

Brukersentrert Utvikling for Microsoft Surface

Idéutvikling, Prototyping og Wizard-of-Oz Brukbarhetstesting

Kristin Haga
Lars Loennechen Skjelbek

Master i informatikk
Oppgaven levert: Mai 2011
Hovedveileder: Dag Svanæs, IDI

Sammendrag

I dette studiet ser vi på hvordan man kan praktisere brukersentrert utvikling i de tidlige fasene av systemutviklingsløpet for applikasjoner til Microsoft Surface.

Vi har sett på to temaer knyttet til dette; introduksjon av ny teknologi for brukere, og metoder for low-fi prototyping og wizard-of-Oz brukbarhetstesting. Ved praktisering av brukersentrert utvikling er det ikke uvanlig at brukerne som involveres har en lav teknisk innsikt. Dette gjelder spesielt for nye teknologier som multi-touchbord, som fremdeles er lite utbredt. Viktigheten av å formidle mulighetene og begrensningene ved Surface til brukerne er derfor stor. I denne oppgaven blir det sett på hvilken merverdi domenespesifikke demonstrasjoner gir fremfor abstrakte når man skal formidle egenskapene ved Surface. Dette blir besvart gjennom to workshops med uavhengige grupper. Den ene gruppen får presentert abstrakte demonstrasjoner, mens den andre får presentert domenespesifikke. Ved å måle og sammenligne iderikdommen på de to workshopene blir det undersøkt hvilken av de to formidlingsmetodene som fungerer best. Resultatene viste imidlertid ingen signifikant forskjell i antall ideer som kom opp i de to workshopene, men typene av ideer hadde et merkbart skille. Det viser seg at domenespesifikke demonstrasjonseksempler ikke har merverdi overfor abstrakte, men at de to metodene snarere utfyller hverandre.

I de tidlige fasene av systemutviklingsløpet til en applikasjon for Surface er det hensiktsmessig å utvikle en prototype som kan brukes i wizard-of-Oz brukbarhetstester, slik at brukerne kan evaluere designet og komme med forslag til endringer, nye krav eller bekrefte designløsninger. Vi ser i denne oppgaven på hvilke metoder, verktøy og materialer som bør benyttes ved prototyping, og hvordan disse kan brukbarhetstestes ved bruk av wizard-of-Oz på flere testbrukere samtidig. Ulike teknikker blir utprøvd og disse resulterer i en prototype som blir brukbarhetstestet for en aktuell brukergruppe. Resultatene fra brukbarhetstestene viser at papirprototyping og wizard-of-Oz flerbrukertesting av Surface-applikasjoner lar seg gjennomføre, men at det krever god planlegging. Blant viktige utfordringer som avdekkes er testpersonenes påvirkning av hverandre og viktigheten av å skille mellom fysiske og virtuelle GUI-elementer i prototypen.

Forord

Denne masteroppgaven ble utført siste året av masterstudiet i informatikk ved Institutt for datateknikk og informasjonsvitenskap (IDI), ved Norges teknisk-naturvitenskapelige universitet (NTNU) i Trondheim, med retning systemarbeid og menneske-maskin interaksjon.

Vi vil først og fremst takke professor Dag Svanæs ved instituttet for god faglig veiledning gjennom hele året. Vi vil også takke Espen Solberg Nygård ved Accenture for støtte og gode innspill, spesielt i startfasen da arbeidet med problemstillingen var utfordrende.

Arnfinn Stendahl Rokne og Vitenskapsmuseet skal ha en stor takk for samarbeidet under viktige deler av masteroppgaven. Uten deres hjelp ville ikke første del av oppgaven latt seg gjennomføre. Norsk senter for elektronisk pasientjournal (NSEP) skal også takkes for utlån av brukbarhetslaboratoriet og utstyr brukt under brukbarhetstestene. Her var overingeniør Terje Røsand til god hjelp ved opplæring av utstyret.

IDI skal ha takk for utlånet av Surface-enheten som ble brukt til utvikling og testing av demoene. Ved instituttet har også stipendiat Ole Andreas Alsos bidratt med inspirasjon og støtte, og stipendiat Simone Mora har vært behjelpelig med tekniske spørsmål knyttet til Surface.

Til slutt ønsker vi å takke deltakerne og pilottesterne på workshopene og brukbarhetstestene for gjennomføringen av disse seansene. En spesiell takk går også til alle medstudentene på arbeidssalen Fiol for gode samtaler, uendelig mengder kaffe og selskap ved sene arbeidsdager.

Trondheim, 31. mai 2011

Kristin Haga og Lars Loennechen Skjelbek

Innhold

Sammendrag	i
Forord	iii
Innhold	v
Figurer	xv
Tabeller	xix
I Introduksjon og Teori	1
1 Introduksjon	3
1.1 Bakgrunn og motivasjon	3
1.1.1 Multi-touch kommer frem til overflaten	3
1.1.2 Utvikling av programvare for multi-touch	4
1.1.3 Utfordringen med multi-touchbord	5
1.2 Tidligere forskning	6
1.2.1 Studier om brukerinvolvering med novise brukere	6
1.2.2 Studier om prototyping for multi-touch	7
1.3 Problemstilling	8
1.4 Forskning	9
1.5 Disposisjon	10

1.6	Begrepsavklaringer	12
2	Menneske-maskin Interaksjon	15
2.1	Bakgrunn	15
2.1.1	Tverrfaglighet	15
2.2	Brukbarhet og brukeropplevelse	16
2.3	Hvordan designe systemer med høy brukbarhet?	19
2.3.1	Affordance	20
2.3.2	Constraints	20
2.3.3	Mentale modeller	21
2.4	Hvem er brukerne?	22
3	Brukersentrerte Prosesser	25
3.1	Brukersentrert utvikling	25
3.2	Hvem skal involveres?	27
3.3	Hvordan kan man involvere brukerne?	28
3.4	Spesifisere bruker og organisasjonens krav	29
3.5	Utvikle designløsninger	30
3.5.1	Low-fi kontra hi-fi prototyper	31
3.5.2	Low-fi prototyper	31
3.5.3	Hi-fi prototyper	34
3.5.4	Horisontale kontra vertikale prototyper	34
3.5.5	Tre dimensjoner av prototyping	34
3.6	Evaluering av design mot krav	36
3.6.1	Testdeltakere	37
3.6.2	Gjennomføring	37
3.6.3	System Usability Scale	40
4	Interaksjonsteknologier	41
4.1	Fra CLI til NUI	41

4.1.1	CLI	41
4.1.2	GUI	42
4.1.3	NUI	42
4.2	TUI	43
4.3	Multi-touch	44
4.3.1	Teknologier	44
4.3.2	Forskning og utvikling av multi-touchbord	45
4.4	Kommersielle multi-touchbord	47
4.4.1	Microsoft Surface	47
5	Forskningsmetoder	53
5.1	Kvalitative og kvantitative studier	53
5.2	Eksperiment	54
5.2.1	Variabler	54
5.2.2	Måling	55
5.2.3	Validitet	55
5.3	Intervju	56
5.4	Observasjon	57
5.5	Spørreskjema	58
5.6	Research through design	59
5.7	Triangulering	60
5.8	Refleksjon	60
II	Forskningsspørsmål 1	63
6	Forskningsdesign for FS1	65
6.1	Forskningsstrategi	65
6.2	Deltakere	66
6.3	Observasjon	67

6.4	Intervju	67
6.5	Spørreundersøkelse	67
6.6	Variabler	68
6.7	Oppsett av workshop	68
6.8	Validitet	69
7	Formidlingsdemoer	71
7.1	Konseptene	71
7.2	Abstrakt kontra domenespesifikt	72
7.3	Direkte manipulasjon	72
7.3.1	Abstrakt demo	73
7.3.2	Domenespesifikk demo	73
7.4	Interaksjon med fysiske objekter	73
7.4.1	Abstrakt demo	74
7.4.2	Domenespesifikk demo	74
7.5	Samhandling	76
7.5.1	Abstrakt demo	76
7.5.2	Domenespesifikk demo	77
7.6	Kombinasjon med andre teknologier	78
7.6.1	Abstrakt demo	78
7.6.2	Domenespesifikk demo	79
8	Gjennomføring av Workshops	81
8.1	Gjennomføring	81
8.2	Workshop 1 - abstrakt formidling	83
8.2.1	Deltakerne	85
8.2.2	Ideer	85
8.2.3	Kommentarer fra deltakerne	88
8.3	Workshop 2 - domenespesifikk formidling	88

