville gjøres åpent, siden vi kom til å være i deres lokaler og et videokamera ville være godt synlig. Vi følte det derfor naturlig å spørre om tillatelse til å utføre filmingen. Jeg forteller mer om hvordan dette møtet gikk i kapittel 9.

#### 4.4.5 Praktiske og metodiske utfordringer

Den siste delen av oppgaven vår spør: *Hvilke praktiske og metodiske utfordringer møter man på?* Vi tenker da på situasjoner hvor man skal inkludere øyesporing i brukbarhetstesting. Dette spørsmålet dukket opp rett og slett fordi vi *opplevde* problemer, og det var dermed naturlig å ta det med i denne undersøkelsen.

I tillegg til å bruke mine egne erfaringer fra testingen og å sammenligne med hva NetLife Research tenker om denne saken, ville jeg også gjennomføre et *semi-strukturert intervju* med Terje Røsand som jobber mye med brukbarhetslaboratoriet og øyesporeren som hører til Institutt for datateknikk og informasjonsvitenskap (IDI) ved NTNU og Norsk Senter for Elektronisk Pasientjournal (NSEP). I kapittel 10 har jeg skrevet om resultatet fra dette. Intevjuguiden jeg brukte vil også her framgå av transkripsjonen av intervjuet som finnes i vedlegg H.



---

## Case: Systemutviklingsprosjekt for Amnesty

---

I dette kapitlet gir jeg en beskrivelse av caset vi benytter oss av i denne undersøkelsen. Jeg presenterer også nettsiden vi testet og forklarer hvordan navigasjonsstrukturen her hovedsaklig er oppbygd (se 5.3).

### 5.1 Hva er Amnesty?

Amnesty er en ideell organisasjon som jobber for å fremme menneskerettigheter rundt om i hele verden. Organisasjonen ble stiftet i 1961 av advokaten Peter Benensons fra England og har siden den gang vokst seg til en verdensomspennende bevegelse med avdelinger i mer enn 150 land og regioner.

“Kjernen i Amnestys arbeid er: Alle mennesker har rett til liv, frihet og personlig sikkerhet; ingen skal utsettes for tortur eller annen grusom, umenneskelig eller nedverdiggende behandling eller straff; og alle har rett til tanke, menings, religions og ytringsfrihet. Ingen skal utsettes for diskriminering på bakgrunn av sin tro, hudfarge, kjønn, etnisk opprinnelse, språk eller religion.”

[1]

Amnesty tar utgangspunkt i FNs Verdenserklæring om Menneskerettighetene og jobber først og fremst ved å etterforske saker der man har mistanke om disse menneskerettighetene blir brutt. Rapporter fra denne etterforskningen blir publisert og de setter i gang tiltak i form av underskriftskampanjer, lobbyisme og andre aksjonsformer for å stoppe disse bruddene [1].

### 5.2 Systemutviklingsprosjektet

Caset vi tar utgangspunkt i under denne studien er et systemutviklingsprosjekt iverksatt av Amnesty Norge for å fornye nettsidene deres.

Prosjektet har pågått siden februar 2009 og en beta-versjon av nettsidene ble lansert i mars 2010. Nexus Consulting har vært leid inn til forprosjektet og programmering i prosjektet, og er fremdeles med i utviklingen. Bemanningen på utviklingsdelen er på 3 personer (juni 2010).

Vi kom i kontakt med Amnesty via Terje Røsand som opprinnelig hadde tilbudt seg å gjøre en brukbarhetstest av den nye nettsiden deres som et frivillig bidrag til organisasjo-

nen. Etter samtale med professor Dag Svanæs hadde de bestemt at det passet å gjøre som en del av en studentoppgave.

### 5.2.1 Min rolle i prosjektet

Min rolle i dette prosjektet blir som en innleid konsulent med ekspertise på brukbarhet. Jeg gjennomfører da en brukbarhetstest av nettsidene deres og presenterer så senere resultatet for de, slik at de kan bruke dette til å forbedre brukbarheten på nettsidene. At jeg skulle gjøre dette hadde vi avtalt på forhånd med Amnesty. Jeg vil ikke få noen lønn for dette arbeidet, men ser på det som et bidrag til en ideell organisasjon, samt at jeg kan få benytte meg av resultatet i denne studien.

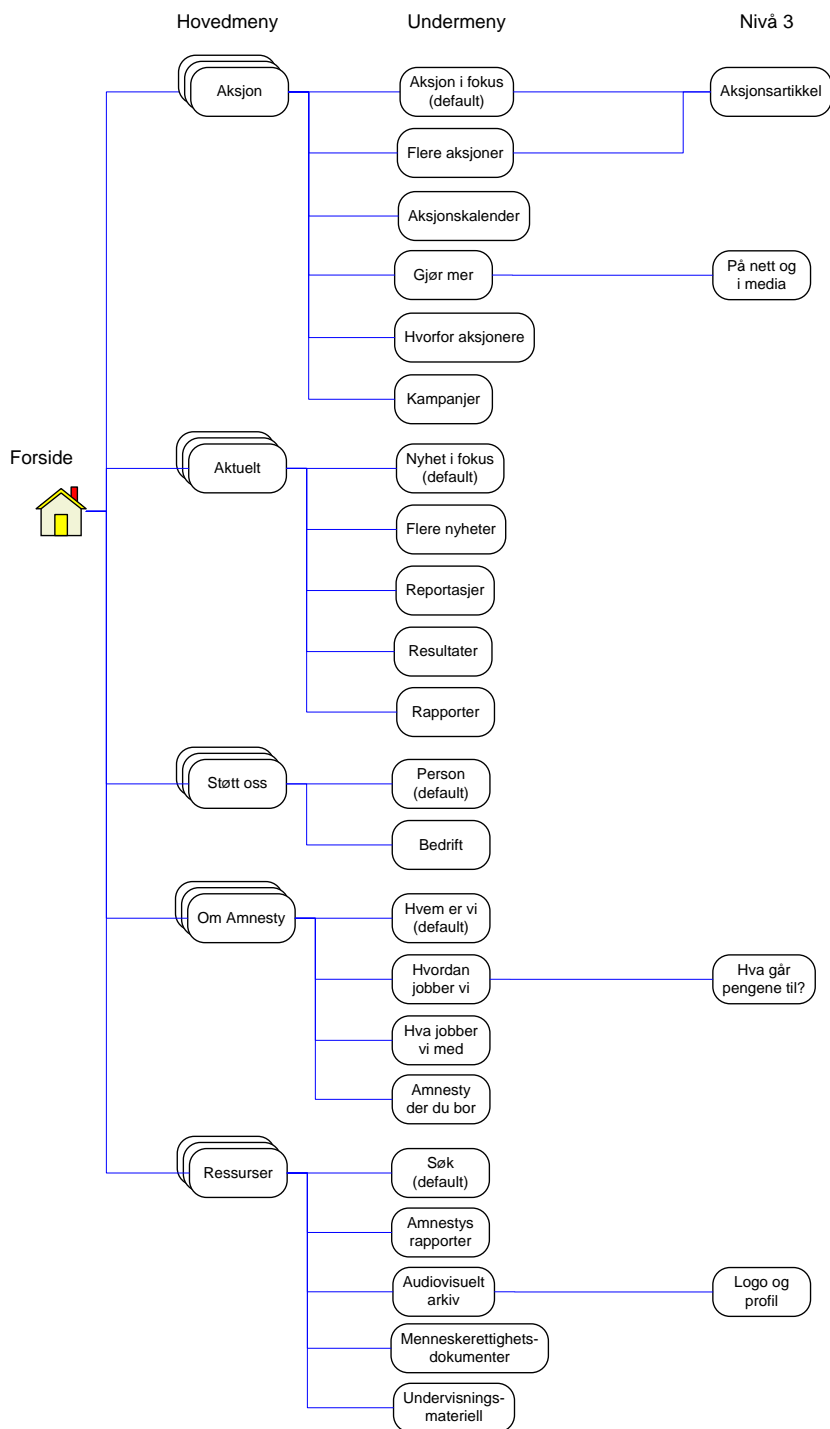
Tidligere har jeg gjennomført en brukbarhetstest av de gamle nettsidene til Amnesty, også i samråd med de, og brukte resultatet av dette i min prosjektoppgave høsten 2009 [25].

## 5.3 Struktur av nettsiden

For å presentere strukturen av nettsiden vil jeg begynne med et konseptuelt sidekart som viser hvordan navigasjonsstrukturen på nettsiden hovedsaklig er lagt opp. Dette viser jeg i figur 5.1. Her har jeg tatt med hvilke alternativer man kan velge på de ulike menyene. På nivå 3 viser jeg kun de sidene det er greit å vite om for senere i rapporten.

### 5.3.1 Forsiden

Jeg vil så begynne med å presentere *forsiden* til nettsiden. Du kan se et skjermbilde herfra i figur 5.2. Øverst til venstre på siden finner vi logoen til Amnesty. En horisontal meny under dette utgjør globalnavigasjonen, vi kaller dette *hovedmenyen*. Vi har også en meny øverst til høyre hvor man finner link til Amnestys eksterne blogg, til et sidekart, innlogging for *Min Side* og en link spesielt for pressen. Under globalnavigasjonen finner vi hovedinnholdet, som på forsiden består av en slags “bildekarusell” som representerer en aktuell nyhet eller aksjon. Man skifter hovedbilde i karusellen ved å klikke på et av de tre små bildene nederst på hovedbildet. Klikker man på hovedbildet, kommer man direkte til selve saken. Til høyre for hovedbildet er en boks med kort info og link til medlemskap i Amnesty. Lenger ned på siden finner vi fire bokser ved siden av hverandre. De to til venstre viser de tre nyeste nyhetssakene og blogginnleggene, vi kaller dette henholdsvis *nyhetsfeed* og *bloggfeed*. Boks nummer tre viser et kart over Norge, hvor man kan klikke på sin lokale region og komme direkte til de aktuelle sidene for sin region. Boksen lengst til høyre er holdt av til annonser. Helt nederst på siden finner man diverse relevant kontaktinformasjon.



Figur 5.1: Konseptuelt sidekart over nettstedet som viser navigasjonsstrukturen på nettstedet. På nivå 3 har jeg bare tatt med de sidene som er nyttige å vite til senere.



Figur 5.2: Her vises forsiden til Amnesty slik den var under testperioden.



### 5.3.2 Aksjon

Klikker man på *Aksjon* i globalnavigasjonen, kommer man til det vi kaller for *Aksjonssiden* (skjerm bilde vises i figur 5.3). Her blir vi også introdusert til undernavigasjonen. Denne finnes som en horisontal linje under hovedmenyen, og også som en vertikal meny til venstre for hovedinnholdet, vi kaller denne for *venstremenyen*. *Aksjon i Fokus*-fanen er valgt som default, og vi ser et illustrasjonsfoto til den aktuelle aksjonen i hovedvinduet. Klikker man på dette bildet, kommer man til selve aksjonsartikkelen. Til høyre for bildet, i stedet for *Bli medlem*-boksen, vises nå det vi kaller *signaturboksen*. Denne består i hovedsak av en teller som viser hvor mange som har signert aksjonen, to tekstfelt for å skrive inn navn og epost, og en knapp for å sende inn signaturen sin. Nederst finner vi igjen fire bokser ved siden av hverandre, som på forsiden, men noe av innholdet er forandret.

Figur 5.3: *Aksjonssiden* som man kommer til ved å trykke på *Aksjon* i hovedmenyen. Man kommer da naturlig inn på det som for tiden er *Aksjon i fokus*.

Velger man *Flere aksjoner* i undermenyen kommer man til skjermbildet vist i figur 5.4. Her vises en liste over de fem nyeste aksjonene som Amnesty jobber med. For å bla til eldre aksjoner, trykker man der det står *Eldre* til høyre over listen.

The screenshot shows the Amnesty International website interface. At the top, there is a navigation bar with links for 'BLOGG', 'SIDEKART', 'MIN SIDE', and 'PRESSE'. Below this is the Amnesty International logo and a search bar. A secondary navigation bar contains 'AKSJON', 'AKTUELT', 'STØTT OSS', 'OM AMNESTY', 'RESSURSER', and 'SØK'. A third bar lists 'AKSJON I FOKUS', 'FLERE AKSJONER', 'AKSJONSKALENDER', 'GJØR MER', 'HVORFOR AKSJONERE', and 'KAMPANJER'. The main content area is a grid of action items:

			ELDRE »	
AKSJON I FOKUS		TEST AV VIDEO Test av Video	01.03.2010	<b>DU</b> Som medlem av Amnesty gir du oss muskler i kampen for menneskerettighetene  <a href="#">BLI MEDLEM &gt;&gt;</a>
FLERE AKSJONER RSS		106 MENN DØMT TIL DØDEN To menn, Abdullah Ali Adam og Al-Murdi Bakheet, ble dømt til døden den 19. januar 2010. Dommen ble avsagt i en av de mange anti-terror...	22.02.2010	
AKSJONSKALENDER		FREDSPRISNOMINERT DØMT TIL 11 ÅRS FENGSEL 1. juledag 2009 ble demokratiforkjemperen Liu Xiaobo dømt til 11 års fengsel for "undergravende virksomhet". Hans forbrytelse var å stø...	18.02.2010	
GJØR MER		RISIKERER LIVET FOR Å GI LIV Hvert eneste år dør over 2.000 kvinner som følge av komplikasjoner under svangerskap og fødsel i Burkina Faso. Signer...	10.02.2010	
HVORFOR AKSJONERE		500 FORSVUNNET. AMNESTY KREVER SVAR Masood Ahmed Janjua, en forretningsmann fra Rawalpindi, "forsvant" 30. juli 2005, på en bussturne til Peshawar. Han skal angivelig ha...	09.02.2010	

Below the main grid are four columns of additional content:

- GJØR MER:** Det finnes flere måter å engasjere seg på for Amnesty. Her finner du en oversikt over hva du kan gjøre hvis du vil gjøre mer enn å signere på aksjonene våre. Gjør mer på Amnesty-kontoret. Gjør mer på nett og i media. Gjør mer på skolen. Gjør mer på jobben. Gjør mer på gata.
- MØT AMNESTY-AKTIVISTER:** Includes a photo of Grete Røder Østby and text: "Navn: Grete Røder Østby, Alder: 65 år, Yrke: Pensjonist (tidligere lektor). Grete har vært aktiv i Amnesty i 25 år og er mest opptatt av de klassiske Amnesty-sakene der mennesker eller grupper blir urettmessig fengslet. Hvorfor er det viktig å være medlem i Amnesty? - Fordi det gir oss tilgang til saker å kjempe for. Det er mer "snøkk" i appeller fra mange, som sammen aksjonerer på gata, på nett og med brev over tid. Les hele intervjuet med Grete her."
- FØLG AMNESTY-BLOGGEN:** REVOLUSJON PÅ LITTERATURHUSET (25.02.2010), OM Å BLI FORBANNA! (23.02.2010), AFGHANISTAN: HVA ER PRISEN FOR DIALOG MED TALIBAN? (17.02.2010).
- ANNONSE:** YOU CAN'T CHANGE HISTORY (Image of a person on a ledge).

At the bottom, contact information is provided: info@amnesty.no, Ansvarlig redaktør: Kristin Rødland Buick, Webmaster: Nils Rahne, Kontaktnummer: 1644 11 69764, Organisasjonsnummer: 970 148 698, Copyright 1996-2008 Amnesty International Norge.

Figur 5.4: Trykker man på *Flere aksjoner* i undermenyen under *Aksjon*, får man opp en liste men flere aksjonssaker man kan velge mellom.

Går man inn på en spesifikk aksjon, kommer man til det vi kaller *aksjonsartikkelen* (se figur 5.5). Øverst, men under hoved- og undermenyen, er et illustrasjonsfoto for den aktuelle aksjonen. Til høyre for dette bildet har vi den samme signaturboksen som på *Aksjonsiden*. Under bildet har vi fire kolonner. Den venstre kolonnen inneholder overskrifter til andre artikler/nyheter. I kolonne nummer to finner vi overskriften til artikkelen med ingress og brødtekst. Nederst i den samme kolonnen finner vi også mulighet til å legge inn kommentarer. Den tredje kolonnen inneholder noen nyttige linker og bilder/videoer til artikkelen. I kolonnen lengst til høyre finner vi bilde og kontaktinformasjon til den som er ansvarlig for aksjonen, en feed med de nyeste blogginnleggene og en med Amnestys nyeste oppdateringer på Twitter, og til slutt noen linker til hvordan man kan engasjere seg videre i Amnesty.

**AMNESTY INTERNATIONAL**
BLOGG | SIDEKART | MIN SIDE | PRESSE

---

AKSJON
AKTUELT
STØTT OSS
OM AMNESTY
RESSURSER
SIKK

CNN RSS | AKSJON | FOKUS | FLERE AKSJONER | AKSJONSKALENDER | GJØR MER | HVORFOR AKSJONERE | KAMPANJER



**500 FORSVUNNET. AMNESTY KREVER SVAR**

Tirsdag 09.02.2010 / Skrevet av: [Kjartan Audevaldsen](#)

Masood Ahmed Janjua, en forretningsmann fra Rawalpindi, "forsvunnet" 30. juli 2005, på en bussreise til Peshawar. Han skal angivelig ha blitt tatt bort av pakistanske etterretningstjeneste. Han er en av mange hundre som er "forsvunnet" i Pakistan.

Masood Ahmed Janjua var på vei til Peshawar sammen med Faisal Faraz, en 25 år gammel ingeniør fra Lahore, 30. juli 2005. Undersøkt skal begge mennene ha blitt arrestert og ført bort fra bussen.

Observert av vitner, myndighetene kjenner til i oktober 2006 begynte den pakistanske høyesterett med haringene om Masood Ahmed Janjuas sak. Flere mennesker som var nær eller "forretnings" har bevitnet å ha sett begge de to mennene i fengsel.

Myndighetene har ved enkelte anledninger offisielt avkåret at de er forsvunnet, mens andre ganger har de bevitnet ethvert kjønneskap til saken.

Rettssaker av 511 "forsvunnet" var registrert og uløst i Pakistan for 2001. Men etter at Pakistan ble en alliert med USA i "krigen mot terror", har et stort antall mennesker "forsvunnet" landet.

Mange av dem har antakeligvis blitt overlevet til USA mot dusør og hænet i amerikanske fengsler eller hemmelige fengsler.

I tillegg har politiske opposenter av regjeringen blitt ofre for "forsvunnet", særlig medlemmer av nasjonalistiske grupper i regionene Sindh og Balochistan som begge utviser mot autoritet.

Stort marked Det er uløst hvor mange personer som har forsvunnet. Men følge den pakistanske organisasjonen "Defence of Human Rights", som jobber på vegne av de "forsvunnet", kan det være mer enn 500 personer.

I oktober 2006 begynte den pakistanske høyesterett å behandle saker som hadde blitt fremmet av familiedommere av de "forsvunnet".

I november 2007 var mer enn 100 mennesker funnet og sluppet fri. Men håpet til ofrenes familier forsvant i november 2007 da president Musharraf erklærte unntaksstilling.

Under unntaksstillingen avvarte Musharraf de fleste høyesterettsdommene, og siden da har ingen av de gjeldende "forsvunnet" saker blitt hørt i haringene om Masood Janjua og Faisal Faraz ble også stanset. Dette ødela håpet til de to menneskene familier.

My regjering ble handlingskraftig Masood Ahmed Janjua og Faisal Faraz sine skyttere er fortsatt uløst. Skytteren til de "forsvunnet" er derfor tatt forbeholdt til gjennomføringen av høyesterettsdommene.

Etter valget i februar 2008 kom en ny regjering til makten som hadde som en av sine løfter å gjennomføre høyesterettsdommene.

De fleste av dommene er nå blitt gjennomført, inkludert høyesterettsjustisarius Iftikhar Muhammad Chaudhry, men så langt har lite skjedd.

Oppdatering 02.02.2010 Den 19. november gjennomtok pakistansk høyesterett haringene i forsvunnet sakene.

Den første haringen var vedrørende sakene til Imran Munir og Mustafa Azam, begge hadde vært utsatt for brutale forsvunnet, men var nå kommet til rette. Ifølge Defence for Human Rights, er fremdeles 195 av totalt 416 forsvunnet saker ubehandlet i rettsystemet.

Den pakistanske regjeringen har også nedsatt en etterforskningsgruppe bestående av seks personer fra etterretningstjenesten, som skal forsøke å finne Masood Janjua.

0000

HAR DU BEMERKT FLERE AKSJONER

Navn:

Epost:

E-postadressen publiseres ikke. Vi registrerer den for å verifisere signaturen din.

SEND DIN SIGNATUR

LES APPELLEN HER

**ANSVARLIG FOR AKSJONEN**



Rådgiver, Gerard Folvord  
Har du spørsmål til denne aksjonen, ta kontakt med Gerard Folvord på epost: [gerard@amnesty.no](mailto:gerard@amnesty.no)

LES AMNESTY SAKER

**AFGHANISTAN: HVA ER PRISEN FOR DIALOG MED TALBÁN?**

17.02.2010  
Afghanske myndigheter har ondt av å se på...

**KUNST GANGER LYS**

11.02.2010  
Det er nå blitt tene til kunstutstillingen...

**NETTETIKETTENS ALLSME – HVEM ER 'JEG'?**

10.02.2010  
Vi jobber mye med Amnesty's tilstedeværelse i...

**TWITTER**

[Amnesty\\_Norge](#)  
Aljosha Bussemakke-Me-Me  
Måe datt til å ha frengst  
for å ha defint (defint)  
demonstrasjon - keev  
istandstet  
<http://bit.ly/s1SgD>

[Amnesty\\_Norge](#)  
Til har det vært vesentlige  
problemer med å vitet  
sammene i dag  
<http://bit.ly/c1vX1L>

[Amnesty\\_Norge](#)  
Nicargua Krivene kan de  
er blitt født han er  
grovst og mettes  
behandling  
<http://bit.ly/c1vX1L>

[Amnesty\\_Norge](#)  
Fellmas foed i krivene  
de brvst mott  
<http://blogg.amnesty.no/mendebidde>

[Amnesty\\_Norge](#)  
Folbusas foed i krivene  
de brvst mott  
<http://blogg.amnesty.no/mendebidde>

**GJØR MER**

Det finnes flere måter å engasjere seg på for Amnesty.

Har finner du en oversikt over hva du kan gjøre hvis du vil gjøre mer enn å signere på aksjonene våre.