8.3.1	Deltakerne	89
8.3.2	Ideer	89
8.3.3	Kommentarer fra deltakerne	92
9	Resultater fra Workshops	93
9.1	Funn	93
9.2	Analyse	94
9.2.1	Fordeler med domenespesifikke demoer	94
9.2.2	Fordeler med abstrakte demoer	95
9.3	Konklusjon	96
9.3.1	Sammenligning med annen forskning	96
III	Forskningsspørsmål 2	97
10	Forskningsdesign for FS2	99
10.1	Forskningsstrategi	99
10.2	Hva ønsker vi å finne ut?	100
10.3	Research through design	100
10.4	Deltakere	102
10.5	Lokasjon	103
10.6	Vår rolle	104
10.7	Observasjon	104
10.8	Spørreundersøkelse	104
10.9	Intervju	105
11	Prototyping og Testforberedelser	107
11.1	Konseptutvikling	107
11.1.1	Valget av idé	108
11.1.2	MultiViten	108

11.1.3	Scenarioer	111
11.2	Prototypingsteknikker	112
11.2.1	Simulering av multi-touch	112
11.2.2	Skille mellom fysiske og virtuelle elementer	115
11.2.3	Hvilke teknikker ble brukt?	127
11.3	Testteknikker	127
11.3.1	Parallele oppgavesett	127
11.3.2	Dedikerte wizards	129
11.3.3	Synkronisering	129
11.4	Prototypen	129
11.4.1	Museumskortene	130
11.4.2	Puslespillbrikkene	130
11.4.3	mediaelementene	132
11.4.4	Faktaboblene	135
11.4.5	Bluetooth-ikonene	137
11.4.6	Mobiltelefoner	137
11.4.7	Surface-bordet	137
11.5	Brukbarhetstesten	143
11.5.1	Parvis testing	143
11.5.2	Testoppgaver	143
11.5.3	Fullskala test på lab	144
11.5.4	Oppsett	146
12	Gjennomføring av Brukbarhetstest	151
12.1	Testforberedelser	151
12.2	Testgjennomføring	152
12.2.1	Testpar 1 - Lasse og Heidi	152
12.2.2	Testpar 2 - Christian og Torbjørn	153
12.2.3	Testpar 3 - Helene og Eivind	154

12.2.4	Testpar 4 - Karl og Øyvind	155
12.2.5	Testpar 5 - Jørgen og Frank	157
12.3	Testresultater	158
12.3.1	Puslespillbrikkene	160
12.3.2	Overføring via bluetooth	161
12.3.3	Trykking på bildeelementene	161
12.3.4	Flytting av mediaelementer	161
12.3.5	Forslag til forbedringer	161
13	Resultater av Brukbarhetstest	165
13.1	Funn	165
13.1.1	Testbrukernes påvirkning av hverandre	165
13.1.2	Skillet mellom fysiske og virtuelle elementer	169
13.1.3	Superrealisme	171
13.1.4	Synkronisering	172
13.1.5	Spørreundersøkelsen	174
13.2	Analyse	174
13.2.1	Testbrukernes påvirkning av hverandre	174
13.2.2	Skillet mellom fysiske og virtuelle elementer	177
13.2.3	Superrealisme	181
13.2.4	Synkronisering	181
13.2.5	Wizard-of-Oz	182
13.2.6	Fungerte brukbarhetstesten til sitt formål?	182
13.3	Konklusjon	182
13.3.1	Materialer og teknikker for prototyping	183
13.3.2	Påvirkning	184
13.3.3	Skillet mellom fysiske og virtuelle elementer	185
13.3.4	Superrealisme	186
13.3.5	Sammenligning med annen forskning	186

IV	Forskningsspørsmål 3	189
14	Anbefalinger	191
14.1	Introduksjon av teknologien	191
14.2	Low-fi prototyping	193
14.2.1	Vær tydelig på skillet mellom fysiske og virtuelle elementer	194
14.3	Wizard-of-Oz flerbrukertesting	194
14.3.1	Parallele oppgavesett	195
14.3.2	Dedikerte wizards	195
14.3.3	Datainnsamling	196
V	Avslutning	197
15	Metodediskusjon	199
15.1	Validiteten ved FS1	200
15.1.1	Ulik grad av forkunnskaper	200
15.1.2	Sosiale faktorer	201
15.1.3	Formidlingsdemoene	201
15.1.4	Modning	202
15.1.5	Validiteten	202
15.2	Validiteten ved FS2	202
15.2.1	Prosess	202
15.2.2	Nyskapning	203
15.2.3	Relevans	203
15.2.4	Utvidningsevne	203
15.2.5	Validiteten	203
16	Konklusjon og Videre Forskning	205
16.1	Konklusjon	205

16.2 Videre forskning	207
16.2.1 Refleksjoner	208
VI Tillegg	209
A Dokumenter Brukt i FS1	211
A.1 Deltakelsesskriv	212
A.2 Samtykkeerklæring	213
A.3 Spørreundersøkelse før workshop	214
A.4 Spørreundersøkelse etter workshop	215
A.5 Idékort	218
A.6 Temaliste for gruppediskusjon	219
A.7 Resultater sp.undersøkelse etter workshop	220
B Scenarioer for MultiViten	223
B.1 Scenario 1	224
B.2 Scenario 2	226
B.3 Scenario 3	229
C Dokumenter Bruk i FS2	231
C.1 Samtykkeerklæring for brukbarhetstest	232
C.2 Spørreundersøkelse om brukbarhet	233
C.3 Resultater sp.undersøkelse om brukbarhet	234
C.4 Spørreundersøkelse om testopplegget	235
C.5 Resultater spørreundersøkelse om testopplegget	236
D Øvrige Tillegg	237
D.1 Samtale med Einar Ingebretsen	238
Bibliografi	239

Figurer

2.1	De forskjellige profesjonene som utgjør fagfeltet MMI. Illustrasjon hentet fra University of Minnesota [53].	16
2.2	Attributter som inngår i akseptansen av et system. Gjengitt fra [47].	17
2.3	Mentale modeller av kalenderåret.	22
3.1	Den gjensidige avhengigheten i brukersentrerte designaktiviteter, ISO 9241-210 [1].	26
3.2	Buxtons skisse av kartnavigasjon på Tablet-PC. Faksimile fra “ <i>Sketching User Experiences</i> ” [7].	32
3.3	To dimensjoner av prototyping. Gjengitt fra [47].	35
3.4	Modell av hva prototyper prototyper. Gjengitt fra [32].	36
3.5	Brukbarhetsfeil som avdekkes i forhold til antall testbrukere.	38
4.1	Durell Bishops <i>Marble Answering Machine</i>	43
4.2	Optisk multi-touchskjerm. Illustrasjon hentet fra Scientific American [3].	46
4.3	Microsoft Surface.	48
4.4	En bruker tar på et bilde og flytter det rundt. Illustrasjon hentet fra Surface User Experience Guidelines [14].	48
4.5	En bruker skalerer bildet. Illustrasjon hentet fra Surface User Experience Guidelines [14].	49
4.6	Surface tags i riktig størrelsesforhold.	49
7.1	Formidlingsdemoer for direkte manipulasjon.	73

7.2	Fysiske objekter med tags festet på undersiden.	75
7.3	Formidlingsdemoer for interaksjon med fysiske objekter.	75
7.4	Formidlingsdemoer for samhandling.	77
7.5	Fysiske objekter med tags festet på undersiden.	78
7.6	Formidlingsdemoer for Surface i kombinasjon med andre teknologier.	79
7.7	Mobiltelefoner som har overført et bilde fra Surface.	80
8.1	Introduksjon.	82
8.2	Demonstrasjon av formidlingsdemoene.	82
8.3	Idédugnad i par.	83
8.4	Kort for gruppe B.	84
8.5	Presentasjon av ideene.	84
8.6	Illustrasjon av ideen <i>Historiske hendelser</i>	86
8.7	Illustrasjon av ideen <i>Videoavspilling</i>	86
8.8	Illustrasjon av ideen <i>Kategorisering av attraksjoner</i>	87
8.9	Illustrasjon til ideen <i>Fargelegging</i>	90
10.1	Skisse over brukbarhetslaboratoriet på NSEP.	103
11.1	Museumskortet for utstillingsgjenstanden isbjørn.	109
11.2	Skisse av papirprototyping for multi-touch.	114
11.3	Skisse av projiseringsteknikken ved prototyping for Surface.	116
11.4	Glassplate på bordben.	118
11.5	Wizard som presser en virtuell puslespillbrikke mot undersiden av en glassplate når et fysisk museumskort ligger på oversiden.	118
11.6	GUI-elementer som ligger oppå andre kan ikke flyttes uten å senke alle GUI-lagene.	119
11.7	Stillinger wizarden kan sitte i for å se hva som foregår på overflaten.	120
11.8	Bruk av magneter på museumskortet og puslespillbrikken.	120