Gjør mer på [Amnesty-kontoret](#)

Gjør mer på [nett og i media](#)

Gjør mer på skolen

Gjør mer på jobben

Gjør mer på gata

**ANDRE NYHETER**

Overgrep i flyktningleirer i Tsjad

Russiske aksjoner i det ferske

Overvaksende og risikabelt fredspris

Fangst mellom militært og regjering

Hindst tilgang til drukkern

Budsjett for 2010

blinde

Amnesty mntar Shellstasjonar i Norge

Stress Quantanomic-kritikk var tomme ord

Storborgst mntar krav om vidtdekkende

Aksjoner i Kina "tyddet av banen"

**ELDRE NYHETER**

>>

Nyttige lenker med stoff om Pakistan:

- Amnesty Internsjonal
- Human Rights Commission of Pakistan
- Human Rights Watch
- Wikipedia

Intervju av Masood Janjuas kone, Amina Masood



Amina Masood, Masood Janjuas kone



Zainab Khanoom, moren til Faisal Faraz

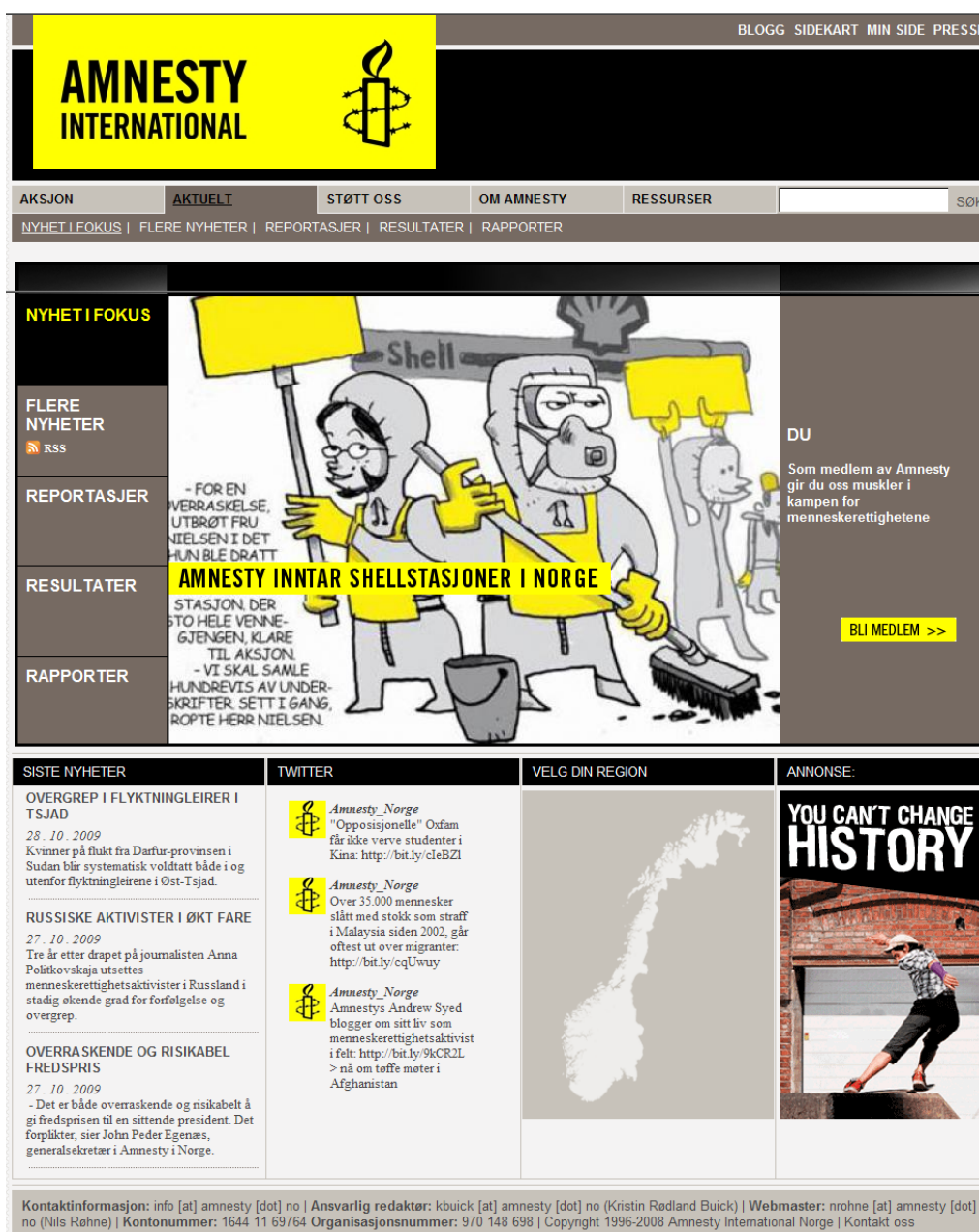
Kontaktinformasjon: [info@amnesty.no](mailto:info@amnesty.no) | Ansvarlig redaktør: Kristin Rodland Buck | Webmaster: Nils Rahne | Kontonummer: 1544 11 63764 | Organisasjonsnummer: 970 148 638 | Copyright 1996-2008 Amnesty International Norge | Kontakt oss

Figur 5.5: Figuren her viser artikkelen man kommer inn på hvis man går inn på en aksjon for å lese om den. Vi kaller denne siden for en *aksjonsartikkel*.

31

### 5.3.3 Aktuelt

Videre på hovedmenyen finnes *Aktuelt*, som inneholder nyheter, reportasjer og resultater som Amnesty publiserer. Dette kaller vi for *Aktueltssiden*. Den er bygget opp på samme måte som Aksjonssiden, hvor man kommer direkte til *Nyhet-i-Fokus*. Skjerm bilde vises i figur 5.6. Ellers kan man velge mellom *Flere nyheter*, *Reportasjer*, *Resultater* og *Rapporter* i undermenyen. Disse er bygget opp med en listefunksjon lik den man finner under *Flere aksjoner* (se figur 5.4).



Figur 5.6: Trykker man på *Aktuelt* i hovedmenyen kommer man til det vi kaller *Aktueltssiden*.

### 5.3.4 Støtt oss

Neste fane på hovedmenyen heter *Støtt oss*. Her finner man informasjon om hvordan man kan støtte Amnesty økonomisk, enten man er en privatperson eller en bedrift. Skjerm bilde vises i figur 5.7.

**AMNESTY INTERNATIONAL**

BLOGG SIDEKART MIN SIDE PRESSE

AKSJON AKTUELT **STØTT OSS** OM AMNESTY RESSURSER

PERSON | BEDRIFT

**PERSON**

**VI LOVER DEG**

- å være en kompromissløs forsvarer for alle menneskerettigheter
- å ansvarliggjøre overgriperne og presse på for endring

**BEDRIFT**

- å gi stemmen din kraft og la deg bruke engasjementet ditt på noe som virker
- å utfordre og presse Norge til å stå opp for menneskenes rettigheter

**DU**

Som medlem av Amnesty gir du oss muskler i kampen for menneskerettighetene

**BLI MEDLEM >>**

GI EN GAVE TIL AMNESTY	SMS-NETTVERKET	TESTAMENTARISKE GAVER	ANNONSE:
<p>En gave til Amnesty er en gave med mening.</p> <p>Du kan gi gaver selv eller på vegne av andre.</p> <p>Dersom bidraget til Amnestys arbeid er en gave fra deg til noen andre, vil du få tilsendt et kort i posten som du kan signere og selv gi til vedkommende.</p> <p>Beløpet bestemmer du selv.</p> <p>Gaven kan være et månedlig beløp, som trekkes automatisk fra kontoen din eller et engangsbeløp.</p>	<p>Redd liv og støtt Amnesty med få tastetrykk.</p> <p>SMS-aksjonene brukes når det er antallet underskifter og rask respons som gjelder. Aksjonene sendes ut to ganger i måneden.</p> <p>Det koster seks kroner å sende og motta SMS, altså 12 kroner i måneden. Halvparten av beløpet går til Amnestys arbeid.</p> <p>Bli med i nettverket av 70.000 SMS-aktivister ved å sende AMNESTY til 1960</p> <p>For å melde seg av send STOPP AMNESTY til 1960.</p>	<p>Det er mulig å gi en gave til Amnestys arbeid i testamentet. Amnesty er finatt for arveavgift og gaven går derfor uavkortet til vårt arbeid for menneskerettighetene.</p> <p>Vi anbefaler å kontakte fagfolk slik at kravene til gyldighet blir ivarettatt. De fleste benytter egen advokat til å opprette testamentet, men Amnesty International kan også hjelpe med juridisk bistand.</p> <p>Amnesty International er svært takknemlig for enhver gave. Det gir oss mulighet til å styrke arbeidet vårt for å redde liv.</p>	<p><b>YOU CAN'T CHANGE HISTORY</b></p>

Kontaktinformasjon: info [at] amnesty [dot] no | Ansvarlig redaktør: kbuick [at] amnesty [dot] no (Kristin Rødland Buick) | Webmaster: nrohne [at] amnesty [dot] no (Nils Röhne) | Kontonummer: 1644 11 69764 Organisasjonsnummer: 970 148 698 | Copyright 1996-2008 Amnesty International Norge | Kontakt oss

Figur 5.7: Her vises skjerm bildet man kommer til ved å trykke på *Støtt oss* i hovedmenyen.



### 5.3.5 Om Amnesty

Trykker man på fanen hvor det står *Om Amnesty* i hovedmenyen, kommer man til siden som vises i figur 5.8. Her kan man blant annet lese om hva Amnesty er og gjør.

The screenshot shows the Amnesty International website. At the top, there is a navigation bar with links for 'BLOGG', 'SIDEKART', 'MIN SIDE', and 'PRESSE'. Below this is a header with the Amnesty International logo and a candle icon. A secondary navigation bar contains 'AKSJON', 'AKTUELT', 'STØTT OSS', 'OM AMNESTY' (highlighted), and 'RESSURSER', along with a search box labeled 'SØK'. Below the navigation is a sub-menu with 'HVEM ER VI', 'HVORDAN JOBBER VI', 'HVA JOBBER VI MED', and 'AMNESTY DER DU BOR'. The main content area is divided into three columns. The left column is a sidebar menu with 'HVEM ER VI' selected. The middle column contains the main text about Amnesty International, including a section 'Vår innsats nytter' and 'Uavhengig'. The right column features a 'DU' section with a call to action 'BLI MEDLEM >>'. The footer contains contact information: 'Kontaktinformasjon: info [at] amnesty [dot] no | Ansvarlig redaktør: kbuick [at] amnesty [dot] no (Kristin Rødland Buick) | Webmaster: nrohne [at] amnesty [dot] no (Nils Röhne) | Kontonummer: 1644 11 69764 Organisasjonsnummer: 970 148 698 | Copyright 1996-2008 Amnesty International Norge | Kontakt oss'.

Figur 5.8: Trykker man på *Om Amnesty* i hovedmenyen kommer man til denne siden, hvor *Hvem er vi* er valgt som default i undermenyen.

### 5.3.6 Ressurser

Den siste fanen i hovedmenyen heter *Ressurser* og her kan man blant annet finne materialer til nedlasting, nyttige linker, og en søkeside med flere filtreringsalternativer (se figur 5.9).

The screenshot shows the Amnesty International website's search interface. The top navigation bar includes links for 'BLOGG', 'SIDEKART', 'MIN SIDE', and 'PRESSE'. The main menu highlights 'RESSURSER'. The search page features a search bar with a 'SØK' button and filters for 'Relevans' and 'Publiseringsdato'. The left sidebar lists categories: 'AMNESTYS RAPPORTER', 'AUDIOVISUELT ARKIV', 'MENNESKE-RETTIGHETS-DOKUMENTER', and 'UNDERVISNINGS-MATERIELL'. The central search area includes date range fields ('Fra dato', 'Til dato'), dropdown menus for 'LAND' (listing countries like Afghanistan, Albania, Bolivia, Bosnia-Herzegovina, Botswana, Brasil) and 'TEMA' (listing topics like Aktivisme, Angrep mot sivile, Åpen, Arrangement, Arrestasjon, Asylsøkere), and a 'KAMPANJER' section with a list of campaigns: 'Dignity: Test-kampanje 1', 'Dignity: Test-kampanje 2', 'Fattigdom handler ikke om penger', 'Shell: Rydd opp', 'Stopp dødsstraff', and 'Stopp vold mot kvinner'. At the bottom, there are checkboxes for 'Aksjon', 'Aktuelt', 'Ansatt', 'Block for rss feed', 'Kampanje', 'Kontaktinformasjon', 'Newsletter subscription block', 'Page', and 'Page with rss feeds'. A footer contains contact information: 'Kontaktinformasjon: info [at] amnesty [dot] no | Ansvarlig redaktør: kbuick [at] amnesty [dot] no (Kristin Rødland Buick) | Webmaster: nrøhne [at] amnesty [dot] no (Nils Røhne) | Kontonummer: 1644 11 69764 Organisasjonsnummer: 970 148 698 | Copyright 1996-2008 Amnesty International Norge | Kontakt oss'.

Figur 5.9: Går man inn på *Ressurser* i hovedmenyen kommer man til denne siden hvor det vises et skjema for enkelt eller avansert søk på siden.





# **Del II**

## **Gjennomføring**

---

## Gjennomføring av brukbarhetstesten

---

Dette kapitlet tar for seg hvordan jeg plala og gjennomførte brukbarhetstesten for finne svar på forskningsspørsmål 1. I 6.1 beskriver jeg de forberedelsene jeg gjorde før selve hovedtesten. Videre sier jeg litt om testdeltakerene og hvordan vi fikk tak i de i 6.2. Til slutt forteller jeg om hvilken testprosedyre vi fulgte og hvilke oppgaver vi brukte i henholdsvis 6.3 og 6.4.

### 6.1 Forberedelser

Som jeg har skrevet om i 2.2.2, må man lage noen oppgaver som testdeltakeren skal gå igjennom på en slik brukbarhetstest. For å finne noen gode oppgaver brukte jeg en del tid på å gjøre meg kjent med nettsiden for å se hvilken funksjonalitet den hadde, og jeg brukte også noen personas som Amnesty hadde sendt meg tidligere, for å se hva slags type mennesker som ville besøkt siden og hvorfor. I tillegg hadde jeg en dialog med Amnesty via epost for å høre om det var noen ting de var spesielt interessert i å få testet. Etter å ha laget et utkast til oppgaver sendte jeg de til Amnesty for å få tilbakemelding. Dette gjorde jeg cirka en uke før pilottesting skulle gjennomføres.

Når det gjelder testopplegg, så ville jeg følge samme mal som jeg brukte i prosjektoppgaven min [25], siden det hadde fungert nokså bra. I korte trekk var det samme opplegg som er beskrevet i 2.2.2. Hele testopplegget finner du nærmere beskrevet i 6.3.

#### 6.1.1 Testlokaler

For gjennomføring av testen var det naturlig at den skulle bli gjennomført på brukbarhet-laboratoriet ved Norsk Senter for Elektronisk Pasientjournal (NSEP) i Trondheim som har tilknytning til mitt institutt og har øyesporingsutstyr. Laboratoriet består av et testrom hvor selve testen blir gjennomført av testpersonen. Her sitter deltakeren framfor en pc med en tilkoblet øyesporer. Dette rommet har også flere kameraer som kan filme fra ulike vinkler og mikrofoner som tar opp lyd. I testrommet er det også flyttbare vegger, slik at man kan forandre omgivelsene alt etter hvilken setting man ønsker.

Rommet ved siden av blir kalt teknisk rom og her kan man observere det som skjer i testrommet via en storskjerm og høytalere. Her kontrolleres også opptak av video og lyd, og innstillingen av kameraene.



Figur 6.1: Bildet viser testrommet hvor testingen ble gjennomført. Øyesøkeren er plassert på bordet mellom skjermen og tastaturet.

### 6.1.2 Utstyr

Til testingen ble det brukt en Intel datamaskin med Windows XP Professional og en 19"-skjerm. Øyesporeren vi brukte var en Tobii X60, som er en frittstående øyesporer (se figur 6.2). Programvaren vi benyttet var Tobii Studio v.2.0.8.



Figur 6.2: Viser bilde av en Tobii Eye Tracker X60. Kilde: [www.tobii.com](http://www.tobii.com)

### 6.1.3 Pilottest

I forkant av en brukbarhetstest er det vanlig å gjøre en pilottest for å sjekke at testoppgavene og testopplegget er gjennomførbart. Vi kjørte en pilottest med én testperson uken før selve hovedtesten.

I første avsnitt i 6.1 skriver jeg at jeg sendte oppgavene til Amnesty for å få tilbakemelding på disse før pilottesten. På grunn av en datafeil fikk jeg dessverre ikke tilbakemeldingen før jeg skulle gjennomføre pilottesten. Jeg hadde selv gått igjennom utkastet en

gang til og endret på noen ting som ikke var gjennomførbart da nettsiden ikke kunne håndtere det ennå. Et eksempel var at de skulle gi en gave til Amnesty via nettsiden, denne funksjonaliteten var ikke mulig å teste ennå, så derfor tok jeg bort denne oppgaven. Oppgavene jeg brukte under pilottesten finner du i vedlegg B.

Testopplegget under pilottesten fungerte bra, så jeg fant ut at dette ikke trengtes å endres på. Når det gjelder oppgavene så oppdaget jeg at på oppgave 5, hvor man skulle finne en reportasje som interesserer en og legge inn en kommentar, at testpersonen fokuserte veldig på å finne en reportasje som virkelig interesserte ham, heller enn å bare finne en. Jeg spurte om dette i intervju-runden etterpå, for meningen med oppgaven var egentlig bare å teste det å legge inn en kommentar. Deltakeren svarte da at han hadde trodd det var det at han var *interessert* i reportasjen som var det viktige, ikke det å legge inn kommentar. Dette fikk meg til å tenke på om jeg burde omformulere oppgaven til hovedtesten, men jeg kom fram til at det var greit at deltakerne ikke bare valgte den første og beste reportasjen de fant, men kanskje lette litt rundt en stund slik at vi oppnådde en mer realistisk situasjon. Jeg ville heller presisere under selve testen at fokuset lå på å legge inn en kommentar, hvis jeg så at folk brukte veldig lang tid på å finne en interessant reportasje.

Bortsett fra denne oppgaven, fungerte resten av oppgavene greit nok. Men etter at jeg hadde gjennomført pilottesten fikk jeg endelig tilbakemeldingen fra Amnesty angående oppgavene. De hadde noen spesifikke ønsker som de ville ha testet, og jeg gjorde noen forandringer i oppgavene. På grunn av dette kjørte jeg ennå en pilottest dagen før selve hovedtesten, for å være sikker på at oppgavene var greit formulert. Dette gikk heldigvis bra, og vi var dermed klare for selve hovedtesten. De endelige testoppgaven er listet i 6.4.

## 6.2 Testpersoner

Til hovedtesten ville vi ha så mange som 15 deltakere, fordi vi her skulle gjøre en metode-sammenligning, og ikke bare en vanlig brukbarhetstest hvor man vanligvis bruker litt færre deltakere. Hvordan vi fikk tak i disse deltakerene er beskrevet under.

### 6.2.1 Rekruttering

For å skaffe testdeltakere tok vi kontakt med Amnesty for å høre om de hadde muligheter til rekruttere noen fra deres nettverk her i Trondheim, på den måten kunne vi få tak i folk som kjente til Amnesty fra før. Terje Røsand tok også kontakt med noen kjente som hadde hatt litt med Amnesty eller lignende organisasjoner å gjøre før. I tillegg hentet jeg inn noen venner av meg, som jeg visste ikke kjente så godt til Amnesty, slik at vi fikk noen på den siden også. Vi passet på at deltakerne var spredt utover flere aldersgrupper.

Vi hadde ikke mulighet til å dele ut noen belønning til de som deltok, så det måtte derfor betraktes som deres frivillige bidrag til arbeidet som Amnesty gjør, eller som en vennetjeneste for meg eller Terje.

### 6.2.2 Testdeltakere

I begynnelsen av testen samlet vi inn demografiske data om kjønn, alder, hvor ofte de bruker internett, hvor godt de kjenner til Amnesty (altså hva Amnesty er og gjør), og hvor ofte de hadde besøkt den gamle nettsiden til Amnesty.

På spørsmålene om hvor ofte de brukte internett og hvor ofte de hadde besøkt den gamle nettsiden kunne de velge et av følgende svar: *Hver dag, ofte, en gang i blant, sjelden* eller *aldri*. Når det gjelder deres kjennskap til Amnesty, kunne de velge blant alternativene *Veldig godt, godt, middels, lite* eller *ingenting*. Vi lot deltakerne selv få vurdere hva som ligger i de ulike alternativene og blandet oss ikke bort i det, siden det er vanskelig å definere en god skala for sånn som kjennskap til Amnesty. En oversikt over hvilke svar testpersonene gav, vises i tabell 6.1.

Deltaker #	Kjønn	Alder	Bruker internett	Kjennskap til Amnesty	Besøkt Amnestys gamle webside
1	Kvinne	40-49	Hver dag	Veldig godt	Aldri
2	Kvinne	20-29	Hver dag	Veldig godt	Ofte
3	Mann	30-39	Hver dag	Godt	Sjelden
4	Mann	10-19	Hver dag	Veldig godt	Hver dag
5	Mann	40-49	Hver dag	Godt	Aldri
6	Kvinne	30-39	Hver dag	Middels	Sjelden
7	Mann	20-29	Hver dag	Veldig godt	Ofte
8	Kvinne	20-29	Hver dag	Lite	Aldri
9	Kvinne	20-29	Hver dag	Godt	Sjelden
10	Kvinne	20-29	Hver dag	Middels	Aldri
11	Kvinne	20-29	Hver dag	Lite	Aldri
12	Kvinne	30-39	Hver dag	Middels	Aldri
13	Kvinne	30-39	Hver dag	Veldig godt	Hver dag
14	Kvinne	40-49	Hver dag	Middels	Aldri
15	Mann	30-39	Hver dag	Veldig godt	Sjelden

Tabell 6.1: Oversikt over testdeltakere.

### 6.3 Testprosedyre

Prosedyren vi fulgte under testen var som følger:

1. Ønske testdeltaker velkommen og presentere oss.
2. Forklare for testdeltakeren formålet med testen og hva vi skal gjøre.
3. Deltakeren skriver under på at han/hun vil delta.
4. La deltakeren fylle inn demografiske data om seg selv.
5. Kalibrering av øyesporeren.
6. Gjennomføring av testoppgavene.
7. Deltakeren svarer på SUS-spørreskjemaet.
8. Kort intervjurunde for å avslutte testen.
9. Takker testpersonen for deltakelsen i testen.

Under punkt 2 forklarte jeg for testdeltakeren at formålet med testen var å teste brukervennligheten til siden, og at vi skulle gjøre det ved at de gjennomførte 9 oppgaver, så skulle de svare på et kort spørreskjema og deretter hadde jeg noen spørsmål om hva de syntes om siden. Jeg viste også utstyret vi skulle bruke og forklarte at vi kom til å ta opp video og lyd, men at vi ikke skulle vise fram bilder eller navn av de når vi skulle presentere resultatet, slik at de ville være helt anonyme som deltakere. Jeg la også vekt på at dette var en test av nettsiden, *ikke* en test av hvor flinke de var. I tillegg forklarte jeg at jeg dessverre ikke kunne hjelpe dem så mye underveis, men at vi gikk videre til neste oppgave hvis de stod fast. Deltakeren ble også spurt om å tenke høyt underveis, og jeg brukte et eksempel for å forklare hva jeg mente med det. På forhånd hadde alle også fått utdelt et infoskriv hvor mye av den samme informasjonen stod, men jeg gjentok det her for at det ikke skulle være tvil om noe. De fikk også anledning til å komme med spørsmål hvis det var noe de lurte på.

På forhånd hadde jeg skrevet ut oppgavene på et ark, og jeg hadde også laget noen fiktive persondata som deltakerne skulle bruke dersom det var nødvendig å legge inn noe personinformasjon under testen (for eksempel til å signere på en aksjon). Oppgavene ville komme opp på skjermen en og en ettervert som de gikk igjennom de, men hvis de skulle glemme hva oppgaven gikk ut på, var greit å ha det på papir også.

Spørsmålene jeg stilte i intervjurunden til slutt finner du i vedlegg C.

## 6.4 Testoppgaver

Under viser jeg testoppgavene vi brukte under testen. Den første oppgaven er spesiell på den måten at den ikke krever at brukeren skal oppnå noe spesielt. Hensikten med oppgaven er å gjøre brukeren litt kjent med nettstedet før han begynner på de mer konkrete oppgavene, siden man ofte bruker litt tid på å orientere seg på et nytt sted.

1. Se deg litt rundt på sidene for å bli litt kjent med de.
2. Finn bloggen til Amnesty.
3. Signér aksjonen om forsvinningen av Masood Ahmed Janjua.
4. Hvor klikker du for å melde deg inn i Amnesty?
5. Finn en reportasje som kan interessere deg og legg inn en kommentar.
6. Finn ut hvilke 3 måter Amnesty vanligvis jobber på for å oppnå sine mål.
7. Finn ut hva pengene som Amnesty Norge samler inn går til.
8. Finn Amnestys logo til nedlasting.
9. Finn ut hvordan du kan være en aktivist for Amnesty gjennom bruk av sosiale medier.





# **Del III**

## **Resultater**

---

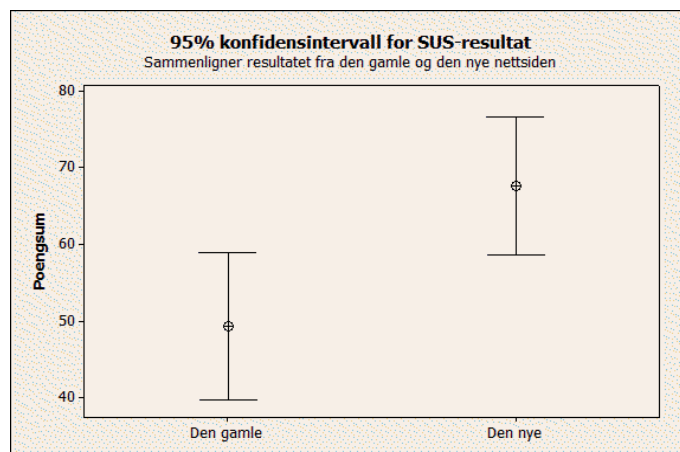
## Problemidentifisering

I dette kapitlet vil jeg presentere resultatene fra brukbarhetstesten. Først vil jeg vise hvilken bedømmelse nettsiden vi testet fikk etter SUS-skalaen (7.1.1) og også vise hvordan nettsiden scoret når det gjelder effectiveness (7.1.2) og da også hvilke områder som voldt problemer for brukeren. I 7.2 vil jeg gå over til å beskrive mer grundig de områdene hvor øyesporingen hadde noe å si for problemidentifisering, og dermed er med på å svare på det første forskningsspørsmålet vårt (se 1.2).

## 7.1 Resultater fra brukbarhetstesting

### 7.1.1 Satisfaction

Satisfaction-faktoren forteller noe om hvor fornøyde deltakerne er med brukbarheten på nettsiden vi har testet. Når vi bruker SUS til å måle dette, vil resultatet havne på en skala fra 0-100, hvor 100 best (se 2.3). En tabell som viser SUS-resultatet fra hver deltaker i denne brukbarhetstesten finnes i vedlegg D.



Figur 7.1: Figuren viser 95% konfidensintervaller for SUS-resultatet for den gamle nettsiden og den nye nettsiden. Den gamle siden har et 95% konfidensintervall med verdier [39.7, 59.0], mens den nye nettsiden har [58.7, 76.6]

Jeg gjorde en tilsvarende test av den gamle siden til Amnesty høsten 2009 og figur 7.1 viser et 95% konfidensintervall av resultatet fra begge testene. Ved å kjøre en to-utvalgs t-test med nullhypotese for like gjennomsnittsverdier i de to utvalgene (se kapittel 2.3.3),

fikk jeg en T-verdi på 3,03 og en p-verdi på 0,006. P-verdien er altså mindre enn 0,05 og vi kan derfor slå fast med nokså stor sikkerhet at folk synes den nye nettsiden er bedre enn den gamle.

Hvis vi skal prøve å koble poengsummen vi fikk fra det samlede SUS-resultatet til et adjektiv, som det gjøres i [2], og vi går ut fra gjennomsnittsverdien som på den gamle siden var 49,3, ville det bli karakterisert som OK (se delkapittel 2.3.2). Til sammenligning ser vi at den nye siden fikk et gjennomsnitt på 67,7 og blir da koblet til adjektivet God. Det er imidlertid to nivå til før man når helt opp, så vi ser at selv folk er mer fornøyd med den nye i forhold til den gamle, så har siden fremdeles forbedringspotensiale. Vi må også huske på at dette ikke er en vitenskapelig målestokk, men det kan gjøre det enklere for utenforstående å skjønne hva det betyr.

### 7.1.2 Effectiveness

Effectiveness handler om at brukeren får gjort det han har tenkt å gjøre, at han oppnår bestemte mål (som jeg har skrevet om i 2.1.1). At effectiveness-faktoren ikke er 100% er en klar indikasjon på at brukbarheten av produktet ikke er helt på topp. I tabell 7.1 har jeg samlet resultatet per testdeltaker.

Testperson	Oppgaver gjennomført riktig uten hjelp	Prosentandel
1	6	75%
2	7	88%
3	8	100%
4	6	75%
5	5	63%
6	5	63%
7	8	100%
8	8	100%
9	8	100%
10	8	100%
11	6	75%
12	6	75%
13	8	100%
14	8	100%
15	8	100%
Gjennomsnitt	7	88%
Standardavvik	1.2	14.9%
Min	5	63%
Maks	8	100%

Tabell 7.1: Tabellen viser hvor mange av oppgavene i brukbarhetstesten hver deltaker klarte uten hjelp og hvor stor prosentandel det utgjør. Dette kalles effectiveness-faktoren på nettsiden (se delkapittel 2.1.1). Gjennomsnitt, standardavvik, minimumsverdi og maksimumsverdi er også inkludert, men oppgave 1 er ikke medregnet fordi det var en åpen oppgave uten et skikkelig mål.

Som vi ser av tabellen er det ikke alle brukerne som får til alt, gjennomsnittet for hele utvalget av testpersoner ligger på 88%. Vi ser også at det er 7 av 15 deltakere har problemer med minst en av oppgaven, altså nesten halvparten av utvalget. Vi bruker så denne informasjonen til å se nærmere på *hvilke* oppgaver som var problematiske. Dette har jeg samlet i tabell 7.2. Ut fra denne ser vi at det er opplagt at oppgave 3, 7, 8 og 9 er problemoppgaver (se oppgavene i delkapittel 6.4).

Oppgave	Deltakere som fullførte korrekt
2	100%
3	60%
4	100%
5	100%
6	100%
7	67%
8	87%
9	87%

Tabell 7.2: Tabellen viser hvor stor andel av testdeltakerne som fullførte de ulike oppgavene korrekt. Vi ser at problemoppgavene var oppgave 3, 7, 8 og 9. Oppgave 1 er ikke inkludert her heller siden det ikke var en oppgave man skulle fullføre noe.

Det er også viktig å huske på at selv om tabell 7.2 gir klare indikasjoner på hvor vi har brukbarhetsproblemer, så betyr ikke det at det er de *eneste* stedene vi har det. Selv om brukerne klarte å fullføre oppgaven, kan det være de gjorde noen småfeil underveis. Dette må vi også ta hensyn til når vi skal analysere brukbarheten av siden. Gjennom høyttenkninsteknikken under testen og intervjurunden etterpå fikk vi også mer innsikt i hvilke ting deltakerne ble forvirret av, misforstod eller syntes var dårlig.

Ved å gå igjennom videoopptakene og intervjuene identifiserte vi problemer og hvor mange som hadde det samme problemet. Deretter kategoriserte vi disse problemene etter hvilken viktighetsgrad problemet hadde. Viktighetsgraden kunne være på tre nivåer: *Kritisk*, *moderat* eller *kosmetisk*. Vi har listet alle problemene vi fant i tabell 7.3 og viser hvor mange som hadde problemer med dette og hvilken viktighetsgrad vi gav det.

Problem	Antall deltakere	Viktighetsgrad
Signerer i kommentarfeltet	7	Kritisk
Audiovisuelt arkiv misforstås	9	Moderat
Kobling mellom bildet og Aksjon-i-Fokus misforstås	3	Moderat
Usikkerhet rundt signering av aksjon, men gjør det rett til slutt	4	Moderat
Linker er ikke tydelige som linker	5	Moderat
Linker har ulike utseende forskjellige steder på siden	3	Moderat
Søkesiden har lite forklaring og avansert søk fungerer ikke skikkelig	4	Moderat

Tabell 7.3 – fortsetter på neste side

Tabell 7.3 – fortsettelse fra forrige side

<b>Problem</b>	<b>Antall deltakere</b>	<b>Viktighetsgrad</b>
Dårlig lesbarhet når mus holdes over elementer i Reportasje-listen og andre lister	2	Kosmetisk
Forventer å finne informasjon om hvordan man kan være aktiv medlem under Støtt oss	1	Kosmetisk
Ønsker eller forsøker å klikke på overskriften til blogg-feeden på forsiden for å komme til bloggen	7	Kosmetisk
Gjør mer-begrepet i undermenyen på Aksjon er ikke helt selvforklarende	1	Kosmetisk
Leter etter hva pengene brukes til under Støtt oss	8	Kosmetisk
Lightbox-begrepet forstås ikke	8	Kosmetisk
Plassering og utseende på POST-knappen var litt rar	7	Kosmetisk
For smal kolonne for artikkeltekst	2	Kosmetisk
Søkelinjen har store bokstaver og fargen på disse på er lite synlig	3	Kosmetisk
Tekststørrelse	2	Kosmetisk

Tabell 7.3: Tabellen viser hvilke problemer vi har identifisert ved å se på videoopptak og se på intervjuene vi hadde med deltakerne. Vi viser også hvor mange av deltakerne som hadde dette problemet, og hvilken viktighetsgrad vi har gitt problemet

## 7.2 Problemområder

Her vil jeg beskrive nærmere de problemene vi fant hvor øyesporingen hadde noe å si for resultatet.

### 7.2.1 Signering av aksjon

I oppgave 3 i testen skulle deltakerne signere under på en aksjon som omhandlet forsvinningen av Masood Ahmed Janjua. Som vi ser av tabell 7.2 var det bare 60% som klarte å gjennomføre denne oppgaven, det var den som hadde dårligst resultat av alle oppgavene. Dette problemet var det eneste vi karakteriserte som kristisk når det gjelder viktighetsgrad. Underskriftskampanjer er en av de viktigste aksjonsformene Amnesty arbeider med, derfor er det veldig ille når de kanskje ikke får like mange underskrifter som de kunne hatt, hvis folk hadde forstått hvordan de skulle skrive under.

Det som gjør denne oppgaven litt interessant er at signaturfeltet opptar ganske stor plass på siden, i et felt til høyre for bildet som illustrerer aksjonen, se figur 5.5 i kapittel 5. Signaturboksen består av en teller som viser hvor mange som har skrevet under denne aksjonen, og to input-felter hvor man skriver inn navnet sitt og epost-adressen sin, i tillegg til en gul knapp hvor det står *Send din signatur*.

Man skulle tro at dette var ganske tydelig og lett å bruke, men det viste seg at det ikke var helt tilfelle. Det som skjedde var at flere av deltakerne signerte i det som faktisk er *kommentarfeltet* under artikkelen. Noen av disse var ganske raske med å signere under her, og var tydelig sikre på at de hadde gjort det riktig. Andre var veldig i tvil om de gjorde det rett. Et eksempel er ei som brukte hele fire minutter fra hun fant riktig side og til hun sendte inn signaturen (i kommentarfeltet). *“Jeg forstår ikke helt dette, jeg er jo vant til å sende inn signaturen min etter at jeg har lest artikkelen, men dette virker jo mer som et kommentarfelt”*, skrev hun til slutt i kommentarfeltet. Hun var altså ganske sikker på at det var en kommentar hun sendte inn, men kunne ikke skjønne hvor signaturfeltet var. Det var også flere som nevnte at signaturfeltet vanligvis brukte å være *under* artikkelen, fordi det var naturlig å lese om aksjonen *først*, og så skrive under. Dette var de også vant med fra de gamle sidene til Amnesty.

### Hva bidro øyesporingen til?

Da jeg så på øyesporingsvideoen til noen av de som ikke fikk til denne oppgaven, oppdaget jeg at testpersonene ikke så noe særlig på den øvre delen av skjermen. Med en gang siden lastet lå øynene på det store skjermbildet, men de gikk så raskt ned til der hvor overskriften og ingressen til artikkelen lå.

Hvis vi ser på blikkdiagrammet til hun som brukte hele fire minutter på å skrive under (se figur 7.2), så ser vi at hun er veldig rask til å gå ned til overskriften på artikkelen, og vi kan se at hun leser ingressen ganske nøye. Deretter ser hun litt på de tre neste avsnittene før hun ser på alle tre bildene som ligger til høyre for teksten. Etter dette går hun enda lenger ned på siden og oppdager input-feltene for navn og epost-adresse for de som skal legge inn kommentar. For å ikke få et altfor rotete blikkdiagram har jeg ikke tatt med hvor hun så videre, men hun gikk opp og ned flere ganger på det hun allerede hadde sett på. Noe som var tydelig, var at hun aldri interesserte seg for det som lå lenger oppe på siden enn overskriften til artikkelen.



Figur 7.2: Blikkdiagram som viser aksjonsartikkelen hvor en bruker ikke la merke til signaturboksen øverst til høyre for det store illustrasjonsfotoet.

Det samme ser vi på neste blikkdiagram (figur 7.3). Dette er en av de andre deltakerne, som heller ikke fikk til oppgaven. Vi ser at fiksering nummer 1 er cirka midt på det store illustrasjonsfotoet, men neste fiksering er på overskriften, under bildet. Det er altså veldig tydelig at det er teksten og overskriften som fanger interessen hennes, det er der hun finner informasjon hun synes er viktig. Vi ser at hun også leser ingressen ganske nøye, men så gjør blikket litt større hopp lenger nedover siden. Hun har tydeligvis funnet ut at hun er



The screenshot shows the Amnesty International website with a gaze diagram overlaid. The diagram highlights the user's visual path, starting from the top navigation bar, moving to the main headline 'Uncle President Please Find my Loving', and then down to the 'SEND DIN SIGNATUR' button. Other fixations are visible on various text blocks and images throughout the page.

Figur 7.3: Blikkdiagram fra en annen bruker, men samme oppgave som i figur 7.2. Fiksering nummer 1 er midt på det store illustrasjonsfotoet, men resten er nedenfor dette.

på rett aksjon, nå må hun finne ut hvor hun skal signere. Etter at hun har sett såvidt på et av bildene til høyre for teksten, oppdager hun et input-felt til venstre som hun undersøker. Hun finner imidlertid fort ut at dette gjelder noe annet. Hun går derfor helt ned på siden og oppdager input-feltene for navn og epost-adresse for å legge inn kommentar. Fra da av kretser blikket hennes kun rundt dette området og hun signerer ganske raskt.

Det som er tydelig i begge disse eksemplene er at de ikke oppfatter den øvre delen av siden som et *relevant* sted å lete etter et signaturfelt.

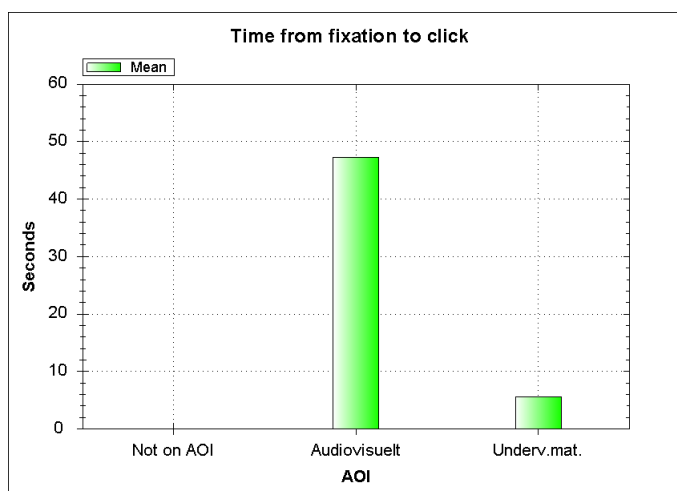
## 7.2.2 Fremmedord

Ser vi på oversikten over identifiserte problemer i tabell 7.3, ser vi at hele 9 stykker hadde problemer med å forstå *Audiovisuelt arkiv*. Dette var en problemstilling de kom bort i da de skulle gjøre oppgave 8 (se kapittel 6.4), hvor de skulle finne Amnestys logo til nedlasting. For å gjøre dette, måtte de gå til *Ressurser*, deretter *Audiovisuelt arkiv*, og så dukket der opp en undermeny hvor de kunne velge *Logo og profil* (se sidekartet i figur 5.1). Tabell 7.2 viser at 87% klarte å gjennomføre denne oppgaven riktig, altså var det bare to stykker som ikke klarte å gjennomføre oppgaven i det hele tatt. Grunnen til at så mange klarte å gjennomføre oppgaven, var at de brukte søkemotoren til å søke på *Logo*. Å bruke søkemotoren er jo en helt gyldig måte å finne fram på, men de fleste brukte denne *etter* at de hadde prøvd å finne logoen til nedlasting ved å navigere litt rundt først, altså som en siste utvei.

Det som var litt interessant var at de fleste gikk inn på *Ressurser*-siden ganske raskt på denne oppgaven, men så stoppet de opp og visste ikke helt hvor de skulle gå videre. En del prøvde å gå inn på menyvalget hvor det stod *Undervisningsmateriell* (se figur 5.9 i kapittel 5 for et skjermbilde). Dette var en side som ikke var utviklet ennå slik at innholdet var tomt, og dermed gikk de raskt ut derfra igjen.

Siden vi brukte høytttenkningsteknikken kom det fram at de forsto ordet *Audiovisuelt* som noe som hadde med lyd eller video å gjøre. “*Audiovisuelt, er ikke det lyd?*” var et typisk sitat fra deltakerne. En av de som fant nedlastingssiden via søkemotoren, og da oppdaget at den lå under *Audiovisuelt arkiv*, uttalte at “*Jo, en logo er jo forsåvidt visuell*”, altså var han enig i at ordet passet til innholdet, men han koblet ikke dette før *etter* at han hadde funnet ut hvor det lå.

I intervjurunden etter at testen var ferdig spurte vi deltakerne om hvorfor de hadde problemer med å finne logoen under *Audiovisuelt arkiv*. De fleste sa at de trodde audiovisuelt betydde video først og fremst. En annen uttalte at “*Audiovisuelt arkiv er litt for fancy ord for meg*”. Det virket altså som om problemet lå i at *Audiovisuelt arkiv* er et fremmedord som folk generelt ikke forstår fullt ut.



Figur 7.4: Figuren viser at det gjennomsnittlig tok cirka 47 sekunder fra en testperson fikserte på menyvalget *Audiovisuelt arkiv* til de klikket på det. Til sammenligning ser vi at det tok bare 6 sekunder fra de oppdaget menyvalget *Undervisningsmateriell* til de klikket på det.

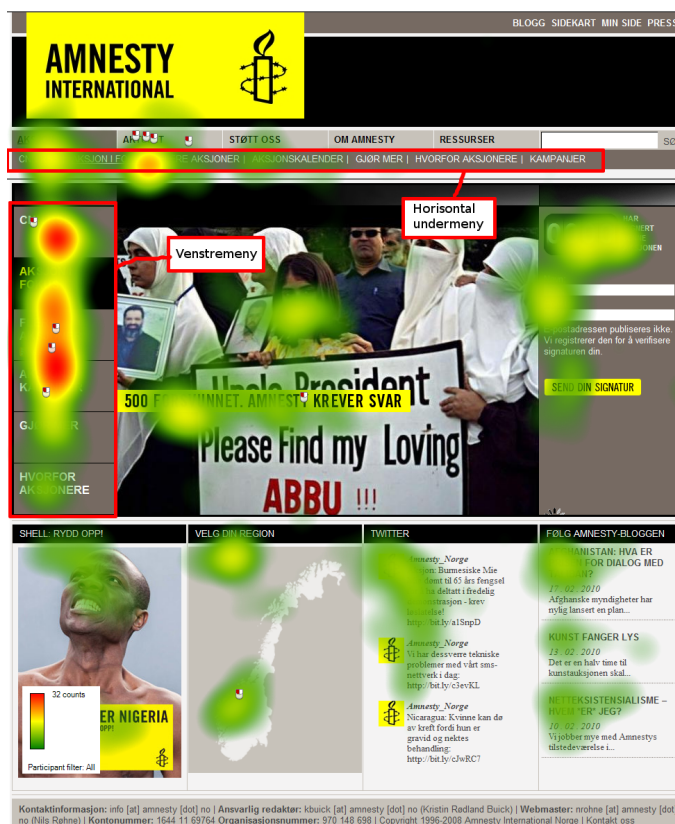
## Hva bidro øyesporingen til?

Jeg markerte menyvalgene for både *Audiovisuelt arkiv* og for *Undervisningsmaterie* som et interesseområde (se 3.4.5) for å kunne se om det var noe forskjell på hvordan folk så på disse to valgene. Jeg valgte å se på hvor lang tid det gikk fra de først oppdaget menyvalget til de klikket på det. Resultatet vises i figur 7.4.

Det som var interessant her var at folk brukte gjennomsnittlig 47 sekunder fra de først oppdaget ordet *Audiovisuelt arkiv* til de valgte å klikke på det, mens for *Undervisningsmaterie* tok det bare 6 sekunder. Dette viser tydelig at folk ikke engang vurderer *Audiovisuelt arkiv* som en mulighet. Dermed forsterker dette konklusjonen vi gjorde på bakgrunn av det tradisjonelle testresultatet.

### 7.2.3 Navigasjon

Mens jeg satt og observerte testdeltakerne under brukbarhetstesten, tenkte jeg over at veldig få brukte den horisontale undermenyen, til tross for at den lå rett under hovedmenyen (se kapittel 5.3), som de brukte flittig. I stedet brukte de menyen til venstre når de skulle navigere utover globalnavigasjonen. Dette var forsåvidt ikke et *problem* i seg selv, men som Jacob Nielsen anbefaler i [14], bør man strebe etter å ha et minimalistisk design, og da er dobbelt opp med menyer unødvendig. Det var også flere som nevnte at de syntes det var unødvendig med to like undermenyer. Derfor var dette noe jeg ville sjekke ut nærmere.



Figur 7.5: Varmekart for hvor folk fokuserte mest de 30 første sekundene på aksjonssiden.

### Hva bidro øyesporingen til?

Figur 7.5 viser et varmekart for hvor testdeltakerne fokuserte mest de 30 første sekundene på *Aksjonssiden* mens de holdt på med oppgave 1, hvor de skulle gjøre seg litt kjent med nettsiden. Vi ser at venstremenyen får mye mer oppmerksomhet enn den horisontale venstremenyen, selv om de i prinsippet inneholder de samme elementene. Av museklikkene (det hvite og røde symbolet) skjønner vi også at de fikseringene som ser ut til å være på den horisontale undermenyen, nok egentlig hører til på hovedmenyen, at det her er en målefeil (mer om unøyaktige målinger i 10.3).

Ut fra det resultatet vi har fått her, så virker det unødvendig å ha to like undermenyer, noe folk også kommenterte på. Vi kan også konkludere med at venstremenyen hadde bedre brukbarhet enn den horisontale undermenyen, siden folk fokuserte mest på den. Altså, hvis en av menyen bør tas bort, bør man ta bort den horisontale undermenyen.