11.9	Skisse av bruk av magneter på GUI-elementer og magnetiske hansker.	121
11.10	Hansker med magneter i fingertuppene interagerer med magnetiske GUI-elementer.	122
11.11	Når en testbruker løfter hånden vil magnetkraften mellom GUI-elementet og hansken opphøre.	123
11.12	Museumskort i papptykkelse.	124
11.13	Bruk av sandpapir på puslespillbrikkene.	125
11.14	Bruk av transparenter for å hindre flytting av statiske GUI-elementer.	126
11.15	Museumskortene.	131
11.16	Puslespillbrikke med sandpapir for å skape friksjon mellom museumskort og underlag.	132
11.17	Skalering av et bildeelement.	133
11.18	Avspilling av et videoelement.	134
11.19	Avspilling av et lydelement.	134
11.20	To puslespillbrikker med mediaelementer.	135
11.21	To puslespillbrikker føres sammen og mediaelementene byttes ut med felles fakta om de to.	136
11.22	Ikon for å representere en mobiltelefon innen rekkevidde.	137
11.23	Mockups av to mobiltelefoner.	138
11.24	Surface-mockup.	139
11.25	Surface-mockup med oppsamlingsboks for brukte kort.	140
11.26	Høytalere på innsiden av Surface-mockupen.	141
11.27	Tekst som vises på skjermen når MultiViten ikke er i bruk.	142
11.28	Planskisse av innredningen av testrommet.	144
11.29	Avdelingene med museumsgjenstander.	147
11.30	Surface-mockupen ble plassert i testrommet med god avstand rundt.	148
11.31	Wizardenes arbeidspulter.	148

12.1	Testpar 1 - Lasse og Heidi.	153
12.2	Testpar 2 - Christian og Torbjørn.	154
12.3	Testpar 3 - Helene og Eivind.	156
12.4	Testpar 4 - Karl og Øyvind.	157
12.5	Testpar 5 - Jørgen og Frank.	158
12.6	Påstand 10.	160
13.1	Jørgen kobler sin puslespillbrikke mot Franks.	168
13.2	To testbrukere forsøker å trykke på fysiske museumskort.	170
13.3	Øyvind legger et kort på mobiltelefon-mockupen.	170
13.4	Øyvind legger et museumskort oppå et annet.	171
13.5	Helene fører hodet nærmere bildeelementet i stedet for å skalere det opp.	172
13.6	Søylediagram for svarene på påstand 5: <i>det var hele tiden lett å skille mellom hva som var fysiske objekter og hva som var virtuelle objekter på Surface-bordet.</i>	178
13.7	Søylediagram for svarene på påstand 6: <i>“jeg forstod lett at det egentlig ikke skal være mulig å flytte på puslespillbrikken uten å bevege museumskortet som lå oppå”.</i>	179
13.8	Flytting av puslespillbrikker uten å bruke museumskortet.	179
13.9	Søylediagram for svarene på påstand 4: <i>“det var vanskelig å forstå at det dukket opp en virtuell puslespillbrikke på ‘Surface-skjermen’ når jeg la ned et fysisk museumskort”.</i>	180
14.1	Demoenes ulike fokus i Houde og Hills modell av hva prototyper prototyper (se Seksjon 3.5.5).	192

Tabeller

10.1 Spørsmål vi ønsker å besvare gjennom brukbarhetstestene. . .	101
11.1 Teknikker brukt til prototyping av MultiViten	128
11.2 Museumskortene som ble laget	131
11.3 Testoppgavene	145
12.1 Påstander om brukbarheten ved MultiViten med tilhørende totalscore	159
12.2 Brukbarhetsproblemer og forslag til forbedringer	163
13.1 Påstander angående prototypen med tilhørende totalscore . . .	175
13.2 Modell av de ulike påvirkningsformene og konsekvensene av disse	176
15.1 Sammenligning av forkunnskaper mellom W1 og W2	200

Del I

Introduksjon og Teori

Kapittel 1

Introduksjon

Dette kapitlet innleder oppgaven med å forklare grunnlaget for forskningen og hva den omfatter. Seksjon 1.1 gir en kort presentasjon av historien bak multi-touch og hvilke utfordringer utviklere står overfor når teknologien blir vanligere i private og offentlige rom. Forskningsspørsmålene fremstilles i Seksjon 1.3 og utledes videre i Seksjon 1.4 der fremgangsmåten blir forklart. Kapitlet avsluttes med en oversikt over rapportens disposisjon i Seksjon 1.5 og begrepsavklaringer i Seksjon 1.6.

1.1 Bakgrunn og motivasjon

1.1.1 Multi-touch kommer frem til overflaten

Multi-touchteknologi innebærer at brukere kan interagere med et brukergrensesnitt ved å være i direkte kontakt med skjermen med to eller flere kontaktpunkter samtidig [80]. Selv om dette er en interaksjonsteknikk som har sett en oppblomstring de siste årene har teknologien eksistert i lengre tid. Bob Boie ved Bell Labs regnes for å ha utviklet den første multi-touchskjermen allerede i 1984 [8]. Da hadde single-touchteknologi eksistert i over ti år med blant annet PLATO IV, en datamaskin utviklet for undervisning. Flere prototyper av multi-touchskjermer har senere blitt utviklet uten at de har blitt kommersialisert. Det var ikke før i 2007 med Apples introduisering av iPhone og iPod Touch at multi-touch for alvor fikk kommersiell eksponering [34]. Apple har siden hatt stor suksess med iPhone, og andre mobiltelefonprodusenter har tatt i bruk multi-touch [37]. Sammen har markedet sett en hurtig vekst i salget av håndholdte enheter med multi-touchskjermer [12]. 2007 var

også året da Microsoft annonserte sitt multi-touchbord; *Microsoft Surface* (heretter kalt Surface). Surface, med sin 30" store touchskjerm, skilte seg fra iPhone ved at størrelsen og antall mulige touchpunkter tillot flere brukere å interagere med skjermen samtidig. Dessuten kunne den gjenkjenne andre objekter på skjermen enn fingre. Surface markerte starten på multi-touchbord som et kommersielt produkt etter årevis med prototyper. Flere kommersielle multi-touchbord har fulgt Surface i ettertid (se Seksjon 4.4). Prisen for en enhet lå på rundt 150 000 NOK [51], og det er uvisst hvor mange enheter som er solgt ettersom Microsoft ikke har offentliggjort salgstallene. Surface er ute av produksjon, men den 6. januar 2011 annonserte Microsoft en ny utgave av Surface, kalt Surface 2.0 (beskrevet i Seksjon 4.4.1). Dette viser at Microsoft fortsatt har tro på multi-touchbord som en kommersiell teknologi.

1.1.2 Utvikling av programvare for multi-touch

11. juli 2008 lanserte Apple *App Store* til iPhone og iPod Touch, en løsning som åpnet for tredjeparts programvare på enheten. Utviklere kunne programmere egne applikasjoner og fikk mulighet til å tilby disse gjennom App Store slik at brukere enkelt kunne laste ned og installere de på enheten sin. App Store ble en suksess og antallet applikasjoner og nedlastinger har økt betraktelig siden lanseringen. Fra 500 tilgjengelige applikasjoner på lanseringsdato har tallet steget til over 300 000 i begynnelsen av 2011 [38]. Den enorme veksten av såkalte *apps* (mobilapplikasjoner) har ført til at brukerne får en stadig større valgfrihet blant apps med samme formål. Suzanne Ginsberg [26] poengterer at tiden er forbi da utviklere kunne legge ut appen sin i App Store så fort som mulig, for så å oppdatere den etter de første negative tilbakemeldingene. Med dagens valgmuligheter i App Store kan brukere raskt og enkelt velge å bruke en annen app dersom den første ikke tilfredsstillt behovet deres. Ginsberg anbefaler derfor å følge *brukersentrerte* metoder for utvikling av iPhone-apps. Ved å studere brukere i deres naturlige omgivelser og attpåtil inkluderer de i utviklingsprosessen vil utviklere få verdifull hjelp. På den måten har appen en større sjanse for å møte behovene for potensielle brukere og passe bedre inn i deres brukskontekst.