### 7.2.4 Hva pengene som samles inn går til

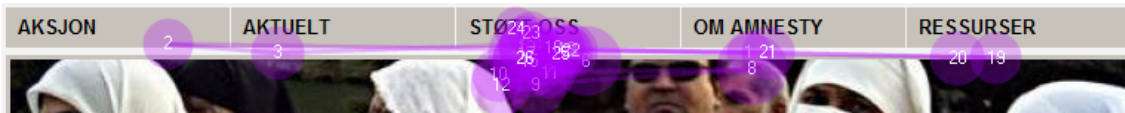
Som vi ser av tabell 7.3 var det hele 8 deltakere som lette etter hva pengene går til under *Støtt oss*-fanen da de skulle gjøre oppgave 7 (se oppgaveteksten i 6.4). Det var også 5 stykker som ikke klarte denne oppgaven i det hele tatt (tabell 7.2). Den korrekte måten å gjennomføre oppgaven på var å gå inn på *Om Amnesty*, og deretter *Hvordan jobber vi*, så var der en link hvor det stod *Hva går pengene til?* under teksten (se sidekart i figur 5.1).

Det er to ting som er interessant å dra fram når det gjelder denne oppgaven. For det første så hadde brukerne allerede besøkt *Hvordan jobber vi* i oppgave 6 (altså like før), og som vi ser av tabell 7.2 så klarte alle den oppgaven. Så hvorfor husket de ikke at det var der de skulle lete? En teori jeg har er at linken de skulle finne ikke var tydelig nok som en link, siden den ikke var understreket slik som linker vanligvis er. Derfor oppdaget de ikke denne da de besøkte siden første gang.

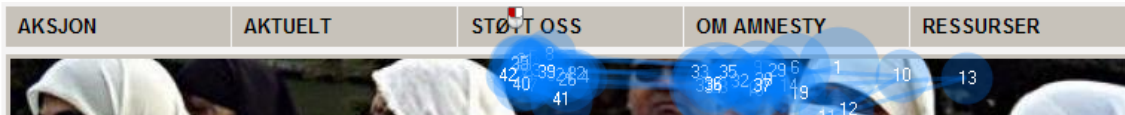
Det andre jeg vil trekke fram er at hva pengene brukes til logisk sett også hører til under *Støtt oss*, siden mange besøkte denne da de prøvde å løse oppgaven. En tendens jeg oppdaget da jeg observerte deltakerne, var at en del først gikk inn på *Om Amnesty*, men så ble de stående litt der og visste ikke helt hvor de skulle lete videre. Det passet kanskje ikke under noen av valgene i undermenyen. Da ombestemte de seg og forsøkte under *Støtt oss* i stedet (hvor de heller ikke fant noe).

### Hva bidro øyesporingen til?

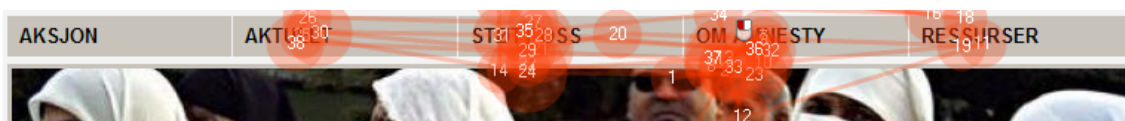
Jeg har valgt å vise fram utsnitt av blikkdiagram fra fire av testdeltakerne (figur 7.6-7.9) som viser hvor de så på hovedmenyen på forsiden da de skulle finne ut hva pengene går til. På figur 7.6 ser vi at deltakeren har vært innom alle fanene i hovedmenyen, men har sirklet mest over *Støtt oss*. Neste figur (7.7) viser en deltaker som har gått mange ganger mellom *Støtt oss* og *Om Amnesty*, og er faktisk opp i hele 42 fikseringer før hun endelig bestemmer seg for å gå inn på *Støtt oss*, noe som er ganske mye. I figur 7.8 er det likedan, men denne deltakeren bestemmer seg til slutt for å gå inn på *Om Amnesty*. Han har også sett på de to menyvalgene til høyre og venstre, men vi ser tydelig hvilke to han vurderer mest. Den siste figuren (7.9) har relativt få fikseringer og brukeren velger raskt riktig fane, men vi ser at hun *har* vurdert *Støtt oss* også, mens ingen av de andre valgene ble vurdert i det hele tatt.



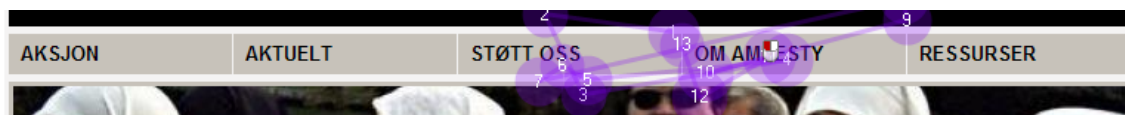
Figur 7.6: Blikkdiagram A. Utsnitt fra forsiden som viser hovedmenyen og hvor deltakeren fokuserte da hun skulle gjøre oppgave 7.



Figur 7.7: Blikkdiagram B



Figur 7.8: Blikkdiagram C



Figur 7.9: Blikkdiagram D

Alle disse figurene viser at i brukernes mentale modell av nettstedet så er både *Støtt oss* og *Om Amnesty* logiske plasseringer for artikkelen om hva pengene går til. Dette betyr ikke at man skal ta bort linken til artikkelen fra *Om Amnesty*, men kanskje gjøre det mulig å finne den via *Støtt oss*.



## Intervju med NetLife Research

---

Dette kapitlet vil ta for seg intervjuet jeg gjorde med Thor Fredrik Eie fra NetLife Research om hvordan de bruker øyesporeren de har. Intervjuet var et forsøk på å få svar på forskningsspørsmål nummer 2: *Hvilken nytteverdi ser utviklere som til daglig bruker øyesporing i brukbarhetstesting?* (se kap. 1.2). Intervjuet fant sted i NetLife Research sine lokaler i Oslo. Jeg brukte lydopptak da jeg gjennomførte intervjuet og det hele varte omtrent en halv time. Transkripsjonen av hele intervjuet finnes i vedlegg E.

### 8.1 NetLife Research

NetLife Research er et norsk konsulentfirma som har spesialisert seg på å levere tjenester innenfor brukeropplevelser og brukervennlighet. De er de eneste i Norge som selger brukbarhetstester med øyesporing. De bruker ikke øyesporing på alle tester de gjennomfører, men selger det som et tillegg til en tradisjonell brukbarhetstest, da med en prisøkning på ca 5%.

#### 8.1.1 Noen kunden vil ha

NetLife Research bruker for øvrig øyesporing på cirka 70-80% av alle testene de gjennomfører. De har foreløpig bare øyesporer på en av to testlaboratorier, men vurderer for tiden å kjøpe en til for å ha på det andre rommet. De opplever at mange kunder gjerne spør konkret etter øyesporing, enten fordi de har hørt eller lest om det på forhånd, eller at de har sett det demonstrert av NetLife Research tidligere.

#### 8.1.2 Hvordan gjennomfører de en test med øyesporing?

Når de gjennomfører en brukbarhetstest med øyesporing, gjør de det vanligvis som en todelt prosess. I første delen lar de testpersonen sitte foran skjermen og gjøre noen oppgaver på egen hånd, helt uten innblanding fra testteamet. Etter dette gjør de en mer vanlig brukbarhetstest hvor de benytter høytttenkningsteknikken (think-aloud).

I forhold til en vanlig brukbarhetstest bruker de cirka et dagsverk ekstra når de benytter øyesporing. Da er alt medregnet, både testing, analysering og uthenting av data.

## 8.2 Øyesporing som analyseringsverktøy

Her forteller jeg litt om hvilke nytteverdier de ser når det gjelder å identifisere problemer.

### 8.2.1 Understøtting av konklusjoner

I forhold til å trekke konklusjoner direkte fra øyesporingen sier NetLife Research at de er veldig forsiktige. Siden de bare bruker opp til 8 personer har de ikke statistisk støtte for å trekke konklusjoner kun basert på øyesporingsdata. Derimot bruker de ofte resultatet fra øyesporingen til å støtte opp under konklusjoner de kom fram til basert på resultatet fra høyttenkningssdelen.

### 8.2.2 Nye oppdagelser

Det er ikke så ofte de oppdager helt nye ting basert på øyesporingen. Men av og til dukker det opp noe som de har kalt for *Visuelle vollgraver*. Dette er store bilder eller grafiske elementer som gjør at andre ting “forsvinner”. Et eksempel er Telenor sine gamle nettsider, hvor de hadde noen store, fine bilder på siden. Globalnavigasjonen lå over dette, mens resten av innholdet lå under bildet. Det som skjedde var da at testpersonene kun konsentrerte seg om innholdet som var under bildet, og ikke det som var over. Dette var ting som ikke var kommet fram på annen måte enn gjennom øyesporingen.

### 8.2.3 Kombinasjonen med testresultater

Thor Fredrik Eie er skeptisk til hvor gode testresultater man vil få utfra rene øyesporingstester når man skal gjøre brukbarhetstester. Han tror det er i kombinasjonen med de funnene man gjør basert på tradisjonell testmetodikk at den viktige informasjonen ligger. I brukbarhetstesting er man ute den kognitive oppfattelsen til testpersonene, altså hvordan de forstår *bruken* av produktet man tester. Hadde man kun vært interessert i hvor mange som ser logoen på en sjampoflaske, slik som et reklamebyrå kunne vært, hadde øyesporingen vært nok. Det blir et helt annet bruksfelt for øyesporing. Men innenfor brukbarhetstesting trenger man litt mer informasjon enn bare øyesporing.

## 8.3 Presentasjon av resultater

Øyesporing har også verdi i dialogen de har med kunder, som jeg trekker fram under.

### 8.3.1 Slagkraftig kommunikasjonsverktøy

Noe av NetLife Research bruker øyesporing mest til er i dialogen med kunden. De opplever at når kunden får se med egne øyne hvor brukeren ser på skjermen og samtidig høre hva brukeren sier, så har de et veldig slagkraftig verktøy i forhold til å overbevise kunden om at de funnene de har gjort er riktige.



### 8.3.2 Varmekart og blikkdiagram

NetLife Research er veldig forsiktig med å bruke varmekart når de skal presentere resultatet. Det er veldig lett å hente ut et varmekart som ser veldig bra ut, men som egentlig ikke sier noen nyttig om brukbarheten på siden. Statistisk sett er også 8 personer altfor lite til å si noe gyldig ved hjelp av et varmekart.

I stedet bruker de heller blikkdiagrammer, som er gyldige individnivå. De bruker det til for eksempel å vise hvor testpersonen ser først på en nettside, altså hva som fanger interessen først.

## 8.4 Som hjelpeverktøy under selve testen

En ting som Thor Fredrik Eie nevnte fra sine egne erfaringer i bruk av øyesporing, var at det gjorde det lettere for ham som testleder å tilpasse spørsmål rettet mot testpersonen underveis i testen. I vanlige tester kan man gjerne bomme litt med timingen når man skal stille oppfølgingsspørsmål til brukeren, og så har kanskje ikke brukeren oppdaget det elementet man spør etter en gang. Når testlederen kan sitte å se hvor testdeltakeren fokuserer til enhver tid, kan han enklere vite når det passer å stille spørsmålet, for eksempel etter at brukeren har sett elementet man spør etter.

## 8.5 Ulemper

NetLife Research ser også noen ulemper med bruken av øyesporing:

### 8.5.1 Testpersoner

Noe man må ta hensyn til når man gjør tester med øyesporing, er at ikke alle personer egner seg til øyesporing. Folk som har briller er ofte vanskelige å teste fordi øyesporeren ikke klarer å lese øynene riktig.

Andre kan sitte ganske urolig, noe som gjør at de kan falle utenfor området hvor øyesporeren kan lese øynene. Da må man eventuelt fortelle brukeren hvordan han må sitte, men det er ikke alle brukere som liker så godt å bli rettet på. Så det blir en avveinings sak hvor mye man skal pirke på testdeltakeren.

### 8.5.2 Tekniske utfordringer

Selv om teknikken har kommet ganske langt, hender det av og til at programmet eller øyesporeren henger seg opp midt i testprosessen, og selv om testpersonen ikke merker så mye til det, betyr det jo at hele resultatet fra den testen går tapt, for man kan ikke ta testen på nytt med samme person.

Nettopp på grunn av disse tekniske detaljene som kan skjære seg, er det ikke alle i NetLife Research som er like komfortable med å gjennomføre slike tester med øyesporing. Det kan jo hende at det sitter fullt av kunder i rommet ved siden av som har betalt for å få se på, og så faller alt i fisk i stedet. Da blir det en psykisk påkjenning for de som utfører testen.



### Presentasjon for Amnesty

---

Dette kapitlet vil ta for seg erfaringene jeg gjorde meg angående nytteverdien av øyesporing i dialogen med kunder og utviklere, som i mitt tilfelle var de som jobber med utviklingen av nettsiden til Amnesty. Jeg hadde en presentasjon for de som jobber med utviklingen av nettsidene til Amnesty, hvor jeg viste fram de funnene vi hadde gjort under brukbarhetstesting og forklarte noen av problemene ved hjelp av øyesporingsdata. Dette møtet ville for min del være en mulighet til å få svar på forskningsspørsmål 3: *Hva er nytteverdien av øyesporing som del av systemutvikling i dialogen med kunder og utviklere?* (se kap. 1.2).

#### 9.1 Setting og gjennomføring av møtet

Møtet var avtalt å finne sted i Amnestys lokaler i Oslo, og ble gjennomført mandag 19. april 2010. Presentasjonen ble holdt i et møterom med fem personer tilstede foruten meg selv. Disse var: Ina Tin (kommunikasjonssjef), Nils Røhne (webmaster), Kristian Hvesser (organisasjonskonsulent), Siri Nielsen (praktikant) og Kristin Rødland Buick (nettredaktør).

Hensikten med møtet var at jeg skulle presentere resultatet fra brukbarhetstesting for Amnesty og de som utvikler siden. Samtidig ville jeg benytte anledningen til å gjøre observasjoner av hvordan øyesporingsdataene bidro til dialogen rundt problemene jeg presenterte. Jeg hadde også fått tillatelse til å gjøre videoopptak av møtet, og dette ble gjort med ett videokamera plassert et stykke bak i rommet.

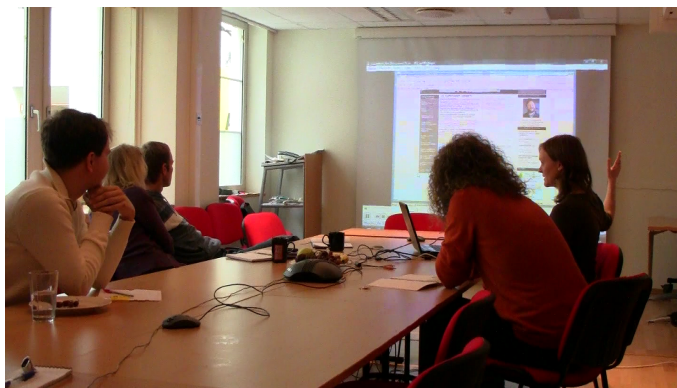
Agendaen min for møtet var å si noe om følgende:

- Generelt om brukbarhet
- Hvordan gjorde vi testen?
- Litt om testdeltakerne
- Statistiske resultater
- Gjennomgang av de ulike problemområdene

Jeg hadde lagt opp presentasjonen slik at jeg først hadde en liten introduksjon til hva brukbarhet og brukbarhetstesting er, og fortalte også litt om hvordan vi gjennomførte testen, og hva slags deltakere vi hadde testet. Deretter viste jeg resultatet fra SUS-skjemaene

og tabellen for effectiveness (se 7.1), og også hvilke av oppgavene som var mest problematiske. Jeg gikk så over til det jeg vil kalle hoveddelen av presentasjonen, hvor jeg beskrev ett og ett problemområde. Jeg begynte med de viktigste problemene først, og tok de mindre problemene etterhvert. Alle lysarkene fra presentasjonen kan du se i vedlegg F. Til sammen tok møtet i underkant av 3 timer.

Hensikten min var å skape en åpen diskusjon rundt hvert av de problemene jeg presenterte, og derfor valgte jeg å sitte under presentasjonen slik at jeg var på samme nivå som de andre tilstede. På den måten håpet jeg at det ville være lettere for alle å være delaktige i diskusjonen.



Figur 9.1: Møterommet hvor jeg holdt presentasjonen for Amnesty. Jeg sitter foran til høyre.

## 9.2 Opplysende og overbevisende for kunden

Da jeg skulle forklare problemet med signering av aksjonen for Masood Ahmed Janjua (se kapittel 7.2.1), prøvde jeg først å forklare med ord hva vi trodde var grunnen til at mange ikke hadde fått til denne oppgaven. Jeg forklarte at mange hadde signert i kommentarfeltet heller enn å signere der de skulle, og jeg presenterte kort noen av teoriene vi hadde om hvorfor dette skjedde. I vedlegg G finner du hele transkripsjonen fra denne diskusjonen. En del av grunnen kunne være at folk var vant til å signere under artikkelen på de gamle sidene som Amnesty hadde. Det kunne også være at det ikke stod en overskrift til kommentarfeltet om at dette gjaldt å legge inn en kommentar. Den siste teorien vi hadde gikk ut på hvordan layouten til siden var, nemlig at folk ikke koblet at bildet og signaturboksen øverst på siden hadde noe med artikkelen å gjøre.

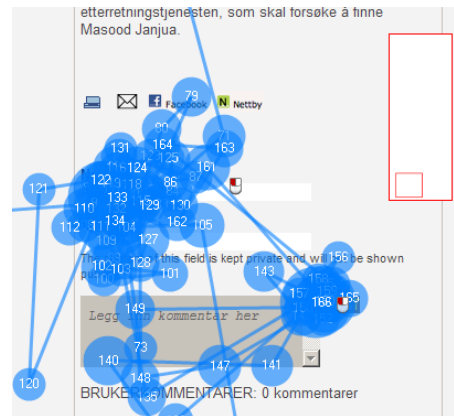
Da jeg nevnte dette under presentasjonen, fikk jeg egentlig ingen reaksjon fra de som var tilstede. Jeg forstod at de ikke helt skjønnte hva jeg mente, de sa ikke så mye, men ansiktsuttrykkene deres fortalte at de forventet å få en god begrunnelse for de teoriene jeg hadde nevnt.

For å illustrere poenget mitt, ville jeg vise to videoer av øyesporingen til to av de som hadde problemer med å fullføre oppgaven. Første videoen varte 45 sekunder, og jeg forklarte raskt at prikkene viste hvor deltakeren fokuserte blikket. Jeg merket med en gang at de tilstedeværende ble veldig interesserte.

Poenget mitt var å vise at deltakeren nesten ikke så på det store bildet i det hele tatt og i alle fall ikke så til høyre hvor signaturboksen var, men scrollet veldig raskt nedover til



Figur 9.2: Bildet viser hvor deltakerne skulle signere for å gjennomføre oppgaven, altså langt oppe til høyre på siden.



Figur 9.3: Bildet viser et blikkdiagram av hvor mange av deltakerne faktisk signerte, altså det som er kommentarfeltet som lå nederst på siden.

overskriften og ingressen og beholdt skjermbildet og blikket herfra og nedover hele tiden etterpå (se blikkdiagrammet i figur 7.2 som illustrerer det samme). Jeg måtte vise den første filmen en gang til fordi det hele gikk litt fort i starten og det tok noen sekunder før de skjønnte hvordan øyesporingen vist.

Den andre filmen spilte jeg også av, mye for å vise at denne deltakeren oppførte seg akkurat som den forrige med å ikke bry seg om det som var over overskriften. Denne deltakeren var også mye mer grundig og søkte lenge etter signaturfeltet, for hun var egentlig ganske sikker på det var et kommentarfelt under artikkelen, men hun fant ingen andre steder å signere. Denne videoen understreket poenget mitt enda tydeligere, og derfor hadde jeg valgt å vise denne til slutt.

Under den første filmen presenterte jeg underveis hvordan vi hadde resonnert for å komme fram til teorien vår, slik at de skulle skjønne hva jeg mente. Jeg merket at under den andre filmen var tilskuerne mye mer med på hva som skjedde og det var tydelig at de fulgte samme resonnement som jeg om hvorfor problemet hadde oppstått. Den andre filmen bekreftet altså det de hadde oppdaget i den første filmen. Det var også tydelig at de syntes de hadde fått mye ut av å se disse videoene. *“Ja, nå skjønner vi alt så meget bedre!”* som en av de uttalte etterpå.

En diskusjon rundt hva dette kan bety kommer i kapittel 11.3.

## 9.3 Informasjon fra varmekart

Etter at jeg var ferdig med presentasjonen og hadde gått igjennom alle problemene vi hadde funnet, begynte vi å diskutere litt rundt hvor folk ser hen når de ikke har en spesiell oppgave de skal gjøre. Jeg hadde med meg noen varmekart som viste startsidene til Amnesty da folk gjorde den første oppgaven, altså når de bare skulle gjøre seg litt kjent med siden, og tenkte at jeg kunne vise et hvor alle deltakerne hadde sett i løpet av det første minuttet de besøkte siden. På den måten kunne de få se hvilke elementer på forsiden som tiltrakk seg oppmerksomheten til brukerne.

Men idet jeg lette etter det riktige varmekartet, så de at jeg også hadde et kart for

kvinnelige deltakere og et kart for mannlige deltakere og dette var de veldig interessert i å studere for å se om det var en forskjell på kjønnene. Disse kartene har jeg vist i figur 9.4 og 9.5. Vi ser at kartene ser litt forskjellige ut, noe vi diskuterte en del. De syntes det var morsomt å se at kvinnene så mer på bildet, mens det så ut som om mennene gikk raskt på overskriftene under bildet.



Figur 9.4: Varmekart av forsiden som viser hvor alle kvinnelige deltakere har sett i løpet av det første minuttet de var på siden.



Figur 9.5: Varmekart av forsiden som viser hvor alle mannlige deltakere har sett i løpet av det første minuttet.

Det jeg har lyst til å trekke fram fra dette eksempelet er at jeg er litt i tvil angående hvor nyttig informasjon de faktisk fikk utfra disse to varmekartene. Var det egentlig noe vits i å vise forskjellen på hvordan menn og kvinner så på forsiden i denne sammenhengen? En av de som var tilstede fleipet med at “*Kanskje vi skulle hatt en side for kvinner, og en side for menn?*”. Utover at det var litt interessant i seg selv at det var forskjell på hvor kvinner og menn fikserte, så tror jeg ikke at det i denne sammenhengen hadde noe å si for brukbarheten.

En annen ting de lurte på var om folk la merke til *Bli medlem*-boksen til høyre på forsiden. Både kvinner og menn hadde sett der, noe som var bra fordi det var viktig for dem å promotere budskapet om at man kan bli medlem. På denne måten har varmekart en nytteverdi i forhold til å bekrefte at elementer de vil skal være synlige faktisk er synlige, *men* dette kunne man like gjerne fått fram ved å vise et samlet varmekart for begge kjønnene.

## Praktiske og metodiske utfordringer

---

Dette kapitlet vil handle om utfordringer man kan komme over når man bruker øyesporing, og vil være et forsøk på å besvare forskningsspørsmål 4: *Hvilke praktiske og metodiske utfordringer møter man på?* (se kap. 1.2). Jeg gjorde et intervju med Terje Røsand som jobber med den tekniske delen på brukbarhetslaboratoriet ved NSEP og IDI for å høre hvilke erfaringer han hadde gjort seg rundt dette temaet. En transkripsjon av hele intervjuet finner du i vedlegg H. I tillegg vil jeg trekke fram noen av mine egne erfaringer fra da jeg gjennomførte brukbarhetstesten tidligere i oppgaven, og fra prosjektoppgaven jeg gjorde høsten 2009, hvor jeg også benyttet øyesporing [25].