Det Ginsberg la merke til vedrørende det tidlige stadiet av app-utvikling til iPhone har fellestrekk med dagens utvikling av applikasjoner til multi-touchbord. Multi-touchbord har foreløpig ingen stor brukermasse. Det er få husholdninger som har råd til Surface som stuebord. Kundegruppen som Microsoft retter seg mot er bedrifter, hoteller, restauranter, museer og lignende organisasjoner som har råd til slik teknologi. Likevel ser vi en trend som peker

i retning av multi-touchbord i mindre offentlige settinger og husholdninger. Introduseringen av Apples iPad den 27. januar 2010 [35] og andre lignende tablets som fulgte viser at brukere er modne for større multi-touchflater enn det mobiltelefoner kan tilby. Det kan dessuten antydes at med Surface 2.0 beveger teknologien seg mot at multi-touchbordene blir tynnere, mer mobile, billigere og i alt mer egnet for installasjon og domestisering i private rom. En stor del av det man har sett av applikasjonsutvikling til multi-touchbord de siste årene har vært gjennomført av softwareutviklere som vil vise teknologiens muligheter uten særlig stor grad av brukermedvirkning. De har hatt ideer til hva teknologien kan brukes til, utviklet ideen og til slutt vist den frem på messer. Vi har også sett tilfeller der interaktive Surface-applikasjoner har vært utviklet på oppdrag fra kunde uten praktisering av en brukersentrert utviklingsprosess. Den 16. november var vi i dialog med Einar Ingebretsen, systemutvikler ved Bouvet (se Tillegg D.1). Han har utviklet Surface-applikasjonen som deltakerne bruker i NRKs underholdningsprogram “De Ukjente”¹. Han kunne fortelle om en prosess preget av hastverk der han fikk en kravspesifikasjon fra NRK, før han laget en konseptskisse i Excel og deretter implementerte applikasjonen. Prosessen var lite preget av brukermerdvirkning.

Ettersom multi-touchbordenes brukermasse tiltar er det nødvendig å ta i betraktning brukersentrert utvikling av applikasjoner for denne plattformen.

1.1.3 Utfordringen med multi-touchbord

Selv om vi kan se stadig mer bruk av multi-touchbord i samfunnet, er det fortsatt en vei å gå før vi kan stadfeste at “mannen i gata” er kjent med teknologien. Utfordringen er derfor desto større når en skal utøve brukersentrert utvikling for slike bord. En av utfordringene er å formidle mulighetene og begrensningene ved teknologien på en forståelig måte så disse brukerne kan involveres i utviklingen. Som teknologer er vi vant med å utlede konkrete bruksområder for teknologi basert på abstrakte beskrivelser av mulighetene. Vår hypotese er at dette er vanskeligere for ikke-teknologiske brukere og at man derfor bør formidle teknologien ved hjelp av flere konkrete eksempler på bruk i brukerens eget domene. Vi anerkjenner at fordelene ved mer abstrakte eksempler kan være nyttig for å unngå at brukerne henger seg for mye opp i de konkrete demonstrasjonseksemplene og hemmer kreativiteten. Likevel tror vi at brukerne vil ha vanskeligheter med å forstå hvordan teknologien kan hjelpe dem, basert på eksempler som ikke har noe med deres situasjon

¹http://www.nrk.no/programmer/sider/de_ukjente/

å gjøre.

Når utviklere benytter seg av en brukersentrert utviklingsprosess er det vanlig å lage prototyper av designet og teste det på utvalgte brukere. I tidlige prototypingsfaser av dataprogrammer og websider blir det gjerne brukt konvensjonelle verktøy og materialer som saks, papir og blyant. Brukbarhetstestene kan være såkalt “wizard-of-Oz” (se Seksjon 3.5.2) der en person simulerer datamaskinens oppgaver, som å bytte skjermbilder og utføre handlinger når brukeren trykker på en papirknapp. Prototyper i senere faser av utviklingen øker i kompleksitet og kan med tiden se ut som en fungerende applikasjon på en datamaskin uten at den er det. Som det fremgår av Seksjon 1.2 er det gjort lite forskning på low-fi prototyping for Surface. Det vil derfor være nødvendig å undersøke om det lar seg gjøre å utvikle slike prototyper for Surface. Særlig er dette viktig med tanke på multi-touchbordets høye kostnader. Utfordringen, for øvrig, ligger i Surface-applikasjoners natur: interaksjonen foregår med en rekke finger- og håndbevegelser, elementer skal kunne skaleres, roteres og flyttes, og fysiske objekter kan skape helt spesielle hendelser på bordet ved å utnytte såkalte “Surface tags” (forklart i Seksjon 4.4.1). Dette innebærer at nye metoder, verktøy og materialer bør vurderes når man skal utvikle prototyper for Surface for at de skal kunne være i stand til å simulere disse egenskapene.

Det overordnede forskningstemaet for vår forskning er hvordan systemutviklere i større grad kan involvere brukerne i utviklingen av applikasjoner for Surface i de tidlige fasene av systemutviklingsløpet.

1.2 Tidligere forskning

For å få en oversikt over hva som er gjort av undersøkelser og forskning rundt vår problemstilling har vi satt oss inn i tidligere studier på tilsvarende områder. Denne seksjonen trekker frem studier som har relevans for oppgaven vår.

1.2.1 Studier om brukerinvolvering med novise brukere

Når man praktiserer brukersentrert utvikling er det ikke uvanlig at brukerne som involveres har en lav teknisk innsikt. Man må da tenke på at det kan være vanskelig for disse å bidra til utviklingen av systemet uten å få en god introduksjon til teknologien. Spørsmålet blir da hvordan man på best måte kan introdusere ny teknologi for brukere uten særlig teknisk innsikt.

Klingsheim og Raae har forsket på nettopp dette i sin masteroppgave “*Introducing New Technologies to Users in User-Centered Design Projects: An Experimental Study*” [56]. Studiet gikk ut på å teste ut om hands-on erfaring og abstrakte konsepter er verdifulle i en introduksjon av nye teknologier. Ved å gjennomføre to workshops med ikke-tekniske brukere for bruk av RFID²-teknologi på et sykehus – én med hands-on forklaring av teknologien og én uten – sammenlignet de ideene brukerne kom opp med. Resultatene tilsa at hands-on erfaring var svært verdifullt, men at abstrakte demonstrasjoner imidlertid var vanskelig for brukerne å overføre til sitt spesifikke domene. De foreslår derfor å la brukerne bli introdusert for flere små domenespesifikke eksempler i stedet for abstrakte konsepter.

I vårt studie anerkjenner vi bruken av hands-on demonstrasjoner av ny teknologi basert på Klingsheim og Raaes forskning.

1.2.2 Studier om prototyping for multi-touch

Derboven et al. skriver i artikkelen “*Comparing User Interaction with Low and High Fidelity Prototypes of Tabletop Surfaces*” [18] om bruken av low-fi og hi-fi prototyping (forklart i Seksjon 3.5.1) for flerbruker- og multi-touch grensesnitt. Studiet var basert på å foreta en serie av brukbarhetstester, både low-fi og hi-fi, for så å sammenligne disse. Low-fi-testene bestod av et vanlig arbeidsbord med konvensjonelle, fysiske verktøy (saks, blyant, teip og lignende), mens hi-fi-testene ble gjort på en implementert applikasjon. Deltakerne på hver av de forskjellige typene tester fikk samme oppgave. Studiet kom frem til at man bør være forsiktig med å generalisere høy-nivå brukerinteraksjoner fra en low-fi til en hi-fi prototype. Dette studiet brukte imidlertid konvensjonelle verktøy i low-fi-testen i stedet for å faktisk papirprototype en virtuell multi-touchapplikasjon. Derboven et al. poengterer at det ikke er etablert noen konkrete metoder for å utvikle papirprototyper til multi-touchbord og at studiet deres ikke var et direkte forsøk på dette. De undersøkte snarere om bruk av konvensjonelle verktøy for å løse en oppgave kan avdekke brukbarhetsfeil som en hi-fi applikasjon for den samme oppgaven klarer. Det blir diskutert i artikkelen at det i videre forskning vil være spennende å se på hvilke innovative, primitive metoder som kan utvikles for å gjøre observasjoner på low-fi prototyper mer forutsigende for hi-fi resultater.