### 10.1 Programvaren

Siden både jeg og Terje Røsand bare har brukt Tobii Studio, kan vi bare snakke om våre erfaringer med programvaren ut fra dette. Terje Røsand synes generelt at programvaren er enkel både å lære seg og å bruke. Men et problem som har dukket opp gjelder det å kunne skille bilder på en god måte. Når en bruker for eksempel har besøkt samme nettside to ulike ganger, men med forskjellig hensikt, da bør disse to gangene skilles fra hverandre. Dessverre er ikke dette spesielt enkelt å gjøre med den versjon av Tobii Studio som vi har brukt. Programmet har innebygd funksjon for å skille på for eksempel URL eller innhold på siden, men dette skiller bare ulike sider, det skiller ikke på tid eller oppgave. Dette problemet hadde jeg også da jeg gjennomførte en brukbarhetstest i prosjektoppgaven jeg hadde i høst. På testen vi gjennomførte tidligere i denne oppgaven klarte vi å løse det, om enn på en *veldig* tungvindt måte. For hver ny oppgave, måtte vi endre pixelstørrelsen på nettleservinduet med én, slik at ingen av oppgavene hadde like vindusstørrelser, og slik kunne vi sortere på URL og vindustørrelse. Brukeren merket selvfølgelig ikke noe av dette, men det krevde en del mer arbeid fra vår side.

Jeg hadde også litt problemer med gruppering av øyesporingsdata fra flere brukere. Selv om de hadde gjort samme oppgave og besøkt de samme sidene, var det ikke alltid de ble plassert i samme gruppe og dermed ville man ikke kunne se for eksempel et samlet varmekart for disse. Grunnen til at de ble plassert i forskjellige grupper kunne være at bittelitt av innholdet på siden hadde endret seg (noe var vanskelig å unngå siden testen pågikk over flere dager) eller at brukeren selv forstørret skjermbildet i nettleseren for å kunne se bedre. For å løse dette måtte jeg manuelt gå inn å flytte data fra en gruppe til en annen, og det var ikke alltid like enkelt, siden det ikke var merket hvilken bruker eller oppgave hvert datasett gjaldt.

I intervjuet med Terje Røsand nevner han også at Tobii Studio krever *veldig* mye maskinressurser, noe som kan gå ut over annen maskinfunksjonalitet. Vi opplevde det på testen vi gjennomførte tidligere i studien ved at tastaturet fikk treg reaksjonstid mens testen pågikk. Brukerne hadde problemer med å skrive skikkelig og de trodde at det var noe galt med tastaturet, siden det hendte at enkelte tastetrykk ikke vistes igjen på skjermen. Det viste seg senere at det ikke var tastaturet som var problemet, men faktisk det at Tobii Studio krevde mer ressurser enn maskinen kunne håndtere bra.

### 10.2 Sittestilling og avstand til øyesporeren

En ting Terje Røsand nevner spesielt når det gjelder den praktiske delen ved å bruke øyesporing, er at avstanden til øyesporeren kan kan av og til skape problemer. For at øyesporingen skal fungere skikkelig så kreves det at avstanden fra øyesporeren til øynene ligger mellom 50-80 cm. Dette stiller man inn ved å flytte stolen i riktig posisjon før kalibreringen starter. Problemet ligger i at denne avstanden ofte er litt større enn mange er vant til å sitte foran en skjerm, og når de blir oppslukt av å løse oppgavene, glemmer de seg og lener seg framover slik at vi mister data.

### 10.3 Nøyaktighet av målingene

En annen ting som både jeg og Terje Røsand har merket oss er at selv om teknologien har blitt ganske bra, så vises ikke alltid resultatet 100% nøyaktig. Terje Røsand nevner også et eksempel fra en hovedoppgave av en student på Dragvoll som skulle bruke øyesporing til å se om det var forskjell på dysletikere og ikke-dysletikere når de las en tekst [11]. Denne studenten hadde kommentert at det var vanskelig å si med sikkerhet hvilken linje testpersonen faktisk *så* på. Dette merket jeg også da jeg skulle se på resultatet fra testingen vår. For eksempel var det enkelt av deltakerne hvor man skjønnte at de hadde sett på hovedmenyen ut fra hvor de klikket og hva de sa, og på blikkdiagrammet kunne man også se at øynene var riktig plassert horisontalt på skjermen, men det hele lå cirka en halv centimeter under hovedmenyen, og da på undermenyen i stedet (se for eksempel figur 7.5 i kapittel 7).

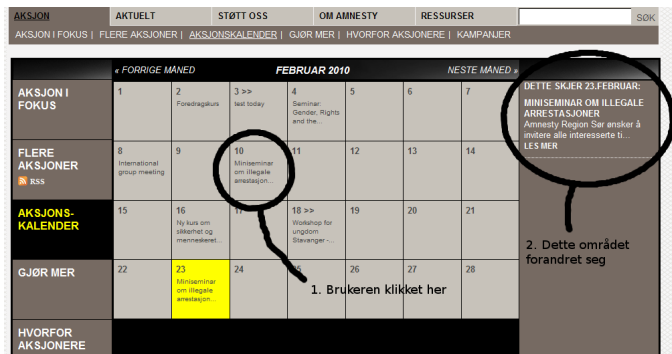
### 10.4 Signal fra øyesporeren faller ut

Et problem som jeg har lyst å trekke fram skjedde under brukbarhetstesten tidligere i studien. Vi hadde en testdeltaker som besøkte aksjonskalenderen til Amnesty under den første oppgaven (se figur 10.1), altså da de skulle gjøre seg litt kjent med siden. Testdeltakeren klikket på en av postene i kalenderen og ventet så en stund før han til slutt sier *“Er det siden som er treg, eller bare skjer det ingenting når man trykker?”*. Etter ennå en stund oppdaget han plutselig at informasjonen han ventet på stod til høyre for kalenderen. Han fortalte at han hadde fokusert blikket på kalenderen og derfor ikke sett at det dukket opp noe til høyre.

Dette hadde vært veldig interessant å få øyesporingsdata på for å se om han faktisk snakket sant, og også for å kunne vise problemet til Amnesty på et senere tidspunkt.



## 10.4. SIGNAL FRA ØYESPOREREN FALLER UT



Figur 10.1: Figuren viser aksjonskalendersiden, hvor en deltaker ikke oppdaget at informasjonen til høyre forandret seg når han klikket på ulike dager i kalenderen.

Dessverre så falt signalet fra øyesporeren ut akkurat da dette skjedde, uten at vi vet hvorfor, så dermed hadde vi ingen data på dette. Det var heller ikke så mange andre testdeltakere som hadde besøkt aksjonskalenderen, altså var dette en unik mulighet til å få øyesporingsdata for dette problemet.



# **Del IV**

## **Drøfting og konklusjon**

---

## Drøfting

---

Dette kapittelet vil ta for seg drøfting og tolkning av resultatene jeg har kommet fram til i kapittel 7-10 og koble dette opp mot forskningsspørsmålene i 1.2.

### 11.1 Problemidentifisering

Forskningsspørsmål nummer 1 handler om i hvilken grad øyesporing bidrar til å identifisere problemer i brukbarhetstesting. For å finne svar på dette gjennomførte jeg en brukbarhetstest hvor jeg skulle se på hvilken ekstra informasjon vi fikk fra øyesporingen i forhold til det vi hadde funnet ut ved bruk av vanlig testmetodikk. Gjennomføringen av testen er beskrevet i kapittel 6 og resultatene fra testen finner du i kapittel 7. Nedenfor vil jeg trekke fram de viktigste konklusjonene vi har funnet ut på bakgrunn av resultatet fra testen, og også se hvordan det passer med de erfaringer NetLife Research har gjort seg når det gjelder å identifisere problemer ved hjelp av øyesporing.

#### 11.1.1 Visuelle vollgraver

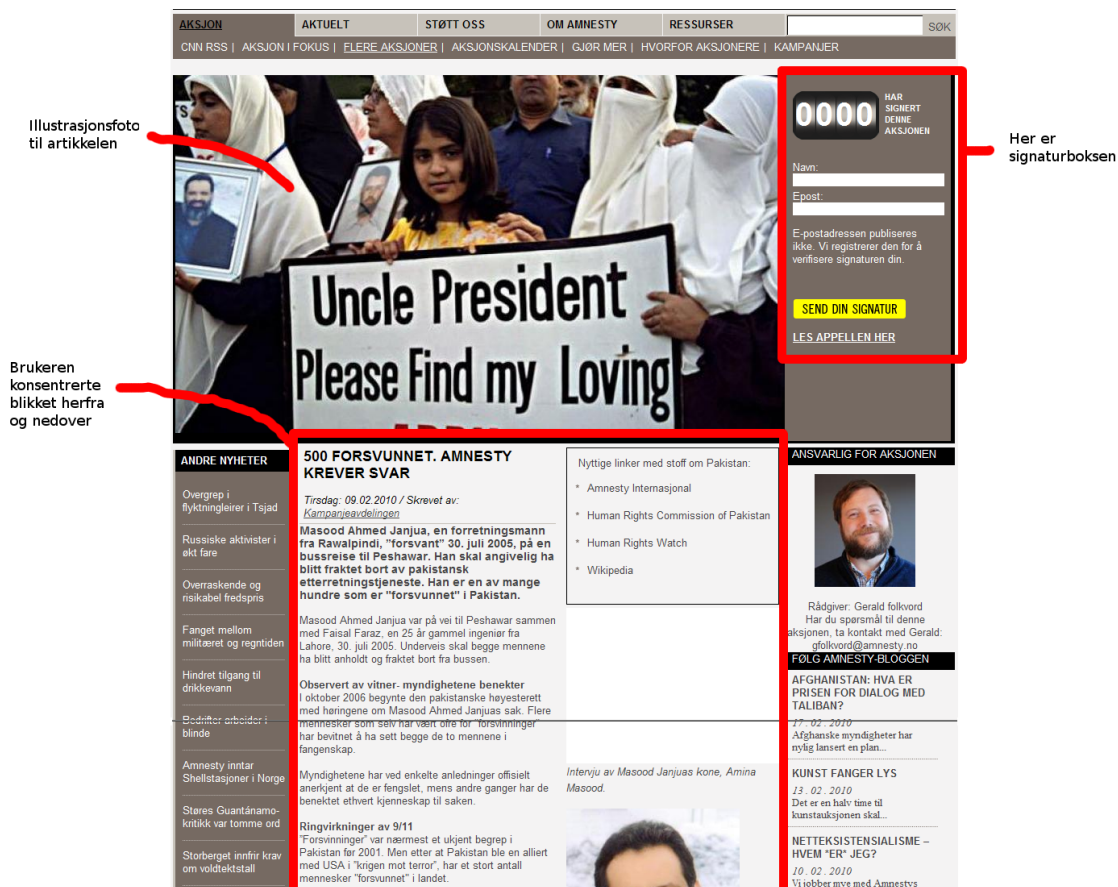
Begrepet *Visuelle vollgraver* har jeg lånt fra NetLife Research som bruker det om elementer i et design som gjør at andre elementer ikke oppdages av brukeren (se 8.2.2). Blant annet nevner de et eksempel fra Telenor sine gamle nettsider hvor noen store bilder førte til at brukeren konsentrerte seg om innholdet som lå nedenfor disse og globalnavigasjonen (som lå over bildene) kom helt i bakleksen.

Dette tror jeg kan sammenlignes med det jeg oppdaget på Amnestys nettsider. I 7.2.1 beskriver jeg en situasjon hvor flere av testdeltakerne gjorde feil. Her kunne vi se på øyesporingsvideoene at mange ikke brydde seg med å se på den øverste del av skjermen, men konsentrerte seg fullt og helt om området nedenfor. Dette førte til at de ikke oppdaget signaturboksen oppe til høyre på siden, som de trengte for å gjennomføre oppgaven (se figur 11.1).

Vi ser altså at øyesporing kan avdekke feil i design som gjør at folk ikke oppdager viktige elementer i brukergrensesnittet.

#### 11.1.2 Bekrefting eller forsterking av konklusjoner

Øyesporing kan brukes til å bekrefte eller forsterke konklusjoner vi har gjort oss på bakgrunn av den tradisjonell testmetodikk. I 7.2.2 beskriver jeg et problem hvor testdeltakerne ikke hadde forstått egentlig hva ordet *Audiovisuelt arkiv* betydde, selv om de trodde de



Figur 11.1: Figuren viser at det store illustrasjonsbildet kan være en visuell vollgrav, siden brukerne konsentrerte blikket nedenfor dette, mens de egentlig skulle signere i signaturboksen.

gjorde det. På bakgrunn av hva de kommenterte underveis, eller sa i intervjuet etterpå, fant vi ut at de fleste trodde det gjaldt lyd eller video, og ikke bilder (hvilket det gjør, og som var det vi *egentlig* var ute etter å finne). De fleste navigerte til rett hovedfane, som var *Ressurser*, men derfra valgte de sjelden riktig alternativ på undermenyen (se sidekart i figur 5.1. Ved hjelp av øyesporingsdata kunne jeg beregne hvor lang tid det tok fra brukerne oppdaget menyvalget *Audiovisuelt arkiv*, til de trykket på det. Sammenlignet med et annet valg som mange trykket på, ser vi at de brukte *utrolig* lang tid før de valgte å klikke på riktig menyvalg. Dette bekreftet det vi allerede hadde visste, at folk ikke trodde *Audiovisuelt arkiv* inneholdt bilder. Men i tillegg forsterket det konklusjonen vår ganske kraftig ved å vise at folk ikke *vurderte* alternativet engang.

Når det gjelder navigeringen på nettsiden, så skriver jeg i 7.2.3 at jeg hadde en teori om at den horisontale undermenyen var litt unødvendig å ha, siden man allerede hadde de samme valgene i venstremenyen. Jeg tok en kikk på øyesporingsdataene, og fikk bekreftet at folk hovedsaklig fokuserte på venstremenyen heller enn den horisontale undermenyen (se figur 7.5).

NetLife Research sier at som analyseringsverktøy så bruker de øyesporing hovedsaklig til å støtte opp under konklusjoner de allerede har kommet fram til basert på øvrig analyse (se 8.2.1), og dette passer godt med mine funn.

Dette viser at øyesporing kan være et nyttig verktøy i problemidentifisering hvis man

er i tvil rundt de konklusjonene man har kommet frem til basert på data fra tradisjonell brukbarhetstesting.

### 11.1.3 Forstå brukernes mentale modell av systemet

I 7.2.4 beskriver jeg en situasjon hvor øyesporingsdataene fortalte meg at de var usikre på hvilken av hovedmeny-fanene de skulle velge for å finne en bestemt artikkel. På blikkdiagrammene (figur 7.6-7.9) ser vi at de fleste vurderte begge alternativene omtrent like mye. Dette forteller meg noe om hvor brukeren synes artikkelen hører hjemme, altså kan vi si noe om hvilken mental modell han har av nettstedet. Med mental modell mener jeg her hvordan brukeren oppfatter navigasjonsstrukturen på nettsiden. Hadde brukeren for eksempel vurdert alle menyvalgene mer likeverdig kunne det kanskje fortalt oss at han ikke synes det høres hjemme noe sted, men siden han stort sett vurderer bare to av de, forstår vi at han tror det kan høre hjemme under begge valgene.

Vi får jo også vite med sikkerhet *hvilke* to alternativer det står i mellom. Vi kan for eksempel tenke oss at alle testpersonene tilfeldigvis hadde valgt *Om Amnesty* til slutt. Da hadde vi ved hjelp av tradisjonell testmetodikk ikke nødvendigvis oppdaget at de også vurderte *Støtt oss* eller noen andre alternativer, og vi hadde trodd at alt var som det skulle. Men ved hjelp av øyesporingsdata kan vi se at de vurderte *Støtt oss* ganske sterkt også.

Dette forteller oss at øyesporing kan være til hjelp for å forstå hvordan navigasjonsstrukturen på for eksempel en nettside bør være for å gi økt brukbarhet. Hvis man kan “lese” brukerens mentale modell gjennom å se hvor han retter oppmerksomheten sin, så kan vi tilpasse navigasjonsstrukturen til denne. Og når en bruker opplever at strukturen av siden stemmer overens med hans/hennes mentale modell, vil han synes at siden er god å bruke.

### 11.1.4 Oppsummering

Nielsen og Pernice påstår i [15, s. 109] at “*Eyetracking technology is very good at exposing design miscues, so you can almost get into users’ heads and see what they are thinking. You know if they look at something and spend time reading it or if they stare at a link, word, or image*”. Vi kan nesten legge til et “*or not*” på slutten av sitatet, for det er minst like viktig å finne ut hvor brukeren *ikke* ser hen på skjermbildet. Alt i alt synes jeg dette oppsummerer ganske bra det vi har diskutert over. Gjennom å se hvor brukeren fester blikket kan vi forstå hva slags mental modell han/hun har. Vi kan oppdage hva som trekker oppmerksomhet, eller forhindrer at noe annet får nødvendig oppmerksomhet (visuelle vollgraver). Og når vi vet hvor brukeren ser eller leser, kan vi bruke dette til å støtte opp under teorier vi har kommet opp med på forhånd.

## 11.2 Nytteverdi for brukbarhetstestere

Her skal jeg trekke inn noen viktige poeng som kom fram under besøket mitt hos NetLife Research hvor jeg intervjuet Thor Fredrik Eie om hvilken nytteverdi de ser angående øyesporing i brukbarhetstesting. Jeg kommer først og fremst til å basere diskusjonen på resultatet du finner i kapittel 8.

### 11.2.1 Tilpassing av oppfølgingsspørsmål under testen

Et veldig interessant poeng som Thor Fredrik Eie fra NetLife Research drar fram da jeg spør om hvilken nytteverdi de ser i å bruke øyesporing i brukbarhetstesting, er det at han som testleder kan benytte seg av øyeporingen i sanntid, altså *imens* testpersonen utfører testen, til å tilpasse oppfølgingsspørsmål underveis (se 8.4). Dette var et nytteområde jeg ikke hadde oppdaget før, men jeg skjønnte umiddelbart det hadde noe for seg. Når man gjør brukbarhetstester, spør man jo gjerne brukeren om ting underveis, for eksempel hva de tenker om spesielle elementer på nettsiden. Hvis brukeren da ikke har oppdaget dette elementet ennå, så kan det gjøre at man “styrer” brukeren i en viss retning, og det vil man jo helst unngå når man gjør brukbarhetstesting. Øyesporing kan altså være nyttig for utviklere ved at man kan gjennomføre brukbarhetstester på en bedre måte, og dermed kan får man riktigere resultat.

### 11.2.2 Nyttig i dialog med kunden

NetLife Research synes for øvrig at det å kunne vise fram øyesporingsvideoer hvor brukeren har hatt problemer med noe under testen er “*en veldig effektiv måte å kommunisere med kundene på*” (sitat fra vedlegg E). I følge Shneiderman og Plaisant så er ikke dette å vise fram videoer for kunden noe nytt, det har også hatt verdi før øyesporing ble aktuelt: “*The reactions of designers to seeing videotapes of users failing with their interfaces is sometimes powerful and may be highly motivated. When designers see participants repeatedly picking the wrong menu item, they realize that the label or placement needs to be changed*” [24, s. 149].

Likevel, ved å vise hvor testpersonen fokuserer, så tilføyer man en ekstra dimensjon i forhold til å vise kunden hvordan brukerne benytter produktet deres. Før kunne man bare følge med på hva brukeren faktisk *gjorde*, ved museklikk eller scrolling for eksempel, eller hva han/hun sa, men med øyesporing kan man i tillegg se hvor testpersonen har oppmerksomheten. NetLife Research har opplevd at når man har denne kombinasjonen av øyesporing og et godt sitat fra testpersonen på et videoklipp, så får man et slagkraftig verktøy overfor kunder som ikke har vært med på testen i sin helhet.

### 11.2.3 Må ha effektiv analyse av data

Når det gjelder analysering av problemer, så bruker NetLife Research øyesporing først og fremst til å understøtte konklusjoner de har kommet fram til på bakgrunn av tradisjonell testmetodikk, for eksempel høyttenkning (se 8.2.1). De har også brukt det noe til å forstå problemer som de ikke har funnet ut av ellers, for eksempel det med at noen elementer stjeler all oppmerksomhet slik at brukeren ikke oppdager andre viktige ting. Men hovedsaklig så bruker de ikke så mye tid på å gå igjennom øyesporingsdataene fordi det er en langsom prosess.

Goldberg m.fl. hevder i [8] at øyesporing ikke har blitt brukt særlig ofte i vanlig brukbarhetstesting blant annet fordi det er tidkrevende og krever omfattende datareduksjon. For bedrifter som lever av å blant annet selge brukbarhetstester så gjelder det altså å finne en måte hvor øyesporing kan brukes effektivt og likevel være nyttig. Sånn sett tror jeg NetLife Research har gjort et fornuftig valg når de sier at de ikke bruker så mye tid på å analysere dataene grundig, men heller noterer seg steder som kan være interessante underveis i testen og ser så bare på øyesporingsdata fra disse stedene i etterkant.



### 11.2.4 Oppsummering

Hvis vi skal oppsummere de poengene vi har snakket om i denne bolken og koble det opp mot forskningsspørsmål 2 (se 1.2), så vil jeg si at det finnes hovedsaklig 3 felt hvor øyesporing kan ha noen nytteverdi. For det første er det nyttig i forhold til å kunne gjennomføre bedre brukbarhetstester (som jeg diskuterte i 11.2.1). Det andre feltet gjelder i forhold til kunden, hvor øyesporingsdata gjør kommunikasjonen lettere mellom utviklerne og kunden (11.2.2). Til slutt har også øyesporing en verdi som hjelp til å gjøre riktige konklusjoner angående problemer man har oppdaget i brukbarhetstesting, men her bruker man bare tid på det som åpenbart gir resultater.

## 11.3 Som kommunikasjonsverktøy

I 8.3 nevner jeg at NetLife Research ser på øyesporing som et *slagkraftig verktøy* når de skal presentere funnene sine for kunden. Med slagkraftig verktøy mener vi at det kan brukes som argumenter til å slå i bordet med overfor kunden, man får mye mer tyngde i argumentene sine når man kan vise til data. Dette opplevde jeg også noe av da jeg skulle presentere resultatet fra brukbarhetstesting vi gjorde for Amnesty. I 9.2 beskriver jeg en situasjon hvor jeg først hadde fortalt om tre teorier vi hadde angående et problem med å få signert på aksjoner. Jeg viste så to øyesporingsvideoer som understøttet den ene teorien om hvorfor problemet oppstod. Dette viste seg å være veldig opplysende for de som var tilstede, og de forstod mye bedre hvorfor vi hadde kommet fram til denne teorien. Dette forteller meg at øyesporing kan være til hjelp for å forklare problemer for kunder, slik at de skjønner mer hvorfor ting ikke fungerer. Men samtidig forteller det meg at dette er et veldig sterkt kort vi som brukbarhetstestere sitter med på hånden i forhold til det å overbevise kunder om at de teoriene vi har kommet fram til er riktige.