Artikkelen “*Problems with Testing Multi-User Multi-Touch Tabletop Applications*” [68] av Sohan ser også på brukbarhetstesting av multi-touchbord.

²Radio Frequency Identification

Han tar for seg utfordringer som utviklere av multi-touchapplikasjoner spesielt står overfor. Fokuset her er på testing av applikasjoner på alle nivåer av fidelity, blant annet behovet for å teste på flerbruker-scenarier. Artikkelen peker på at det allerede er gjort mye forskning rundt små multi-toucheenheter som dagens smarttelefoner, hvor fokuset er rettet mot en enkelt bruker som utfører forskjellige multi-touch gestures (forklart i Seksjon 4.3). Små enheter som dette skiller seg imidlertid fra multi-touchbord [68]. En av grunnene som nevnes er at noen applikasjoner er ment for å brukes av flere brukere samtidig og det vil dermed være nødvendig å teste nettopp flere personer på en gang. Denne artikkelen tar ikke for seg problemer som påvirkning, samarbeid og synkronisering mellom brukere, men handler i større grad om interaksjonsmuligheter som oppstår når man er flere brukere på bordet samtidig. Sohan etterlyser gode metoder for å flerbrukerteste multi-touchapplikasjoner.

Disse studiene peker alle på momenter som er interessante å se på i relasjon til vår forskning. De gir dessuten grunnlag til å hevde at dette er et aktuelt forskningsområde.

1.3 Problemstilling

Motivasjonen bak forskningen er mangelen vi ser på brukermedvirkning i utviklingen av applikasjoner for Surface. For å bidra til økt brukskvalitet og tidligere avdekning av designfeil ser vi et behov for forskning på brukersenterte prosesser for denne plattformen. Med motivasjon i Klingsheim og Raaes forslag om bruk av domenespesifikke eksempler i stedet for abstrakte når teknologi introduseres for novise brukere, ønsker vi å se nærmere på disse to måtene når Surface skal introduseres for ikke-tekniske brukere under designworkshops. Når en slik workshop i et typisk utviklingsløp har resultert i en designløsning vil vi dessuten se på hvordan denne løsningen kan prototypes og brukerevalueres.

Problemstillingen for dette studiet blir å:

- Undersøke hvordan man bør gjennomføre en brukersentrert utviklingsprosess for Surface i de tidlige fasene av utviklingsløpet.
- Komme frem til anbefalinger for gjennomføring av designworkshops, prototyping av designløsninger, og brukbarhetstesting.

På grunnlag av problemstillingen er det utledet tre forskningsspørsmål:

- **FS1:** I en brukerdeltakende designworkshop, hvilken merverdi gir domenespesifikke demonstrasjonseksempler fremfor abstrakte ved formidling av mulighetene og begrensningene til Surface?
- **FS2:** Hvilke metoder, verktøy og materialer bør benyttes ved prototyping av applikasjoner for Surface i de tidlige fasene av systemutviklingsløpet, og hvordan kan disse prototypene brukbarhetstestes ved hjelp av wizard-of-Oz?
- **FS3:** Hvilke anbefalinger kan vi komme med til prosessen rundt brukersentrert utvikling for Surface i de tidlige fasene av systemutviklingsløpet?

FS1 og FS2 vil sammen danne grunnlaget for å kunne svare på FS3.

1.4 Forskning

For å svare på FS1 vil vi gjennomføre to workshops med brukere i et bestemt domene. På den ene workshopen vil vi formidle sentrale konsepter ved Surface som viser dets muligheter og begrensninger ved hjelp av abstrakte eksempler. På den andre workshopen vil vi på tilsvarende måte bruke konkrete, domenespesifikke eksempler. Eksemplene vil være applikasjoner som utvikles og kjøres på Surface. Under workshopene vil deltakerne bli bedt om å komme frem til ideer til hva Surface kan brukes til i det aktuelle domenet. Ved å måle og sammenligne iderikdommen på de to workshopene vil vi undersøke hvilken av de to formidlingsmetodene som fungerer best.

Workshopene vil resultere i flere ideer til applikasjoner som kan implementeres på Surface. For å besvare FS2 kommer vi til å velge én av disse til et konsept som det skal jobbes videre med. Det vil bli eksperimentert med ulike teknikker og materialer for å utvikle en prototype for dette konseptet. Prototypen vil så brukbarhetstestes ved å gjennomføre wizard-of-Oz brukbarhetstester på flere brukere samtidig. Eksperimenteringen med prototyping og brukbarhetstestene vil avdekke utfordringer knyttet til bruk av slike metoder til Surface.

1.5 Disposisjon

Rapporten er delt inn i seks hoveddeler; *Introduksjon og Teori, Forskningsspørsmål 1, Forskningsspørsmål 2, Forskningsspørsmål 3, Avslutning, og Tillegg*. I delene som omhandler FS1 og FS2 blir forskningsdesign, gjennomføring, resultater og konklusjon til hvert av spørsmålene presentert. I delen *Avslutning* blir forskningsmetodene som er benyttet i FS1 og FS2 diskutert og forsvart med tanke på validitet. I denne delen blir også forskningsspørsmålene konkludert under ett.

Denne disposisjonen vil gi et kort innblikk i hva de ulike kapitlene tar for seg.

Kapittel 1 Introduksjon gir en introduksjon til prosjektet, ved å presentere motivasjon, problemstilling og forskningens kontekst.

Kapittel 2 Menneske-maskin Interaksjon presenterer teori som omhandler fagfeltet MMI som blir anvendt i forskningen.

Kapittel 3 Brukersentrerte Prosesser tar for seg teori om prosesser knyttet til brukerinvolvering. Dette inkluderer metoder og teknikker som benyttes for å involvere brukere i evalueringen et system, hovedsaklig prototyping og brukbarhetstesting.

Kapittel 4 Interaksjonsteknologier gir en presentasjon av interaksjonsteknologiens evolusjon, samt innføring i multi-touch, multi-touchbord og Surface.

Kapittel 5 Forskningsmetoder gjør rede for de ulike forskningsmetodene som er relevante for forskningen.

Kapittel 6 Forskningsdesign for FS1 forklarer hvordan vi har tenkt å besvare FS1.

Kapittel 7 Formidlingsdemoer presenterer formidlingsdemoene som benyttes i workshopene.

Kapittel 8 Gjennomføring av Workshops tar for seg hvordan workshopene ble gjennomført.

Kapittel 9 Resultater fra Workshops presenterer og analyserer resultatene etter workshopene.

Kapittel 10 Forskningsdesign for FS2 forklarer hvordan vi har tenkt å besvare FS2.

Kapittel 11 Prototyping og Testforberedelser omhandler utvikling, prototyping og brukbarhetstesting av et applikasjonskonsept for Surface.

Kapittel 12 Gjennomføring av Brukbarhetstest tar for seg gjennomføringen av brukbarhetstestene.

Kapittel 13 Resultater av Brukbarhetstest presenterer og analyserer resultatene av brukbarhetstestene.

Kapittel 14 Anbefalinger utleder svaret på FS3 i form av anbefalinger for brukersentrert utvikling for Surface i de tidlige fasene av systemutviklingsløpet.

Kapittel 15 Metodediskusjon diskuterer resultatene og validiteten ved de anvendte forskningsmetodene.

Kapittel 16 Konklusjon og Videre Forskning besvarer oppgavens tre forskningsspørsmål på bakgrunn av resultatene fra studiet. Kapitlet kommer også med forslag til videre forskning.

1.6 Begrepsavklaringer

Viktige begrep og forklaringer er listet opp i denne seksjonen. Siden IT-fagfeltet er sterkt dominert av engelske termer vil det bli brukt engelske ord og bøyninger av disse der det virker naturlig. Eksempelvis vil vi bruke ordet workshop og flertallsformen workshops. Vi vil imidlertid bruke norsk form i bestemt en- og flertall på den engelske termen (workshopen og -e).

Demo	Implementert applikasjon som viser mulighetene ved Surface.
Enbrukertesting	Brukbarhetstesting med én bruker av gangen.
Flerbrukertesting	Brukbarhetstesting med to eller flere brukere samtidig.
Gesture	Hånd og fingerbevegelser på en multi-touchskjerm som tolkes som en egen inpuhandling.
GUI	Grafisk brukergrensesnitt (eng: Graphical User Interface).
Hands-on	Praktisk og aktiv bruk av teknologi for å få erfaring.
Hi-fi prototype	High fidelity. Komplekse prototyper med mer detaljer. Ligner mer på et implementert system.
Low-fi prototype	Low fidelity. Enkle, primitive prototyper uten mye detaljer.
MMI	Menneske-maskin interaksjon.
Mockup	Full-skala modell av et produkt eller system.
Muldvarp	Person/agent som opererer “undercover” med skjulte hensikter. Brukes i denne oppgaven om en testbruker som opptrer på lik linje med andre testbrukere, men som samtidig er en del av teststaben.
Multi-touch	En overflates evne til å gjenkjenne to eller flere kontaktpunkter samtidig.
Multi-touchbord	En multi-touchskjerm som er plassert på ben eller en kropp slik at det opptrer som et bord. Også kalt “tabletop” i industrien.

MultiViten	Applikasjonskonseptet som fremstilles gjennom oppgaven.
NSEP	Norsk Senter for Elektronisk Pasientjournal ved St. Olavs Hospital, NTNU.
NUI	Naturlig brukergrensesnitt (eng: Natural User Interface).
PD	Participatory Design. En tilnærming hvor brukere involveres i designprosessen.
Superrealisme	For GUI-elementer som er ment å simulere virkeligheten er superrealisme egenskaper som går forbi det realistiske.
Surface	Microsoft Surface, multi-touchbordet som benyttes i oppgaven.
SUS	System Usability Scale. Spørreskjema for å måle brukstilfredshet.
Tablet	Enhet med større touchskjerm enn dagens smarttelefoner. Eksempel: Apple iPad.
Tag	Surface-tags. Tags er spesiallagede strekkoder som Surface klarer å gjenkjenne.
TUI	Tangible User Interface.
TUIO	Tangible User Interface Objects. Åpent rammeverk som definerer en felles protokoll og API for multi-touchflater med TUI.
W1	Workshop 1 (fikk presentert abstrakte formidlingsdemoer).
W2	Workshop 2 (fikk presentert domenespesifikke formidlingsdemoer).
Wizard	Person som simulerer datamaskinens handlinger.
Workshop	Designworkshop. Samling hvor utviklere, representanter fra kunden og brukere jobber sammen for å forsøke å designe en løsning.

Kapittel 2

Menneske-maskin Interaksjon

Dette kapitlet presenterer relevant bakgrunnsteori for forskningen vår fra fagfeltet MMI.

2.1 Bakgrunn

Menneske-maskin interaksjon, heretter forkortet MMI, er kort fortalt studiefeltet som omhandler menneskers samhandling med datamaskiner. Det finnes ingen fastslått, entydig definisjon av begrepet, men ACM Special Interest Group on Computer Human Interaction (ACM SIGCHI) [20] har kommet frem til følgende (vår oversettelse):

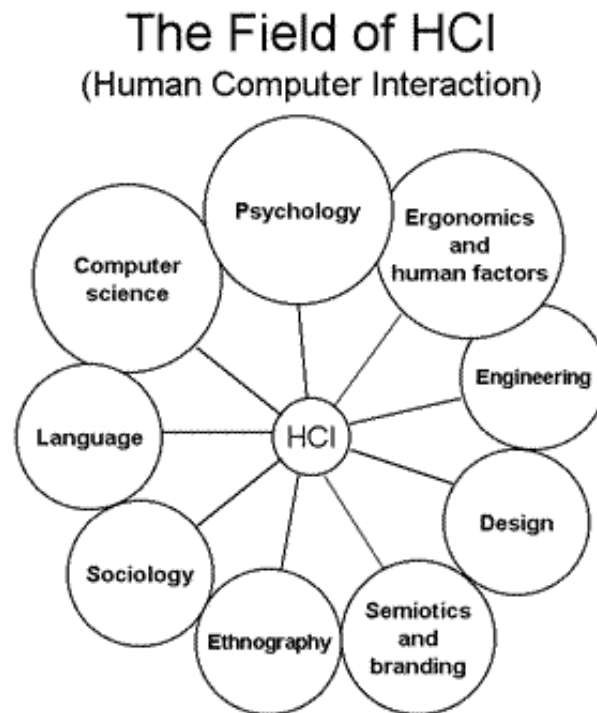
“MMI er fagområdet som dreier seg om design, evaluering og implementering av interaktive datasystemer for menneskelig bruk og studiet av store fenomener tilknyttet dette.” - ACM SIGCHI [20, s. 5]

Fagområdets fokus er interaksjonen mellom mennesker og maskiner og hvordan brukeropplevelsen kan gjøres best mulig.

2.1.1 Tverrfaglighet

Selv om MMI særlig studeres innen informatikk er det ikke utelukkende et datateknisk fagområde. En tverrfaglighet bestående av informatikk, psykologi, sosiologi, antropologi og industriell design utgjør forskjellige fokus innen

MMI (se Figur 2.1). I tillegg til kunnskap om systemutvikling og design av GUI innebærer dette teorier om kognitive prosesser, empiriske analyser av brukeres atferd og organisasjonsforståelse. Det kan ofte være fordelaktig å ha innsikt innen flere av disse områdene når man arbeider med interaksjon [20].

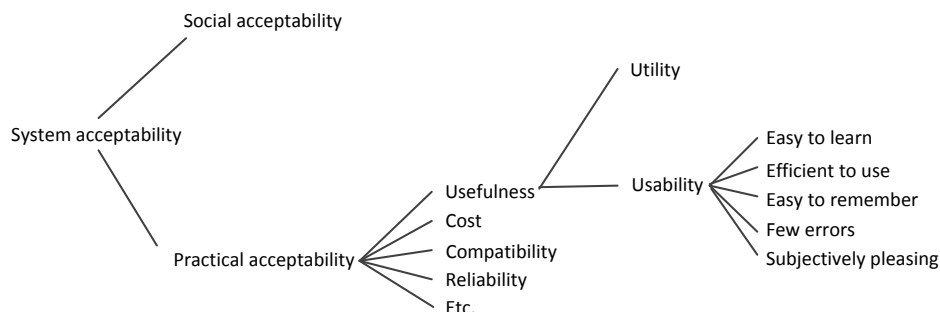


Figur 2.1: De forskjellige profesjonene som utgjør fagfeltet MMI. Illustrasjon hentet fra University of Minnesota [53].

2.2 Brukbarhet og brukeropplevelse

Det finnes flere publikasjoner som tar for seg retningslinjer og mønsterpraksiser innen MMI for god interaksjonsdesign. Blant flere anerkjente fageksperter innen fagfeltet MMI finner vi Jakob Nielsen. Han publiserte i 1993 boka “*Usability Engineering*” [47] der hovedmålet med boka var å gi konkrete råd og metoder for hvordan systemet kan nå høyere grad av *brukbarhet*. Begrepet brukbarhet er svært sentralt innen MMI. Nielsen presenterer en modell av den totale akseptansen av et system der brukbarheten inngår som en del av den (se Figur 2.2).

Usefulness beskriver hvor nyttig systemet er for brukeren. Den er sammensatt



Figur 2.2: Attributter som inngår i akseptansen av et system. Gjengitt fra [47].

av *utility* og *usability* (brukbarhet). *Utility* er spørsmålet om hvorvidt systemet inneholder de funksjonene som brukeren trenger. *Usability* uttrykker hvor lett og tilfredsstillende det er for brukeren å bruke disse funksjonene. For å oppnå god brukbarhet mener Nielsen at man må betrakte fem attributter:

- **Lærbarhet:** Systemet må være lett å lære så brukeren kan ta det i bruk innen kort tid.
- **Effektivitet:** Når brukeren har lært seg hvordan systemet fungerer skal han kunne yte høy produktivitet.
- **Hukommelse:** Det skal være lett å huske hvordan systemet fungerer, så brukeren kan returnere til det etter et visst opphold og slippe å lære alt på nytt.
- **Feilhåndtering:** Det skal være vanskelig for brukeren å gjøre feil og hvis det skjer skal den være lett å gjenopprette.
- **Tilfredsstillelse:** Brukeren skal synes at systemet er behagelig å bruke.