Det er flere farer forbundet med dette. Som jeg nevnte, så hadde vi i utgangspunktet tre teorier om hvorfor problemet hadde oppstått. Disse var ikke gjensidig utelukkende, men til sammen var kanskje det grunnen til at så mange ikke hadde fått til denne oppgaven. To av disse teoriene hadde vi kommet fram til kun på bakgrunn av hva testdeltakerne sa underveis og i intervjuet etterpå, mens det at de ikke koblet at illustrasjonsfotoet og signaturboksen hørte til artikkelen under bildet, fant vi ved hjelp av øyesporingen (se figur 7.2). Det at vi faktisk hadde data på at folk hoppet glatt over bildet og gikk direkte til teksten under bildet, og holdt fokuset der resten av tiden, gjorde at vi trodde den teorien var viktigere enn de andre. Det kan hende vi hadde rett, men man bør være klar over at øyesporingsdata *kan* gjøre oss “blinde” for andre feil.

Dette kan videre føre til at vi som brukbarhetstestere fokuserer for mye på de feilene vi kan underbygge ved øyesporingsdata i forhold til andre brukbarhetsproblemer når vi skal presentere funnene våre for kunden.

### 11.3.1 Bruk av varmekart (heat maps)

En annen fare med øyesporing som kommunikasjonsverktøy er at man i prinsippet har mulighet til å mikse og trikse med dataene slik at det “passer” til det man er ute etter. Dette gjelder først og fremst varmekart, fordi det er utrolig mange faktorer som spiller inn på hvordan et varmekart ser ut (Bojko skriver en del om dette i [3]). For eksempel kan litt endring i tidsperioden varmekartet skal vise data for, forandre hele bildet. Eller

hvis man har 10 deltakere og en av disse har brukt veldig mye tid på å se på et element som de andre ikke har brydd seg særlig mye om i det hele tatt, så kan man forandre varmekartet fullstendig ved å ta bort data fra denne ene personen fordi det ikke “passer” med det generelle resultatet. Dette er en alvorlig felle å gå i, da man potensielt sett kan vise uriktige resultater til kunden. Det blir derfor snakk om et tillitsforhold mellom kunden og brukbarhetstesterne, hvor kunden må stole på genuiniteten til de dataene som vises fram.

NetLife Research er for øvrig forsiktige med å bruke varmekart når de presenterer funnene sine for kunden (se 8.3.2). Dette er fordi de vanligvis har bare 8 deltakere på en brukbarhetstest, noe som er vanlig praksis (se kapittel 2.2.2). Et så lite utvalg er for lite til å trekke statistisk gyldige slutninger av varmekart, da det ligger i varmekartets natur at til flere deltakere, til mer robust blir det (statistisk sett).

Likevel vil mange kunder gjerne ha øyesporing i brukbarhetstesting, til tross for et tillegg i pris, som NetLife Research opplever (se 8.1). Noe av grunnen til dette kan være at datarepresentasjoner som for eksempel varmekart er intuitivt enkle å forstå seg på, som Bojko forteller om i [3]. Varmekart er også veldig enkle å generere, og man trenger ikke mye erfaring med øyesporing for å kunne lage flotte varmekart.

I 9.3 trekker jeg fram en episode fra presentasjonen min for Amnesty hvor de var veldig interesserte i å se på to varmekart som viste forskjellen på hvor menn og kvinner så på forsiden. Det som var poenget mitt her var at alle varmekart sier ikke nødvendigvis noe nyttig om brukbarheten til siden, noe Thor Fredrik Eie fra NetLife Research også sier (se 8.3.2).

Det jeg vil fram til her er at man bør være klar over hvorfor man viser bestemte varmekart. Det er lett å lage varmekart som er stilige og morsomme å se på, men det betyr ikke at de nødvendigvis formidler den viktige informasjonen på en god måte. Samtidig er det tydelig at kunder ofte er veldig interesserte i å se på slike visuelle representasjoner av øyesporingsdata, kanskje nettopp fordi det er noe håndfast, noe mer enn bare ord fra de som har gjennomført brukbarhetstesten.

### 11.3.2 Oppsummering

Det vi har diskutert her kan kobles opp mot forskningsspørsmål nummer 3, altså om hvilken nytteverdi øyesporing kan ha i dialogen mellom brukbarhetstesterne og kunden. Jeg tenker da på kunden som både de som betaler for testen og de som skal bruke resultatet videre (altså utviklere). Vi ser at øyesporing kan bidra til å gjøre det lettere for kunden å forstå problemene vi har funnet. Det er også et sterkt verktøy for brukbarhetstesterne fordi det gir mer tyngde og troverdighet når man kan vise til data heller enn bare formidle påstander.

## 11.4 Praktiske og metodiske utfordringer

Her vil jeg diskutere noe av det vi har funnet ut gjennom studiet som kan knyttes opp mot forskningsspørsmål nummer 4 (se kapittel 1.2), altså om hvilke praktiske og metodiske utfordringer man kan møte på ved bruk av øyesporing i brukbarhetstesting. Jeg baserer diskusjonen først og fremst på resultater jeg har beskrevet i kapittel 10, men trekker også inn noen poeng som kom frem under intervjuet med NetLife Research fra kapittel 8.

### 11.4.1 Utstyret kan være ustabil

Til tross for at store tekniske framskritt innenfor øyesporing de siste årene, hender det likevel at både programvaren og øyesporeren slutter å fungere når man trenger den. Som jeg skrev om i 10.4 så hadde vi en episode under brukbarhetstesten i første del av studien hvor vi gjerne skulle fått øyesporingsdata fra, men hvor utstyret da sviktet og vi mistet den ene sjansen vi hadde. NetLife Research har også erfart at dette skjer av og til (se 8.5.2), og at slike problemer har videre har ført til at noen ansatte ikke trives med å gjennomføre tester med øyesporing. Denne typen problemer merker ikke testdeltakerne noe av heldigvis, men det er et alvorlig problem for de som trenger dataene.

Men det finnes også problemer som faktisk kan være til hindring for deltakerne under testen. I 10.1 skriver jeg om et problem vi kom over under brukbarhetstesting hvor programvaren brukte så mye datakraft at reaksjonstiden på tastetrykk fra brukeren ble påvirket. De opplevde at når de trykket på en tast, så reagerte ikke maskinen slik at det vistes på skjermen, og de trodde at det var noe galt med tastaturet. Denne feilen var et nokså stort irritasjonsmoment for noen av deltakerne, og slike ting vil man jo helst unngå under en brukbarhetstest.

Vi må altså være klar over at tekniske feil kan oppstå når vi gjør brukbarhetstester med øyesporing og det er derfor viktig i et systemutviklingsprosjekt å ikke være 100% avhengige av øyesporingsdata for å få et vellykket resultat.

### 11.4.2 Nøyaktighet av målinger

Et annet poeng jeg har lyst til å dra fram er dette med nøyaktighet av målingene fra øyesporeren. Terje Røsand nevner i 10.3 en student som hadde hatt litt problemer med å finne ut hvilken linje en bruker leste på skjermen bare ved å se på øyesporingsdataene. Jeg skriver at dette var noe vi også erfarte under brukbarhetstesten, hvor det kunne se ut som om deltakeren så på undermenyen, men vi skjønnte at det var hovedmenyen han egentlig så på. I mange tilfeller vil ikke dette ha noe særlig å si fordi man skjønner at for eksempel en bruker har sett *på* en knapp heller enn like under den hvor det ikke er noe.

Jeg støtte likevel på et problem på grunn av dette da jeg analyserte resultatet fra brukbarhetstesting. Underveis i testingen oppdaget jeg at nesten ingen av deltakerne brukte den horisontale menyen med hvit bakgrunn som lå over venstremenyen og under hovedmenyen og undermenyen. Jeg hadde en teori om at de ikke oppdaget dette feltet, siden det virket litt usynlig der det lå klemt mellom menyene med sterke bakgrunnsfarger, og ville se om jeg kunne fått bekreftet dette i øyesporingsdataene. Dessverre var målingene såpass unøyaktige at jeg ikke kunne konkludere med at de ikke hadde oppdaget feltet, og jeg kunne dermed heller ikke bruke det som argumentasjon da jeg skulle presentere resultatet for Amnesty.

Øyesporingsteknikken har nådd langt i dag og er *som regel* nøyaktig nok til at man kan få fornuftige resultater, men disse to avsnittene viser at det finnes situasjoner da man gjerne skulle hatt mer presise målinger.

### 11.4.3 Ikke alle egner seg som testdeltakere

Noe NetLife Research nevner er at man ikke kan bruke øyesporing på alle personer (se kapittel 8.5.1). Øyesporeren har for eksempel problemer med å lese øynene til folk som har briller. Dette opplevde jeg også da jeg gjennomførte prosjektoppgaven min høsten

2009, hvor jeg skulle teste en deltaker som brukte briller [25]. Vi fikk ikke til å kalibrere øyesporeren i det hele tatt, og det er jo frustrerende både for deltakeren og testteamet når man ikke engang får begynt testen.

Et annet problem forårsaket av testdeltakerne er at de kan sitte nokså urolig slik at de faller utenfor “boksen” hvor øyesporeren kan lese av øynene. Det hyppigste problemet vi støtte på under testen vår, og som Terje Røsand nevner spesielt i 10.2, var at deltakerne hadde en tendens til å lene seg framover når de ble oppslukte av oppgaven og kom da for nær øyesporeren. Noen av testdeltakerne kommenterte også før vi gjorde testen at de følte de satt enten for nært eller for langt i fra skjermen da vi hadde stilt stolen slik at øyesporeren tok inn signal.

En brukbarhetstest bør jo reflektere omgivelser som ville vært naturlig for brukerne, og hvis man legger begrensinger på hvordan brukeren kan sitte, så handler man i mot dette prinsippet.

## 11.5 Metodediskusjon

Her vil jeg diskutere litt styrker og svakheter ved de metodene jeg har valgt i forhold til å besvare forskningsspørsmålene mine (se 1.2).

### 11.5.1 Forskningsspørsmål 1

For å få svar på det første forskningsspørsmålet, om hvilken grad øyesporing bidrar til å identifisere problemer, valgte vi altså å gjennomføre en brukbarhetstest. En brukbarhetstest er i prinsippet en kvalitativ undersøkelse fordi man hovedsaklig ønsker å sette fingeren på *hvilke* feil eller mangler systemet man tester har slik at man kan rette opp disse, heller enn at man ønsker å få et tall for *hvor mange* som gjorde feilen. I møte med de som leser funnene fra slike brukbarhetstester blir resultatet dessverre ofte bedømt som om det var en kvantitativ studie i stedet, når det egentlig er helt andre kriterier som bedømmer kvaliteten på kvalitative studier. Michael Hughes skriver om dette i [9] og foreslår også metoder for å sikre at kvaliteten på resultatet fra brukbarhetstester blir god.

Det er i hovedsak tre ting han trekker fram for å kunne bedømme dette:

- *Intern validitet* som handler om hvor troverdig er det som blir testet. Tester vi funksjonaliteten som brukeren faktisk ville ha brukt (og kanskje vært avhengig av) ute i den virkelige verden? Og forstår vi faktisk det vi observerer under testen?
- *Ekstern validitet* (også kalt *generaliserbarhet*) som handler om hvor sannsynlig er det at situasjonen vi tester kan overføres til en reell situasjon. Her kommer blant annet valg av testpersoner inn i bildet. Tester vi folk fra de rette brukergruppene for eksempel?
- *Reliabilitet* handler om hvor sannsynlig det er at forsøket kan gjentas med samme resultat.

Jeg vil nå prøve å relatere disse begrepene til brukbarhetstesten vi gjorde for å besvare det første forskningsspørsmålet. Når det gjelder *indre validitet*, så kan jeg vise til at for å komme fram til realistiske oppgaver, hadde jeg en dialog med Amnesty før vi gjennomførte testingen. Amnesty kjenner bedre til målgruppene for sine nettsider enn jeg

gjør, og vet derfor litt mer om hvilke ting en bruker ville være interessert i å bruke. I tillegg benyttet jeg meg av 5 personas som Amnesty hadde utarbeidet tidlig i prosjektet sitt, for å se litt hva slags type mennesker som ville bruke nettsidene deres.

Under dette punktet vil jeg også trekke fram måten vi valgte å gjennomføre øyesporing i brukbarhetstesten på. Vi valgte å gjennomføre øyesporing med *samtidig* bruk av høyttenkning, og som Hughes nevner i [9, s. 493] så er høyttenkningsteknikken en viktig metode for å vite at du faktisk forstår det du observerer. Men når man jobber med øyesporing må man være litt forsiktig når man bruker denne teknikken, fordi “*studies have shown the [think-aloud] method affects where participants look on the display*” [6, s. 1139]. En person har nemlig en tendens til å se mer på de elementene han snakker om, og hvis testlederen da for eksempel har spurt deltakeren et spørsmål, er det med på å bestemme hvor deltakeren ser, noe som drar ned den indre validiteten. Jeg tror at hvis jeg skulle gjort en ny, lignende studie av dette, så ville jeg valgt en annen metode for å skille data fra øyesporing og data fra tradisjonell brukbarhetstesting.

I forhold til *ekstern validitet* så ser vi av tabell 6.1 i kapittel 6 at vi hadde en blanding av deltakere som kjente til Amnesty fra før og de som ikke gjorde det, og det samme gjelder hvor ofte de hadde besøkt den gamle nettsiden. Denne fordelingen var grei fordi vi ville teste om det var noen forskjell i brukbarheten for de som kanskje ikke vet så mye om Amnesty fra før, og de som er godt kjent (med for eksempel terminologi som brukes, osv.). I tillegg ser vi at vi har et overtall av kvinnelige deltakere, men vi tror likevel det er realistisk nok siden 4 av de 5 personasene Amnesty hadde utarbeidet var kvinner, altså tror Amnesty selv at det er et overtall av kvinner i målgruppen deres. Når det gjelder aldersgrupper, så ser vi av tabellen at vi har bare 1 under 20 år, og ingen fra 50 og oppover. Dette er det eneste minuset jeg kan se angående utvalget av testdeltakere. Ungdommer under 20 år hadde vært interessant å teste fordi de kanskje hadde sett annerledes på noen ting, mens de over 50 kanskje har generelt dårligere datakunnskap og kunne dermed gitt et annerledes resultat enn resten.

Når det gjelder *reliabiliteten* så kan jeg vise til at både jeg og Terje Røsand gikk igjennom resultatene fra brukbarhetstesten og vi diskuterte problemene med hverandre. På den måten fikk jeg dobbeltsjekket at resultatene jeg hadde kommet fram til stemte, noe som er med på å øke reliabiliteten. Jeg vil også trekke fram testopplegget jeg fulgte er en standardisert metode, noe som øker reliabiliteten ytterligere.

## 11.5.2 Forskningsspørsmål 2

For å få svar på forskningsspørsmål nummer 2, hadde vi valgt å gjøre et *intervju* med NetLife Research. Forskningsspørsmålet spør hvilken nytteverdi utviklere ser i bruken av øyesporing som til daglig benytter det i brukbarhetstesting. Her kunne vi for eksempel ha sendt ut en spørreundersøkelse til mange firmaer som gjør dette, men problemet var at i Norge er det ikke så mange som bruker øyesporing, slik at vi ville ikke hatt tilstrekkelig med informanter til en kvantitativ undersøkelse. På grunn av dette var det nokså naturlig å velge å gjøre et intervju med en av de som bruker øyesporing og heller gå litt i dybden på spørsmålet enn å finne generelle tendenser i et bredere utvalg. Alt i alt synes jeg at å velge intervjumetoden her var et godt valg, og jeg fikk veldig rikt utbytte av dette intervjuet med hensyn til å besvare forskningsspørsmålet, men også til å få bekreftet noe av resultatene vi hadde fått på de andre forskningsspørsmålene. Intervjuer blir (i følge [22, s. 131]) ofte kritisert for å være subjektive og ikke objektive, men slik forskningsspørsmålet vårt er

formulert, så er det faktisk det subjektive synspunktet vi er interesserte i å få vite.

Som jeg nevnte såvidt i forrige avsnitt så fokuserte vi ikke på å finne generelle tendenser i et bredt utvalg. Generaliserbarhet er vanskelig å oppnå ved intervjumetoden, siden man trenger mange respondenter for at man skal kunne trekke statistisk gyldige konklusjoner [16, s. 199]. Dette må eventuelt komme i en senere studie. Men som Robson skriver i [21, s. 271], så er intervjuer passende når man søker å være utforskende på et felt hvor det har vært gjort lite forskning fra før av, for eventuelt å kunne bane vei til mer grundige, kvantitative undersøkelser senere. Det er nettopp dette vi har gjort et forsøk på her.

### 11.5.3 Forskningsspørsmål 3

I forskningsspørsmål 3 var vi interesserte i å finne ut hva nytteverdien av øyesporing som del av systemutvikling er i dialogen med kunder og utviklere. Vi valgte her å benytte oss av caset vårt (kap. 5) slik at Amnesty ville komme i rollen som kunde, og jeg selv ville være en praktiker-forsker, altså en deltakende observatør som allerede har en reell rolle i caset.

Rollen min i dette caset var som en innleid konsulent med ekspertise på brukbarhet, hvor jeg så skulle presentere resultatene fra brukbarhetstesten jeg hadde gjort tidligere. Dette er jo en rolle hvor jeg er forholdsvis uerfaren, siden jeg aldri har opptrådt som brukbarhetseksperter tidligere. Men gjennom skolegang og studier har jeg lært og fått en del øvelse i hvordan man skal holde presentasjoner generelt, så på det området følte jeg meg sikker nok. Jeg følte også at møtet gikk ganske bra og at jeg fikk formidlet det jeg hadde tenkt, til tross for min uerfarenhet. Som praktiker-forsker har man ofte en utfordring i å skille mellom det som er spesielt og det som er naturlig av det man observerer, slik som utenforstående ville gjort, siden man selv er en del av settingen til vanlig [16, s. 211]. Her tror jeg at min uerfarenhet på området var med på å gjøre denne utfordringen mindre.

Oates sier at et case-studie blant annet karakteriseres ved at det foregår i dets naturlige setting, hvor forskeren forsøker å forstyrre denne så lite som mulig [16, s. 142]. Caset vil som oftest eksistere før forskeren kom inn i bildet, og høyst sannsynlig fortsette å eksistere etter studien er ferdig. Dette stemmer for vår del, prosjektet vi studerer er et virkelig systemutviklingsprosjekt og det gjorde at møtet vi skulle observere for å besvare forskningsspørsmålet vårt var realistisk nok. Jeg fikk også gode tilbakemeldinger fra deltakerne i etterkant av møtet. *“Dette var jo veldig nyttig for oss, da! [...] Så bra at vi fikk vite dette såpass tidlig, nå kan vi jo til og med fikse noen av feilene før vi fjerner Beta-stempelet vårt”*. Dette viser at møtet gjenspeilet en virkelig situasjon.

### 11.5.4 Forskningsspørsmål 4

For å besvare forskningsspørsmål 4, om hvilke praktiske og metodiske utfordringer man møter ved bruk av øyesporing i brukbarhetstesting, valgte vi også å gjøre et intervju (se 4.4.5). Som nevnt under 11.5.2 bør man være forsiktige med å trekke generelle konklusjoner fra slike intervjuer, siden man har så få respondenter. Men til forskjell fra resultatet for forskningsspørsmål 2, så brukte jeg her også mine egne erfaringer fra bruken av øyesporing i brukbarhetstesting i denne studien og tidligere studier, til å støtte opp under og utfylle resultatet fra intervjuet med Terje Røsand. Dette blir å betrakte som observasjon hvor jeg da er en praktiker-forsker, siden jeg faktisk var i rollen som en testleder under en

brukbarhetstest.

En ulempe med slike deltakende observasjoner hvor det som regel finnes bare én observatør, er at leseren må stole på at de resultatene som kommer fram er riktige, som Oates beskriver i [16, s. 211]. Man kan jo ikke garantere at andre ville observert akkurat det samme som du, rett og slett fordi man som mennesker er forskjellige. Men man kan gjøre visse grep som gjør resultatet mer troverdig. I vårt tilfelle gjorde vi jo videoopptak av brukbarhetstestene vi gjennomførte, og da kom selvfølgelig de praktiske problemene vi opplevde der også med på film. Verre er det med de utfordringene som oppstod utenfor hver av brukbarhetstestene, for eksempel under dataauthenting i etterkant.





## Konklusjon

---

Dette kapittelet inneholder konklusjonene vi kan slutte basert på det vi har funnet ut i denne rapporten (12.1) og jeg nevner også noe om hva som kan være interessant å forske videre på (12.2).

### 12.1 Konklusjon

Konklusjonen min her er et forsøk på å sammenfatte resultatet vi har funnet ut i denne studien på en noenlunde kortfattet måte. Jeg skal prøve å gi et svar på forskningsspørsmålene som er beskrevet i introduksjonskapittelet (1.2). Det vi i hovedsak lurer på er altså: *Har øyesporing nytteverdi som del av brukbarhetstesting i systemutvikling?* Eller for å sitere min veileder Dag Svanæs: *“Er det verdt bryet?”*

For å svare på dette vil jeg ta for meg ett og ett av forskningsspørsmålene jeg har jobbet med i denne studien. Mer utfyllende resultater og drøfting av disse finner du i kapittel 7-11.

1. *I hvilken grad bidrar øyesporing til å identifisere problemer?* På bakgrunn av brukbarhetstesten jeg gjorde for å finne svar på dette, kom jeg hovedsaklig fram til tre områder hvor øyesporing var til hjelp. For det første kan det hjelpe oss til å forstå hvorfor noen brukere ikke “oppdager” et viktig element i brukergrensesnittet. Det kan for eksempel være at noe annet stjeler oppmerksomheten til brukeren, og dette kan man se tydelig ved hjelp av øyesporing. Det andre jeg vil trekke fram er at data fra øyesporing kan bekrefte eller forsterke antagelser du har gjort deg på bakgrunn av vanlig testmetodikk. I vårt tilfelle hadde nettsiden vi testet to like menyer på sidene, og vi merket at brukerne benyttet nesten utelukkende bare den ene av dem. Vi hadde derfor en teori om at de ikke så den andre menyen, noe øyesporingen bekreftet. Til slutt kan øyesporing også hjelpe oss til å forstå brukerens mentale modell av for eksempel navigasjonsstrukturen på siden, altså hvor han forventer at ting skal være. Vi kan se hvilke alternativer han vurderer før han eventuelt klikker på et, og slik kan vi se om navigasjonsstrukturen på siden er logisk oppbygd i forhold til slik brukeren tenker.
2. *Hvilken nytteverdi ser utviklere som til daglig bruker øyesporing i brukbarhetstesting?* Her gjorde jeg et intervju med NetLife Research som bruker øyesporing mye i brukbarhetstesting. De opplever at øyesporing er et veldig slagkraftig kommunikasjonsverktøy overfor kunden når man kan vise øyesporingsvideoer og spille

av lydopptak fra testpersoner som bruker produktet deres. I tillegg synes de også øyesporing er et hendig hjelpeverktøy under selve testen, på den måten at testleder eller moderator kan følge med på hvor testpersonen ser hen og dermed stille oppfølgings spørsmål på passende tidspunkt. Når det gjelder problemidentifisering så tar det mye tid å gjøre en skikkelig analyse av øyesporingsdata, så de benytter seg stort sett av dette hvis de er i tvil om noe, og en sjelden gang oppdager de også nye problemer.

3. *Hva er nytteverdien av øyesporing som del av systemutvikling i dialogen med kunder og utviklere?* For å finne svar på dette hadde jeg et møte med Amnesty, som da blir kunde, hvor jeg presenterte resultatene fra brukbarhetstesten. Konklusjonene er først og fremst basert på erfaringene mine derfra. Det er tre ting jeg vil trekke fram her. For det første så opplevde jeg at når de tilstedeværende fikk se med egne øyne hvordan brukeren beveget øyene, forstod de problemet mye bedre. På den annen side var det et kraftig verktøy for meg som brukbarhetstester da det gav mer tyngde og troverdighet til argumentene mine når jeg kunne vise til øyesporingsdata. Men jeg vil også nevne at man må være bevisst hva man viser fram av øyesporingsdata, spesielt når det gjelder varmekart, siden det er veldig enkelt å lage noe som ser stilig ut, men som egentlig ikke forteller deg noe som helst.
4. *Hvilke praktiske og metodiske utfordringer møter man på?* Selv om teknikken i dag er kommet ganske langt, så er det fremdeles noen ting som ikke fungerer optimalt. Det kan hende at øyesporeren plutselig faller ut midt i en test slik at vi mister data fra resten av testen, eller programvaren forårsaker problemer på maskinen fordi den krever for mye datakraft. Det er heller ikke alle personer som egner seg som testdeltakere, siden øyesporeren kan ha vanskelig for å lese folk med briller for eksempel, eller at de har lett for å bevege seg utenfor området hvor den tar inn signal. En annen ting er at målingene kan være litt unøyaktige, og det passer kanskje ikke til alle formål.

For å oppsummere kort så kan altså øyesporing være til hjelp i problemidentifiseringprosessen til en viss grad. Men det jeg synes er mest oppsiktsvekkende her er dette med at det er et veldig nyttig i verktøy i kommunikasjonen mellom brukbarhetstestere og kunden/andre utviklere. Vi ser også at utviklere som bruker dette til vanlig ser en klar nytteverdi i dette. Samtidig må man være klar over at man vil støte på utfordringer her også, som i så mange andre felt.

## 12.2 Videre forskning

Som jeg nevnte i introduksjonskapittelet er ikke dette en oppgave hvor vi vurderer de økonomiske betraktningene av øyesporing i systemutvikling, men dette er absolutt et område som er mulig å forske videre på.

Det kan også være interessant å se nærmere på hvilken testmetodikk som er mest hensiktsmessig i en systemutviklingssituasjon hvor man ikke har ressurser til å gjøre grundig analyse av dataene. Som jeg skrev om i metodediskusjonen (under 11.5.1), så kan høytteningsteknikken påvirke øyesporingen på en måte som ikke er ønskelig. Det finnes flere metoder som kan brukes, men det er ikke sikkert alle er gode til industrimessig bruk.

# **Vedlegg**

---

---

## System Usability Scale

---

Nedenfor har jeg listet opp de påstandene som brukes på norsk når man skal beregne SUS-verdiene. Brukerne skal svare ved å velge hvor enig eller uenig de er på en skala fra 1 til 5, hvor 1 tilsvarer *Sterkt uenig* og 5 tilsvarer *Sterkt enig*. Påstandene er de samme som i [4] og oversatt av Dag Svanæs i 2006.

1. Jeg kunne tenke meg å bruke dette systemet ofte.
2. Jeg synes systemet var unødvendig komplisert.
3. Jeg synes systemet var lett å bruke.
4. Jeg tror jeg vil måtte trenge hjelp fra en person med teknisk kunnskap for å kunne bruke dette systemet.
5. Jeg syntes at de forskjellige delene av systemet hang godt sammen.
6. Jeg syntes det var for mye inkonsistens i systemet. (Det virket “ulogisk”)
7. Jeg vil anta at folk flest kan lære seg dette systemet veldig raskt.
8. Jeg synes systemet var veldig vanskelig å bruke.
9. Jeg følte meg sikker da jeg brukte systemet.
10. Jeg trenger å lære meg mye før jeg kan komme i gang med å bruke dette systemet på egen hånd.



## Pilotoppgaver

---

Følgende oppgaver brukte vi under pilottestingene beskrevet i 6.1.3. Oppgavene ble endret til de som vises i kapittel 6.4 før gjennomføringen av hovedtesten.

1. Se deg litt rundt på sidene for å bli kjent med de.
2. Finn bloggen til Amnesty.
3. Finn den viktigste aksjonen for Amnesty og gi din underskrift.
4. Finn ut hvordan du kan være en aktiv Amnesty-medlem på din arbeidsplass.
5. Finn en reportasje som interesserer deg og legg inn en kommentar.
6. Finn ut hvilke 3 måter Amnesty vanligvis jobber på.
7. Finn ut hva pengene som Amnesty Norge samler inn går til.
8. Finn ut når det neste arrangementet for Amnesty Norge er og når påmeldingsfristen er.





## Intervjurunde etter test

---

Etter at testdeltakeren hadde gjennomført oppgavene og svart på SUS-spørreskjemaet (se kapittel 6.3), stilte jeg noen spørsmål om hva de syntes om nettsiden og tok opp noen tema som kanskje ikke hadde kommet fram underveis i testen, men som var interessant for oss å vite. Jeg hadde lagt det opp som et semi-strukturert intervju, og spørsmålene jeg hadde forberedt er listet under.

1. Hva synes du om websiden generelt?
2. Hva synes du om designet? Farger, layout, osv.
3. Hva synes du om lesbarheten? Fonttype og størrelse?
4. Var det vanskelig å finne fram til det du lette etter?
5. Var det noen ord eller andre ting du ikke skjønnte meningen med? Noe du stusset over?
6. (Spesielle ting under testen som var vanskelig)



---

## SUS-resultater

---

Tabell D.1 viser resultatet hver av deltakerne endte opp med etter å ha svart på SUS-spørreskjemaet etter testen. Dette blir omhandlet nærmere i kapittel 7.1.1.

Testperson	SUS score
1	62,5
2	87,5
3	52,5
4	77,5
5	75
6	50
7	70
8	82,5
9	52,5
10	72,5
11	32,5
12	60
13	87,5
14	65
15	87,5
Gjennomsnitt	67,7
Standardavvik	16,2
Min	32,5
Maks	87,5

Tabell D.1: Tabellen viser resultatet fra SUS-skjemaene brukerne svarte på etter testen, inkludert gjennomsnittsverdi, standardavvik og minimum- og maksimumsverdi. Dette representerer satisfaction-faktoren for brukbarheten på nettsiden (se delkapittel 2.1.1), altså hvor fornøyde de er med siden etter å ha brukt den.



---

## NetLife Research - intervju

---

Her følger en transskripsjon av intervjuet jeg hadde med Thor Fredrik Eie som er Senior Usability Specialist i NetLife Research. Jeg bruker dette som basis for resultatene i kapittel 8.

*Bruker dere eyetracking på alle tester?*

Nei, vi gjør ikke det. Det er et slags tillegg som kundene våre kjøper. Men jeg vil vel anslå at vi bruker det i 70-80% av testene, så vi bruker det ganske jevnlig.

*Er det kundene som spør etter det eller er det dere som foreslår det for kundene?*

Når vi sender tilbud til kundene våre, så sender vi også med tilbud for eyetracking i tillegg som da er ca 5% påslag i prisen. Men vi har også en del kunder som har sett at vi har demonstrert eyetracking før eller de har lest litt om det på forhånd, og da spør de konkret etter eyetracking. Så det er egentlig bare et tillegg i forhold til en vanlig brukertest.

*Hva ser dere etter når dere analyserer eyetracking?*

Vi bruker eyetracking først og fremst som et støtteverktøy i forhold til vanlig brukertestingsmetodikk. Så rent metodisk pleier vi å gjøre eyetracking ved først å la testpersonen løse noen oppgaver helt uten innvirkning fra oss, altså at vi blander oss ikke inn i det hele tatt. Jeg bruker faktisk å si at jeg har glemt å hente kaffe, slik at jeg ut av rommet sånn at folk får sitte helt i ro og fred her inne.

Så kjører vi en del 2 som er mer vanlig testmetodikk med think-aloud. Og da bruker vi eyetracking-dataene som vi har samlet inn i den første delen av testen til å bekrefte eller forsterke konklusjonene som vi har sett i den vanlige think-aloud-delen av testen. Vi har en metodisk utfordring i og med at vi vanligvis bruker bare 8 testpersoner, og det er for få til egentlig å si noe statistisk om det. Så vi er veldig forsiktige med å trekke konklusjoner kun basert på eyetrackingen, men vi bruker det heller til å forsterke ting som vi allerede har sett. F.eks. hvis folk har slitt med navigasjonen på en side, så kan vi vise frem en artikkelside der folk ikke har sett på noe av navigasjonsbiten på siden, eller kanskje hvis det er noen store bilder som har "slått i hjel" resten av siden [fordi det tar for mye oppmerksomhet].

*Hvor lang tid bruker dere på å analysere resultatet av en eyetrackingstest?*

Vi bruker kanskje rundt et dagsverk i tillegg til det som vi bruker på en vanlig brukbarhetstest.

*Ser dere noe på filmene i etterkant av testene?*

Lite. Det hender at vi tar ut klipp fra filmene for å ha med oss når vi skal presentere resultater. For det hender jo ofte at testpersonene kommer med et fantastisk sitat kanskje 5 minutter ut i testen som med en gang setter fingeren på det store problemet. Så da bruker vi ofte å ta ut noen sånne små videoklipp på et halvt til ett minutt varighet hvor det er et bra sitat, og kunden kan da både se øyebevegelsene og høre stemmen til testpersonen, noe som er ganske interessant. Det er et ganske slagkraftig verktøy når man skal presentere for kunder som kanskje ikke har vært med på testen og dermed ikke har sett problemet selv, det er med andre ord en veldig effektiv måte å kommunisere med kundene på.

*Dere bruker altså ikke filmene for å finne problemer?*

Det er lite. Vi har brukt det noe, men da har det mest gått på navigasjonsprinsipper og andre ting som har feilet helt for eksempel.

*Er det fordi det er veldig tidkrevende å analysere på den måten?*

Ja, det tar veldig mye tid hvis man skulle gjort det skikkelig. Vi bruker ca et dagsverk eyetrackingsdataene, og da er det medregnet å finne fram til og eksportere de riktige klippene. Men vi bruker ikke noe tid på å analysere videoene videre, det er mest for å finne fram til stedene hvor vi allerede vet at noen sa noe bra for eksempel.

*Hva ser dere på som den største nytteverdien av eyetracking? Er det det å kunne presentere det for kunden?*

Nei, det er flere nivåer her. Det ene er at i den første delen, da brukeren sitter for seg selv som om han var hjemme, så kan vi sitte på observasjonsrommet og se på øyebevegelsene, og siden vi styrer såpass lite kan vi i større grad anta hva logikken til brukeren er, altså hva det er han leter etter og hvordan han vil bruke nettstedet. Vi kan også se f.eks. hvis det er noen primære navigasjonsproblemer, at navigasjonen ikke blir brukt, så kan det hende at folk ikke har sett det i det hele tatt.

En annen ting som er veldig fint med eyetracking er at for meg som moderator som kjører en brukertest, så kan jeg sitte ved siden av og se hvor testpersonen ser hen, og det er med på å hjelpe meg til å time spørsmålene bedre. Et problem man av og til kommer bort i som moderator er jo at man begynner å stille oppfølgingsspørsmål før testpersonen faktisk har fått lest en tekst eller at de fremdeles er på jakt etter og er aktive for å løse en problemstilling, og så spenner man litt bein på de fordi man stiller et oppfølgingsspørsmål før de er ferdige. Så beste måten er da å sitte å se på eyetrackingen mens bruker jobber for seg selv, og så bruker du dette til å stille spørsmålene på et bedre tidspunkt, f.eks. når man ser at nå begynner testpersonen å se litt over alt uten noe spesielt mål med det han holder på med. Så akkurat testmessig er det et veldig fint verktøy å ha.

*Merker dere noen ulemper når dere har med eyetracking å gjøre?*

Ja, det er jo ikke alle personer som lar seg teste med eyetracking så godt. Så i løpet av en brukertest så har vi kanskje 1-2 testpersoner der eyetrackingen fungerer dårlig. Det kan være mange årsaker til det, noen har briller som kaster mye skygge i forhold til sensorene på eyetrackeren, noen sitter veldig urolig slik at de ikke holder seg innenfor bevegelsessonen til eyetrackeren, andre sitter å støtter seg veldig mye. Og da er det jo også et spørsmål om hvor mye man skal blande seg inn i det og si at "nei, du kan ikke

---

sitte sånn og sånn”, for da vil det bli dårlig stemning når man sitter og plukker på måten testpersonen sitter på.

Vi har også hatt litt utfordringer teknisk, f.eks. at det har hengt seg, og det har vært litt ustabil sånn. Utviklingen går jo den riktige veien, men av og til skjer det jo ting som man ikke kan forklare av den tekniske delen. Vinduer som åpner seg på en pussig måte eller at eyetrackinga rett og slett krasjer underveis, og testpersonen merker jo ikke noe til akkurat det, men vi vil jo da mangle data fra den testen. Så i forhold til det har vi hatt litt utfordringer.

I tillegg så er det ikke alle som er så glade i å kjøre eyetrackingstester fordi det er så mye teknisk som kan skjære seg og så har man da et helt rom med kunder som betaler for å sitte å se på. Alle her kjører stort sett eyetrackingstester, men noen føler seg litt usikre på akkurat det.

Vi er jo de eneste som selger tester med eyetracking [i Norge], men jeg vet også at finn.no har en eyetracker, skattetaten kjøpte en i fjor, og statistisk sentralbyrå har også en som de bruker. Disse bruker jo eyetracking for evaluering av nettsteder. Men det er jo et veldig nyttig verktøy, og vi har jo foreløpig bare eyetracker på det ene laboratoriet vårt, men vi har snakket om å skaffe oss en på det andre også, for vi har sett at det er noe kundene våre gjerne vil ha.

Noe som man må passe på er jo at det er utrolig lett å ta ut heat maps som ser veldig kule ut, men som forteller deg egentlig ingenting, så vi er ganske forsiktige med. Heatmaps basert på 8 personer er såpass dårlig statistisk sett, så det er ikke noe man kan slå i bordet med. Men vi bruker heller en del gazeplots og drar ut f.eks. hva er de første elementene folk ser på, for det kan man vise på individbasis. Men i forhold til samlede resultater fra forskjellige testpersoner, så er vi veldig forsiktige med å konkludere for mye ut i fra det.

Ofte så har vi jo hørt eller sett ting brukeren ikke har fått til, og da finner vi gjerne ting i eyetrackingdataene som kan støtte det og i alle fall gi noen hypoteser om hvorfor det er sånn.

*Så det er veldig sjelden dere finner noen nye problemer basert på eyetrackingsdata?*

Det er noen tilfeller hvor vi har funnet ganske kule ting egentlig. For eksempel noe som vi kaller for visuelle vollgraver, som er at store bilder eller store grafiske elementer i et webdesign faktisk dreper det innholdet som ligger rundt, ett blikkfang rett og slett. Brukeren trenger ikke nødvendigvis se så mye på det heller, men det tar bare så mye plass. Et eksempel er fra de gamle Telenor-sidene, der de hadde mange fine, redaksjonelle bilder. Globalnavigasjonen lå over bildet, mens innholdet lå under bildet, og da så vi at globalnavigasjonen ble fullstendig slått i hjel av det store bildet sånn at alt fokuset til folk lå under det bildet, fordi de tenkte at åja, det er bare her det er noe innhold, alt det andre er bare tull. Vi gjorde også en test av Aschehaug sin sider, hvor vi så at globalnavigasjonen på toppen var i en veldig kraftig, oransje farge, mens søkefeltet lå over dette oransje. Det vi oppdaget da, var at brukerne så ikke søkefeltet, det kom ikke fram rett og slett. Og da fikk vi en teori om at den oransje fargen tok så mye plass, at det var vanskelig å finne søkefeltet. Så vi har hatt noen tilfeller der vi har sett ting på eyetrackingen som ikke har kommet frem på andre måter, og det er litt interessant. Men det blir jo også på individnivå.

*Men hovedresultatet er stort sett det som bygger opp under ting dere har funnet ellers?*

Ja, det er det. Vi har forsåvidt prøvd å kjøre rene eyetrackingtester, men for det først så trenger man da så og så mange testpersoner, og da kommer all den administrasjonsbiten med å rekruttere testpersoner inn i bildet også, så vi har ikke fått solgt så mange av de. Men jeg er litt i tvil om hvor mye den rene eyetrackingen vil tilføre av informasjon, for slik som jeg ser det, så er det det at du faktisk har eyetrackingdata og det du har sett brukeren gjøre og si i testen som forteller deg mest om problemstillingene. Hvis du får et heatmap for eksempel, så er det så utrolig mange faktorer som har noe å si for hvordan denne heatmapen vil se ut, det trenger ikke være at menyen er dårlig utformet. Så jeg tror det er vanskelig å kjøre brukertesting med ren eyetracking, og få mange gode resultater.

Men eyetracking brukes jo også mye til reklametesting, og der er det mer interessant å vite hvor mange som f.eks. så logoen på sjampoflasken. Der er det helt andre ting som er viktig. Men jeg mener at man går glipp av de mer kognitive tingene når man gjør brukertesting med kun eyetracking.

Det som er med eyetracking er jo at man plutselig har fått teknologien god nok til å gjøre brukbarhetstesting med dette, og mange veiver litt med armer og bein for å finne den rette metoden å bruke dette på. Jeg var på en konferanse i 2007, og mange snakket om da var at man kjører en retrospektiv think-aloud, dvs at man lar brukeren gjøre oppgavene først uten noe innblanding fra oss, og så spiller du deretter av eyetrackingvideoen for testpersonen og lar han kommentere underveis. Vi har egentlig ikke kjørt så mye av det av forskjellige årsaker. Sånn som vi ser det så er det egentlig i kombinasjonen av hva brukeren gjør og sier, og eyetrackingdataene at gullet kommer fram.



---

## Presentasjon for Amnesty

---

Følgende sider viser lysarkene jeg brukte under presentasjonen jeg hadde for Amnesty angående resultatene fra brukbarhetstesten. Mer om dette møtet i kapittel 9.

Resultater fra brukbarhetstest  
av ny webside til

Amnesty Norge

**Elin Standal**  
Masterstudent ved NTNU  
Institutt for datateknikk

**Terje Rødsand**  
Staff Engineer ved NTNU  
Institutt for datateknikk

# Agenda

- Generelt om brukbarhet
- Hvordan gjorde vi testen?
- Litt om testdeltakerne
- Statistiske resultater
- Gjennomgang av de ulike problemområdene

2

# Brukbarhet

God brukbarhet handler om at brukerne får gjort det de skal på en effektiv og tilfredsstillende måte.

*Effectiveness*  
*Efficiency*  
*Satisfaction*



3

# Brukbarhetstesting

- Hvem er nettsiden for?
  - Medlemmer/Ikke-medlemmer? Journalister? Pensjonister?
  - Hva kjennetegner disse menneskene i bruk av data?
- Hva skal de gjøre der?
  - Hvorfor besøker folk siden?
  - Finne relevante oppgaver.
- Hvor befinner de seg når de besøker den?
  - Testlokalet bør være så likt en virkelig situasjon som mulig.

4

# Gjennomføring av brukbarhetstesten

## Brukeren:

- Utforsket nettsiden på egen hånd.
- Gjorde 8 spesifikke oppgaver
- Svarte på et standard spørreskjema om brukbarheten til nettsiden (SUS).
- Svarte på spørsmål om ulike aspekter ved nettsiden (f.eks. lesbarhet, layout, farger, osv.)

5

# Oppgavene

1. Se deg litt rundt på sidene for å bli litt kjent med de.
2. Finn bloggen til Amnesty.
3. Signér aksjonen om forsvinningen av Masood Ahmed Janjua.
4. Hvor klikker du for å melde deg inn i Amnesty?
5. Finn en reportasje som kan interessere deg og legg inn en kommentar.
6. Finn ut hvilke 3 måter Amnesty vanligvis jobber på for å oppnå sine mål.
7. Finn ut hva pengene som Amnesty Norge samler inn går til.
8. Finn Amnestys logo til nedlasting.
9. Finn ut hvordan du kan være en aktivist for Amnesty gjennom bruk av sosiale medier.

6

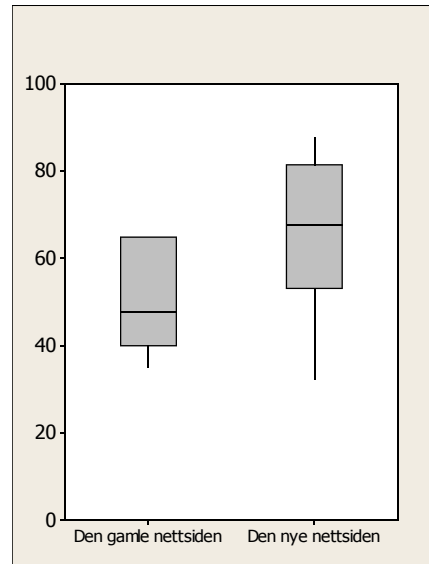
# Testdeltakerne

- 10 kvinner og 5 menn
- Alle brukte internett hver dag
- Kjennskap til Amnesty:
  - 9 god eller meget god
  - 6 middels, lite eller ingenting
- Besøkt den gamle hjemmesiden til Amnesty:
  - 4 hver dag eller ofte
  - 11 sjelden eller aldri

7

# SUS – System Usability Scale

- Hvor tilfredse var testdeltakerne med nettsiden?
- På en skala fra 0-100 fikk den nye nettsiden et gjennomsnitt på 68.
- Til sammenligning fikk den gamle siden et gjennomsnitt på 49.



8

## Effectiveness

- Klarte brukerne å gjennomføre oppgavene riktig?
- 7 av 15 (47%) klarte ikke å fullføre minst én av oppgavene.

Bruker	Fullførte oppgaver
1	75%
2	88%
3	100%
4	75%
5	63%
6	63%
7	100%
8	100%
9	100%
10	100%
11	75%
12	75%
13	100%
14	100%
15	100%
Gjennomsnitt	88%

9

# Problemene

Oppgave	Brukere som fullførte korrekt
2	100%
3	60%
4	100%
5	100%
6	100%
7	67%
8	87%
9	87%

- Problemoppgaver:
  3. Signering av aksjon
  7. Finne ut hva pengene går til.
  8. Laste ned logo.
  9. Finne artikkel om å være aktivist gjennom sosiale medier.
- Problemområder som ikke vises igjen i tabellen ble også identifisert.

10

## Signering av aksjon

**Hva:** Brukerne signerte i kommentarfeltet.

**Hvorfor:**

- Fra tidligere var de vant til å signere under artikkelen.
- Tenkte ikke at det var sammenheng mellom artikkelen og bildet øverst.
- Manglet overskrift til kommentarfeltet.



11

# Linker

**Hva:** Folk finner ikke linker.

**Hvorfor:**

- Ser ikke ut som linker.
- Inkonsistens i hvordan de ser ut.