Når man betrakter disse attributtene er det viktig å merke seg at et system som oppleves som brukbart av én bruker samtidig kan føles lite brukbart for en annen. Det er derfor avgjørende å definere hvem som er målgruppen for systemet og hvilken kontekst de er i.

Nielsen var ikke alene om å bruke begrepet brukbarhet og for å gi en felles definisjon ble brukbarhet standardisert i ISO 9241-210 [1]. ISO-en forklarer også hvordan man identifiserer relevant informasjon når man arbeider med brukbarhet for brukergrensesnitt. Ifølge ISO 9241-210 er brukbarhet:

“Extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use” - ISO 9241-210 [1, s. 3]

Definisjonen sier altså at brukbarhet innebærer at et system er *anvendbart, effektivt og tilfredsstillende* for *bestemte brukere* med *bestemte mål* i *bestemte omgivelser*. Videre defineres effektivitet, anvendbarhet og tilfredsstillelse som, henholdsvis:

- Høy nøyaktighet og kompletthet av målet.
- Lav ressursbruk i forhold til nøyaktigheten og komplettheten.
- Ingen ubehageligheter ved produktet, men snarere en positiv holdning til bruken av det.

Som definisjonen tilsier er det essensielt å ta i betraktning hvem som er brukerne når man utvikler et interaktivt system, hvilke oppgaver disse ønsker å løse og i hvilken kontekst de opererer i (utdypet i Seksjon 2.4). Det er derfor viktig å kjenne målgruppen, studere brukerne og jobbe nært med brukerne under utvikling av datasystemer. Utviklere og brukere bør i fellesskap komme frem til krav for produktet og brukbarhetstester bør utføres. Mangel på nærhet til brukerne resulterer ofte i systemer som ikke imøtekommer brukernes behov [62]. Begrepet *“brukersentrert utvikling”* oppstod gjennom Donald Normans arbeid på 80-tallet og blir ofte brukt om utviklingsmetoder som inkluderer brukerne i utviklingsprosessen [2]. Dette er nærmere forklart i Kapittel 3.

De siste årene har det blitt et stadig sterkere fokus på det ISO 9241-210 nevner som *tilfredsstillelse*. Fenomenet blir betegnet som *user experience* (forkortet UX), eller *brukeropplevelse* på norsk. Brukeropplevelsen relateres til brukerens følelser og er ifølge ISO 9241-210 en konsekvens av presentasjonen, funksjonaliteten, ytelsen, den interaktive atferden og assistanseegenskapene ved et interaktivt system (både maskinvare og programvare) [1]. Det er dessuten et produkt av brukerens tidligere erfaringer, holdninger, kompetanse og personlighet. Utviklere må ta dette brede spekteret av faktorer til etterretning for å skape gode brukeropplevelser.

2.3 Hvordan designe systemer med høy brukbarhet?

I Donald Normans "*The Psychology of Everyday Things*" [49] presenterer han filosofier for hvordan utviklere kan bli flinkere til å designe hverdagslige ting så disse blir mer intuitive for brukerne. Han poengterer først og fremst to enkle prinsipper; *sørg for at brukeren (1) forstår hva han skal gjøre og samtidig (2) forstår hva som foregår*. Det handler om å gjøre interaksjonen så intuitiv at brukeren kan si "ja, selvfølgelig!" når han gjør en handling for å få systemet til å utføre det han vil. For å oppnå dette fremhever han at designet bør:

- Gjøre det lett å se hvilke handlinger som er mulig til enhver tid (ved å bruke constraints, se Seksjon 2.3.2).
- Gjøre ting synlig. Dette inkluderer den konseptuelle modellen av systemet, alternative handlinger og resultatet av handlingene.
- Gjøre det lett å evaluere den nåværende tilstanden i systemet.
- Følge naturlige koblinger mellom
 - brukerens intensjon og de nødvendige handlingene som må gjøres,
 - handlingene og den resulterende effekten,
 - den synlige informasjonen og tolkningen av systemtilstanden.

Design som ikke følger disse grunnpilarene vil gjøre systemet vanskelig å forstå for de som skal bruke det. Det er likevel en utfordring å forvandle oppgaver som i utgangspunktet er kompliserte om til intuitive funksjoner for brukeren. Norman presenterer derfor syv retningslinjer for å forvandle vanskelige oppgaver til enkle:

1. **Bruk både kunnskap fra verdenen og kunnskap fra hodet.**
 - Lag gode konseptuelle modeller og skriv enkle brukermanualer.
2. **Simplifiser strukturen av oppgaver.**
 - Brukeren har begrenset kapasitet til å huske ting på kort og lang sikt. Organiser derfor strukturen slik at hukommelsen ikke overbelastes.

3. **Gjør ting synlig: skap bro mellom utførelse og evaluering.**
 - Det skal være lett for brukeren å se hvilke muligheter han har på utførelsessiden og han skal kunne se effekten av en handling.
4. **Gjør de riktige koblingene.**
 - Det bør være en intuitiv kobling mellom handling og effekt. En teknikk kan være å bruke grafikk på knapper.
5. **Utnytt både naturlige og kunstige constraints.**
 - Constraints gjør det lett for brukeren å se den riktige måten å håndtere systemet på. Et eksempel på constraints i GUI-sammenheng er å bruke kombinasjonsbokser som inneholder de eneste valgmulighetene.
6. **Design for feil.**
 - Forutse hvilke feil som er mulig og planlegg systemet slik at brukeren lett kan unngå disse eller i verste fall angre og gjenopprette.
7. **Standardiser hvis alt annet feiler.**
 - Når det ikke er mulig å lage intuitive koblinger mellom system og brukers intensjon er det beste alternativet å lage en standard. Fordelen med standarder er at de må læres én gang og fungerer så i mange sammenhenger.

2.3.1 Affordance

Norman adopterte begrepet *affordance* til MMI-kontekst. Affordance henviser til de sansede og faktiske egenskapene ved ting som avgjør hvordan den tingen kan brukes. Affordancer gir sterke ledetråder om hva brukeren kan gjøre med gjenstanden. Det dreier seg med andre ord om hvilke handlinger gjenstandene signaliserer. Utformingen til disse tingene og konteksten de står i kan si mye om hvordan de brukes. Norman bruker blant annet dørhåndtak som eksempel: dører som er ment å dyttes ut har håndtak som ikke kan gripes eller intet håndtak i det hele tatt, mens dører som må trekkes tilbake har håndtak som er lette å gripe fatt i.

2.3.2 Constraints

Norman bruker begrepet constraints om begrensinger på hvilke handlinger brukeren kan gjøre med en gjenstand. Han skriver: "*affordances suggest the range of possibilities, constraints limit the number of alternatives*" [50, s. 82].

Ved å bruke constraints ved designet av ting kan man begrense mulighetene til å bruke tingen på feil måte.

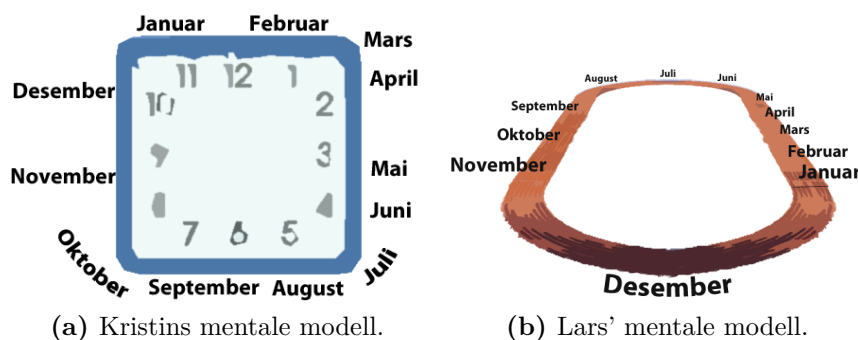
2.3.3 Mentale modeller

Vi mennesker skaper oss *mentale modeller* av alt rundt oss. Mentale modeller stammer fra fagfeltet kognitiv psykologi og er indre, subjektive representasjoner av eksterne realiteter [16]. Brukere interagerer med et system basert på sin mentale modell av hvordan systemet fungerer. Det er derfor svært viktig å forstå brukernes mentale modell av den eksterne virkeligheten for å kunne designe brukervennlige systemer [49]. Det er ikke nødvendigvis slik at utvikleren og brukeren har samme mentale modell av et fenomen, ei heller at omfanget av systemets brukere deler den samme modellen. Systemet må derfor designes slik at brukerne vil danne seg en mental modell som stemmer overens med slik det faktisk fungerer.