<b>HVEM ER VI</b>	<b>Aksjoner</b> Amnesty International aksjonerer hver dag for å stoppe rettferdige rettssaker og protestere mot diskriminering i verden. Her er oversikten over de aksjonene som pågår.
<b>HVORDAN JOBBER VI</b>	<b>Temaer Amnesty fokuserer på i 2010</b> <b>Fattigdom</b> 923 millioner mennesker er kronisk underernærte. Hver fattigdom handler ikke om penger, men om at de som menneskerettighetsbrudd. Fattigdom er verdens mest løst løsning er ikke mer penger, men helse- og utdanningskvinnens posisjon. Amnesty mener investering i menneskelig sikkerhet og menneskerettigheter.
<b>HVA JOBBER VI MED</b>	<b>Sikkerhet og menneskerettigheter</b> Antiterror-tiltak i mange land bryter menneskerettighetsdommer. Eller de stilles for militære domstoler som uten fangeleiren er fremdeles ikke stengt. Mennesker spert menneskerettighetene i en rekke land som ledd i å bekjempe mot terror skal respektere menneskerettighetene.
<b>AMNESTY DER DU BOR</b>	ANDRE TEMA

12

# Audiovisuelt arkiv

**Hva:** Problemer med å finne logo til nedlasting.

*"Audiovisuelt, er ikke det lyd?"*

**Hvorfor:**

- Audiovisuelt forbindes med lyd eller video, ikke bilder.
- Fremmedord

*"En logo er for såvidt visuell... men den er jo ikke audio!"*

*"Audiovisuelt arkiv er litt for fancy ord for meg"*

13

# Søking

**Hva:** Avansert søk var vanskelig å forstå og bruke.

**Hvorfor:**

- Ingen forklaring til skjemaet.
- Filtrering av søket fungerte ikke.

14

# Menyer

**Hva:** Forvirrende og unødvendig med to undermenyer.

**Hvorfor:**

- Av og til er undermenyene like, av og til ikke.
- To like menyer tar opp mye plass på siden.
- Undermenyen på nivå 3 er lite synlig.

15



# Mindre problemer

## Forsiden:

- Bloggen
- "Bildekarusellen"
- Søkelinjen har bare blokkbokstaver og lite synlig skrift.
- Søkeknappen gir ikke feedback som de andre fanene.



16

# Mindre problemer

## Aksjon:

- "Aksjon"-begrepet ikke helt selvforklarende.
- Kobling mellom bildet og "Aksjon i fokus"-fanen.
- Dårlig lesbarhet på listene når musen holdes over.
- Vanskelig å finne eldre artikler.
- Fargekoder i aksjonskalenderen.
- "Gjør mer"-begrepet ikke helt selvforklarende.
- "Infobokser" nederst mangler under Aksjonskalender og Gjør mer.

17

# Mindre problemer

## **Støtt oss:**

- Folk forventet å finne informasjon om hvordan pengene de f.eks. gir i gave blir brukt.
- Unødvendig og lite pent med så store faner på høyremenyen.
- Noen forventet å finne informasjon om hvordan man kan være aktiv som medlem her.

## **Om Amnesty:**

- Inkonsekvens mellom linker og undermeny på nivå 3.

18

# Mindre problemer

## **Ressurser – Audiovisuelt arkiv**

- Unaturlig at portretter av de ansatte vises først når Amnesty er en veldedig organisasjon.
- "Lightbox, det ville jeg sannsynligvis aldri klikket på siden jeg ikke vet hva det er?"
- Unødvendig med "like" logoer under hverandre.
- Hva med rettigheter til bruk av bildene?





19

# Mindre problemer

## Artikler

- Rar plassering og utforming på POST-knappen for å legge inn kommentar.
- Printer-ikonet nederst ser ikke ut som en printer, og har heller ingen annen forklaring til hva det er eller gjør.


Den palestinske regjeringen har også etablert en etterforskningsgruppe bestående av seks personer fra etterretningstjenesten, som skal forsøke å finne Masood Janjua.

   Facebook  Nettby

**Navn: \***  
Anonymous

**Epost: \***

The content of this field is kept private and will not be shown publicly.

Legg inn kommentar her 

BRUKERKOMMENTARER: 0 kommentarer

20

Takk for oppmerksomheten!

21



## Episode fra møtet med Amnesty

---

Under følger en transkripsjon fra en episode hvor øyesporingen hadde noe å si for diskusjonen rundt ett av problemene vi hadde funnet. Det var 5 personer tilstede, foruten meg selv. Episoden blir tatt fra i rapporten i kapittel 9.2.

### Videor for problemet med signering av aksjon

**Elin:** Jeg skal vise to filmer der vi følger øyene deres når de skal signere en aksjon. Det som var feilen med denne oppgaven, var at de signerte i kommentarfeltet av en eller annen grunn. Det vi tror er grunnen, er at på den gamle siden, så signerte de alltid nederst etter at de hadde lest artikkelen, og det stod heller ingen overskrift ved kommentarfeltet som sa “*Legg til en kommentar*”, noe det gjorde på for eksempel nyheter.

Jeg skal også vise sammenhengen mellom bildet og artikkelen. For det virket som om folk hadde litt problemer med dette, at de ikke koblet at artikkelen hørte til bildet [øverst på siden].

**Person 1:** Okay...

**Elin:** [henter fram den første videoen]

Denne videoen er ganske kort. Den røde prikken viser hvor testpersonen fokuserer. Vi ser veldig fort at han scroller raskt nedover på siden for å lese artikkelen.

Vi ser at han registrerer at det står “*Navn*” og “*Epost*”. Han er kanskje litt i tvil om hva kommentarfeltet er, men han finner ingen annen plass å fylle ut dette.

**Person 2:** Kan du spille den forfra en gang til? Det gikk så fort. Hva var det han gjorde aller først?

**Elin:** Ja, det kan jeg gjøre.

**Person 1:** Nå prøver han å signere en aksjon, ikke sant?

**Elin:** Ja, nå skal han signere en aksjon. Og så har han søkt etter denne, for han fant den ikke på en annen måte, men han fant den slik.

Vi ser at han ser såvidt på bildet, men han ser ikke bort på høyremargen. Og så scroller han nedover, for han ser at det er *der* artikkelen begynner, på en måte.

**Person 3:** Og streifer rundt her ja... [da øynene til testpersonen fokuserer på kommentarfeltet] “*Er det her jeg skal signere, eller er det ikke?*” Dette her var interessant altså!

**Person 1:** Ja, veldig! Han bruker *lang* tid på det også!

**Person 2:** Han går altså *ikke* videre, han bare fortsetter [å streife rundt kommentarfeltet]

**Person 3:** [Da filmen var ferdig] Var dette en som fullførte oppgaven?

**Elin:** Nei, det var det ikke. Det var en som ikke klarte oppgaven. Etter at han hadde gjort dette her, så sa jeg til ham at “Det du *egentlig* skulle gjøre, det var å signere aksjonen”. Da svarte han: “Åja! Det var jo det jeg trodde jeg hadde gjort!”

Da lette han litt videre og fant til slutt riktig sted oppe til høyre, men det blir ikke registrert som fullført oppgave, siden han jo trodde han hadde gjort oppgaven tidligere og han fikk *hjelp* til gjøre den riktig senere. Nå skal jeg vise en video til. Den er litt lenger enn den forrige.

**Person 1:** Dette er samme oppgave, eller?

**Elin:** Ja, dette er samme oppgave. [Starter filmen]  
Hun fant den gjennom å gå inn på *Aksjon*.

**Person 2:** Men, lå ikke aksjonen på startsidene?

**Elin:** Jo, men det var veldig få som fant den på denne måten. De måtte ofte søke på navnet [til den aksjonen gjaldt].

Vi ser her at hun også scrollet veldig fort nedover. Hun begynner å se på overskriften og leser ingressen. Så går hun litt lenger ned, og opp igjen. [Testpersonen går opp og ned flere ganger på siden]

**Person 2:** Oi, oi, oi...

**Elin:** Vi ser at hun går aldri *helt* opp slik at hun ser alt.

**Person 1 og 2:** Nei...

**Elin:** Så det virker som om de tenker at artikkelen begynner ved overskriften og går nedover. Det som er over er noe reklame eller en kampanje eller noe sånt.

**Person 1:** Ja [mens hun nikker].

**Person 3:** Her skriver hun inn navnet sitt, ja...

**Elin:** Ja, og her var hun litt usikker, hun skjønnte ikke hvorfor hun skulle legge inn en kommentar, hun skulle jo bare signere.

---

**Person 1:** Ja.

**Elin:** Hun var på en måte ganske sikker på at dette var et kommentarfelt, men hun skjønnte ikke hvor ellers hun skulle signere.

**Person 3:** Men altså, tenkte hun at navn og epost var en signaturdel og at kommentar var noe annet?

**Elin:** Altså, hun var veldig usikker på dette. “*Signerer jeg nå, eller gjør jeg noe annet?*” For hun skulle jo bare skrive navn og epost, det var det hun tenkte. Og så skulle hun *Sende* den, ikke *poste* den.

**Person 2:** Men altså, det betyr jo at de forholder seg ikke til bildet i det hele tatt.

**Elin:** Nei, ikke det øverste bildet i alle fall.

**Person 4:** [til person 3] Alle de *Legg inn kommentar her*-innleggene, det er kanskje nettopp dette de gjør.

**Person 3:** Ja, det var det jeg tenkte på jeg også. For det er jo helt naturlig at man ikke fyller inn noen kommentar da, man bare legger inn navnet og eposten sin og trykker på knappen. That’s it! Og da ligger den default-teksten der som et innlegg.

**Elin:** Ja.

Dette var det mest alvorlige problemet vi fant da.

**Person 2:** Det er jo skikkelig alvorlig også! Det er jo å signere vi virkelig vil at folk skal gjøre!

**Elin:** Dere ser her også at hun går flere ganger lenger opp for å sjekke om det er noe annet sted hun kan signere. Men hun går aldri *helt* opp.

**Person 2:** Jeg skjønner veldig godt at folk blir forvirret, for det er jo faktisk ikke noen pekepinn på hvor de kan signere [lenger nede på siden].

Så folk forholder seg simpelthen ikke til bilder altså, som del av innholdet?

**Person 3:** Ja, altså, det er jo hvert fall ikke naturlig å forholde seg til bilder når man skal *gjøre* noen ting. Det skjønner jeg jo! Men det er veldig overraskende at de bare feier rett forbi det og ned på teksten altså! Det har jo nesten ikke vært *ett* punkt [en fiksering] oppe på bildet.

**Elin:** Nei, det var bare helt i starten.

Bildet er jo det samme som var på forsiden, så jeg tror at de bare registrerte at det var samme bildet, og så tenkte at det var samme kampanje eller noe. Så da gikk de heller direkte ned på artikkelen.

**Person 3:** Ja. Jeg fant jo en kommentar i dag senest hvor det stod “*Signed!*” i kommentarfeltet. Det er jo en helt tydelig kobling.

Ja, nå skjønner vi alt så meget bedre!

**Person 1:** Ja, dette er jo *veldig* nyttig for oss da!

**Person 3:** Ja, veldig!

**Person 1:** Vi ville jo ikke drømt om det, i hvert ikke jeg, at dette var tilfellet. Det er helt utrolig!

**Person 3:** Nei, det hadde vi ikke tenkt på i det hele tatt.



## Intervju med Terje Røsand

---

Her er transkripsjonen av intervjuet jeg gjorde med Terje Røsand om hvilke erfaringer han hadde med den praktiske og metodiske delen av øyesporing. Dette er omhandlet nærmere i kapittel 10.

*Hvordan var det å sette opp utstyret i starten?*

Jeg synes jo at det var veldig enkelt. Bruksanvisningen var grei. Første gangen jeg koblet det opp, så brukte jeg ikke så veldig lang tid på å måle nøyaktig, og egentlig så vet jeg ikke hvor nøyaktig man trenger å være, men jeg går ut at det er et visst slingringsmonn i og med at det var med et målebånd på 3 meter, når man egentlig bare trengte det til å måle korte avstander.

Så jeg synes at det var kjempelett å komme i gang med, og det gjelder jo under bruk også. Og det med kalibreringen synes jeg går såpass fort at jeg kan ikke skjønne at det er et problem for brukeren. Det som *kan* forstyrre brukeren til en viss grad, det er jo det at man må justere inn stolen og sånn. Og det passer rett og slett ikke for alle, for man må ha en viss avstand [fra eyetrackeren], 50-80 cm eller noe. Det har vi jo sett her, at når de er ferdige med å justere inn sittehøyde og avstand fra skjermen, så vil de lene seg framover når de begynner på testen, og så mister vi data. Men generelt så kan vi jo si at eyetrackeren gjør ikke så mye ut av seg i forhold til at brukeren blir forstyrret av at det er noe teknisk utstyr her. Men i forhold til det med oppsettet så synes jo jeg at det var enkelt, men jeg jobber jo med tekniske ting da.

*Hva tenker du om programvaren, Tobii Studio, har du brukt det noe særlig?*

Jeg har jo brukt det mest for å bare teste det ut, men jeg synes at det er veldig lett å bruke. Det virker som at det er et ganske kraftig verktøy, men det er jo også noen ting som er problematisk. For eksempel når brukeren besøker samme side flere ganger i løpet av en test, men er ute etter forskjellige ting, så er det å få skille dette en utfordring.

Ellers så har jeg jo inntrykk av at de legger inn noe funksjonalitet som man egentlig ikke får brukt så mye, for eksempel den statistikk-biten. Det er godt mulig at det er mange som bruker akkurat det, men jeg tenker at dette er sånn som man legger inn bare fordi det er gratis”og kjekt å ha siden teknologien allerede er tilstede. For det med statistikk er jo veldig enkelt å legge inn, når man først har alle dataene tilgjengelig, da skal det ikke så mye ekstra arbeid til for å legge det til. Men som sagt, det kan jo hende at det er noen som bruker dette mye.

*Har du sett noe på for eksempel Bee Swarm-funksjonaliteten?*

Nei, for det må vi ha egen lisens på, og det er vanskelig å si noe om hvor nyttig det er uten å ha testet det. Men det *ser* jo ut som om det kan være interessant da. Nå husker jeg ikke akkurat hva det går ut på, men jeg har jo lest om det i manualen [til Tobii Studio].

*Vi hadde jo et problem med keyboardet som ikke fungerte da vi testet tidligere i år. Fant du ut hva som var grunnen til dette?*

Ja, Tobii Studio i seg selv krever ganske mye maskinressurser. Og i tillegg bruker vi Xvid som er en videokoder for å lagre videoen, og den er satt opp med ganske høy frame rate sånn at videoen får god kvalitet og det krever mer ressurser av maskina. For den codecen som ligger innebygd gjør at filmen blir litt mer stakkato, men det krever ikke så mye ressurser. Det Tobii anbefaler som minstekrav til maskinvare var litt i høyeste laget egentlig, så vi kommer til å bygge en litt kraftigere maskin for å kunne bruke det.

*Så det er hovedsaklig maskinvare som har vært problemet her da?*

Ja, for oss så har det vært det. Vi fikk en forsinkelse som gjorde at brukerne opplevde at det var noe feil med tastaturet, og det var det jo ikke. Så det får vi jo rettet opp etter hvert.

Vi kjøpte jo den eyetrackeren som er helt uavhengig av skjerm. Man kan jo også kjøpe en som er innebygd i skjerm, men vi kjøpte denne for å få en litt mer fleksibel løsning. Så i prinsippet så kan vi kjøre eyetracking på alt, store skjermer, prosjektorer, osv.

*Har du prøvd det?*

Jeg har prøvd litt på en stor skjerm, men poenget mitt er at det er visse begrensinger for eyetrackeren. For det første så er det jo avstanden fra eyetrackeren til øyet til testpersonen som må være innenfor 50-80 cm. I tillegg så må vinkelen fra eyetrackeren ut til alle kanter være mindre enn 35 grader. Dette er jo tross alt litt, men avstanden skaper et problem. Jeg prøvde nemlig å koble det opp til en touch-skjerm, og det vil jo si at du må kunne stå så nær skjermen at du når bort med fingeren, og det var egentlig ikkje forenelig med avstanden til eyetrackeren. Men hvis man setter opp eyetrackeren på et stativ og brukeren ikke trenger å ta på skjermen, så kan man jo gå litt lenger unna, og da vil det gå fint vil jeg tro.

Det skal jo i prinsippet gå an å bruke dette på objekter også, det trenger ikke bare være en dataskjerm. Men da må det kalibreres med en plate som er tegnet, og et kamera som er festet over brukeren som filmer det som skjer. Det har jeg også prøvd, og det fungerer. Det som er svakheten med det, er at eyetrackeren forholder seg til et *plan* i rommet. Så hvis man har 3-dimensjonale objekter og hvis det da er langt fra den forreste delen til den bakerste delen av objektet, så vil det bli en feil. Men hvis det er noenlunde innen en flate, så vil det sikkert kunne fungere ganske greit. Men jeg kan ikke se for meg at dette er noe vi kommer til å bruke noe særlig.

Jeg tenkte også på dette med nøyaktigheten av målingene. Jeg har ikke jo ikke gjort noe systematisk testing av nøyaktigheten, men ved å kjøre den *Verify*-funksjonen [under kalibreringen], så vil du jo få et brukbart innblikk i hvor nøyaktig den er.

*Men akkurat det får du vel ikke tatt opp på videoen, når du bruker Verify-funksjonen? [Tobiis innebygde videooptaker starter etter kalibreringen]*

Nei, du får det jo ikke med på videoen, men det vil jo likevel gi en pekepinn på hvor nøyaktig den er. Det jeg kan si generelt om det forresten, er at den er mest nøyaktig midt

---

i skjermbildet, og så vil unøyaktigheten bli større til lenger ut mot kanten man kommer. Jeg kommer ikke på i farten om den er mest nøyaktig horisontalt eller vertikalt, det tror jeg varierer litt. Men jeg har også sett at hvis den bommer litt, og man kjører en ny kalibrering på samme bruker, så kan det bli bedre. Du har jo også mulighet til å legge inn 9 kalibreringspunkter [i stedet for 5]. Det er mulig at det kan gjøre ting bedre, men da har du også ulempen med at det tar lenger tid for brukeren [å kalibrere].

Jeg har lest en hovedoppgave fra Dragvoll, fra språkvitenskap eller noe sånt, der det ble brukt eyetracker. Hun hadde dysletikere og ikke-dysletikere som skulle lese en tekst, og hun kommenterte at hun klarte ikke alltid å se hvilken linje brukeren var på, men at hun fant det ut ved hjelp av telling i stedet. Så det må jo vi også ha i bakhodet, at det kan bomme litt.

*Ja, det har vi jo litt sett på de resultatene vi hadde, at det er ikke alltid det ser ut som de har sett direkte på menylinjen, men du ser at det er der de har trykket.*

Ja, du skjønner automatisk at det er der han har lest. Så til usability-ting så vil jeg tro at den er nøyaktig nok. Da jeg begynte å bruke denne, for cirka et år siden, så vart jeg egentlig overrasket over hvor bra det fungerte, det var *ganske* nøyaktig. Det er jo en viss boks hodet kan være innenfor, og den er egentlig nokså stor. Så i forhold til det at den ikke skal skape noe *problem* for brukeren, så synes jeg at det fungerer veldig bra. For, jo mer slike tekniske ting vi dytter på brukeren, og som er synlig for brukeren, jo mer vil det forstyrre, tror jeg. Til testing brukte vi tidligere en mikrofon som vi festet på brukeren, men nå har vi gått over til å bruke takmikrofoner. Selv om lyden var veldig mye bedre på den første siden vi slipper mye støy, så anser vi det som viktigere at brukeren slipper å bli betenkt med å ha på seg mikrofonen. Eyetrackeren oppfatter jeg som veldig bra i så henseende.



---

## Referanser

---

- [1] Amnesty. Hva er amnesty international, 2007. url = <http://www.amnesty.no/web.nsf/pages/0E5D37D1628A914BC1256F0400423B39> [Accessed 04. feb 2010].
- [2] Aaron Bangor, Philip Kortum, and James Miller. Determining what individual sus scores mean: Adding an adjective rating scale. *Journal of usability studies*, 4(3):114–123, 2009.
- [3] Agnieszka Bojko. Informative or misleading? heatmaps deconstructed. In J. A. Jacko, editor, *Human-Computer Interaction. New Trends*, volume Volume 5610/2009 of *Lecture Notes in Computer Science*, pages 30–39. Springer Berlin/Heidelberg, 2009.
- [4] John Brooke. Sus - a quick and dirty usability scale.
- [5] Andrew T. Duchowski. *Eye Tracking Methodology: Theory and Practice*. Springer, 2003.
- [6] Joseph S. Dumas and Jean E. Fox. Usability testing: Current practice and future directions. In *Human-Computer Interaction: Development Process*. CRC Press, 2009.
- [7] Joseph H. Goldberg and Xerxes P. Kotval. Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics*, 24:631–645, 1999.
- [8] Joseph H. Goldberg, Mark J. Stimson, Marion Lewenstein, Neil Scott, and Anna M. Wichansky. Eye tracking in web search tasks: Design implications. In *Eye Tracking Research & Application*, pages 51–58, New Orleans, Louisiana, 2002. ACM New York, NY, USA.
- [9] Michael Hughes. Rigor in usability testing. *Technical Communication*, 1999.
- [10] Robert J. K. Jacob and Keith S. Karn. Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In Jukka Hyönä, Ralph Radach, and Heiner Deubel, editors, *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, pages 573–605. Elsevier Science BV, Oxford, England, 2003.
- [11] Marthe-Julie Lundeby. *Dysleksi: en sammenligning av leseforståelse, lesehastighet og øybevegelser hos dyslektiske og normallesende studenter*. PhD thesis, Norges teknisk-naturvitenskapelige universitet, 2008.

- [12] Martina Manhartsberger and Norbert Zellhofer. Eye tracking in usability research: What users really see. In Usability Symposium, editor, *Empowering Software Quality: How Can Usability Engineering Reach These Goals?*, volume vol. 198, pages pp 141–152. OCG publication, 2005.
- [13] Jakob Nielsen. Jakob nielsen’s alertbox, march 19, 2000: Why you only need to test with 5 users, 2000. url =<http://www.useit.com/alertbox/20000319.html>.
- [14] Jakob Nielsen. Ten usability heuristics, 2005. url = [http://www.useit.com/papers/heuristic/heuristic\\_list.html](http://www.useit.com/papers/heuristic/heuristic_list.html) [Accessed June 11th 2010].
- [15] Jakob Nielsen and Kara Pernice. *Eyetracking Web Usability. Voices That Matter*. New Riders, Berkeley, CA, 1st edition, 2010.
- [16] Briony J. Oates. *Researching Information Systems and Computing*. SAGE Publications, London, 2006.
- [17] Alex Poole and Linden J. Ball. Eye tracking in human-computer interaction and usability research: Current status and future prospects, 2005.
- [18] Jenny Preece, Helen Sharp, and Yvonne Rogers. *Interaction Design: Beyond Human-Computer Interaction*. John Wiley and Sons Ltd, Chichester, West Sussex, England, 2nd edition, 2007.
- [19] NIST Industry USability Reporting project. Common industry format for usability test reports, 2001.
- [20] Kristen Ringdal. *Enhet og mangfold*. Fagbokforlaget Vigmostad & Bjørke, Bergen, 2. edition, 2009.
- [21] Colin Robson. *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*. Blackwell Publishers, 2nd edition, 2002.
- [22] Anne Ryen. *Det kvalitative intervjuet: Fra vitenskapsteori til feltarbeid*. Fagbokforlaget, 2002.
- [23] Jeff Sauro. The user is in the numbers. *interactions*, 13(6):22–25, 2006.
- [24] Ben Shneiderman and Catherine Plaisant. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Pearson Education, Inc., fourth edition, 2005.
- [25] Elin Standal. Eye tracking and usability testing, 2009.
- [26] Tobii. User manual: Tobii studio, 2008.
- [27] Tobii. What is eye tracking?, 2008. url = [http://www.tobii.com/corporate/eye\\_tracking/what\\_is\\_eye\\_tracking.aspx](http://www.tobii.com/corporate/eye_tracking/what_is_eye_tracking.aspx) (Accessed: 11. des 2009).