For å illustrere bedre hvordan våre mentale modeller fungerer vil vi vise et eksempel: Det samme filosofiske spørsmålet har dukket opp flere ganger i omgangskretsen vår – hvordan hver enkelt av oss ser for seg *kalenderåret* i hodet. Dette skaper ofte humoristiske diskusjoner der vi deler våre egne abstrakte representasjoner av dette fenomenet. For selv om kalenderen stort sett representeres på den standardiserte måten i almanakker og lignede – med 12 måneder i hver sin firkant med tilhørende tabeller som representerer ukene og ukedagene – er det ingen av oss som bruker denne i hodet. Kristin ser for seg en firkantet klokke uten visere. Rundt denne er månedene plassert i retning med klokka (se Figur 2.3a). Gjennom årets gang føler hun at vårmånedene “går nedover” mens høstmånedene “går oppover”. Denne mentale modellen er svært annerledes enn det Lars har (se Figur 2.3b). Lars forestiller seg en tredimensjonal løpebane der månedene er plassert utover mot klokka (med andre ord *i løperetningen på en løpebane*). Sommermånedene er plassert på den ene kortsiden, mens hele desember måned okkuperer den andre kortsiden. Banen heller dessuten slakt nedover mot sommeren og slakt oppover mot desember.

Et søk på Internett viser at det ikke bare er vi som filosoferer rundt dette. Flere bloggere deler sine erfaringer med spørsmålet rundt den mentale modellen av kalenderåret i vennegjengen [25][63]. Det viser seg dessuten at mange er sterkt knyttet til sin modell. De synes andre sine representasjoner er latterlige og sier blant annet at de aldri kan klare å tenke seg at kalenderåret går “mot klokka” eller at det er en sirkulær gang i det hele tatt.

Dette eksempelet illustrerer godt hvordan et fenomen som er allmennkjent



Figur 2.3: Mentale modeller av kalenderåret.

kan skape forskjellige mentale modeller. Kalenderen er for øvrig et eksempel på noe som har blitt standardisert [49]. Oppfattelsen av det abstrakte fenomenet tid har blitt standardisert gjennom kalenderen og klokka for at vi skal kunne forstå hverandre når vi omtaler tiden.

Norman [49] skiller mellom tre forskjellige *konseptuelle modeller* ved et system; *designmodellen*, *brukermodellen* og *systembildet*. Designmodellen er konseptet som utvikleren har av systemet som den vil prøve å formidle til brukeren. Brukermodellen er den mentale modellen brukeren danner seg av hvordan systemet fungerer. Brukeren skaper denne mentale modellen basert på systembildet; systemets karakteristik (grafiske utseende, funksjonene, responsen, brukermanualene). Ønskesituasjonen er at designmodellen og brukermodellen stemmer overens. For å oppnå dette må utvikleren forstå brukeren og danne en god konseptuell modell for hvordan systemet skal oppføre seg. Fordi brukeren til syvende og sist kun kommuniserer med systembildet uten kjennskap til utvikleren er det like viktig at utvikleren klarer å formidle dette konseptet gjennom systemets ansikt utad (GUI, funksjoner, brukermanualer). Hvis brukeren ender opp med en mental modell som ikke speiler systemet slik det faktisk fungerer vil det føre til feil [16].

2.4 Hvem er brukerne?

Som utledet i Seksjon 2.2 bruker ISO 9241-210 [1] begrepet “bestemte brukere” (eng. *specified users*) når brukbarhet skal defineres. Hvem er disse brukerne? En betydningsfull del av arbeidet med å designe et system er å vite hvem systemet designes for. Det er mange ulike definisjoner på hva en bruker er. Carmel et al. skrev i 1993: “*An unambiguous definition of ‘user’ is impossible*” [10, s. 40]. Videre mener de at Nord-Amerikanere forstår begrepet “bruker”

som alle ikke-tekniske individer som berøres av systemet. Skandinaver på sin side mener at begrepet kun inkluderer de operative menneskene som er i direkte kontakt med systemet. Det siste virker å sammenfalle med ISO 9241-210 som definerer en bruker slik: “*person who interacts with the product*” [1, s. 3]. For øvrig, ifølge Nielsen kan brukerne sies å inkludere alle som blir påvirket av det endelige produktet [46]. Et eksempel er innsjekking i skranken på en flyplass. Her vil ISO 9241-210 betrakte de ansatte i skranken som brukerne, siden det er de som direkte interagerer med det. Nielsen ville nok på sin side påstå at de reisende også er brukere i den grad at de blir påvirket av hvordan systemet fungerer, da med spesielt tanke på effektiviteten på kundebehandling, papirene som skrives ut og om systemet fungerer som det skal.

Uansett definisjon er det viktig å gjøre seg oppmerksom på hvem som har interesse av systemet og la stemmene til disse bli hørt under utviklingen.

Kapittel 3

Brukersentrerte Prosesser

Dette kapitlet omhandler teori om prosesser knyttet til brukerinvolvering i systemutviklingsprosjekter. Denne teorien anvendes videre i forskningen og er dermed aktuell som bakgrunnsmateriale.

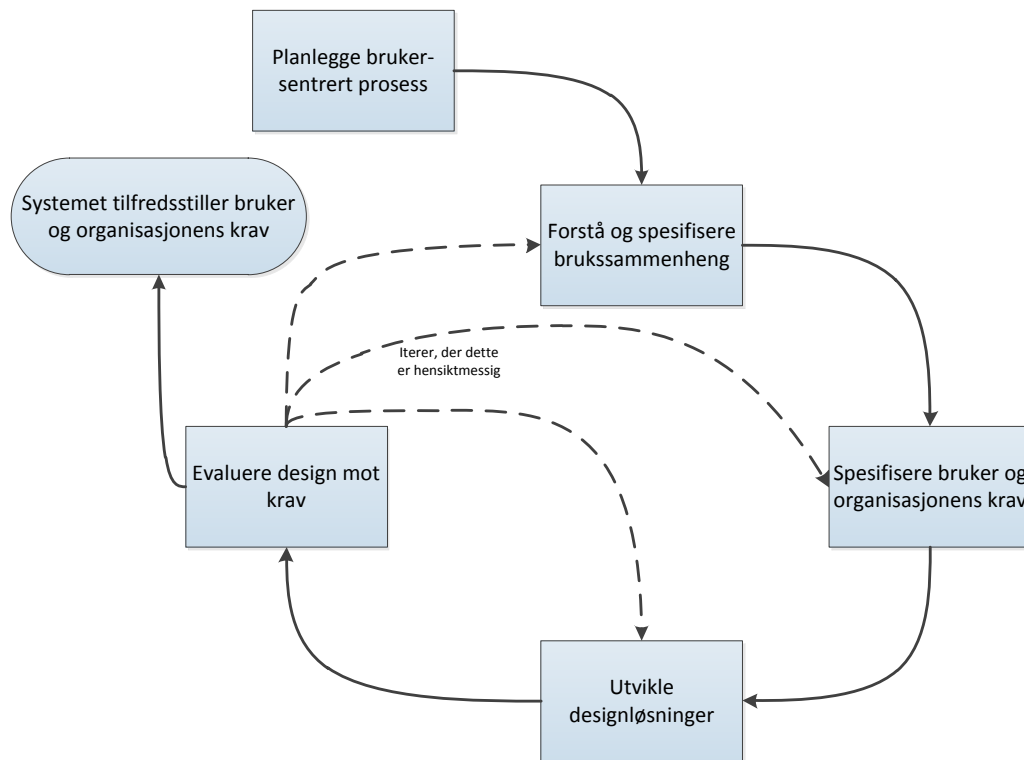
3.1 Brukersentrert utvikling

Brukersentrert utvikling er en tilnærming til utvikling av interaktive systemer som fokuserer spesifikt på å oppnå høy brukbarhet [1], og er mye brukt av dagens systemutvikling. Denne utviklingsformen tar utgangspunkt i at brukerne aktivt tar del i designarbeidet allerede på et tidlig tidspunkt, fremfor passivt å komme med tilbakemeldinger.

ISO 9241-210 beskriver fire brukersentrerte designaktiviteter som bør finne sted i et systemutviklingsprosjekt (se Figur 3.1):

- Forstå og spesifisere brukerkonteksten.
- Spesifisere brukernes og organisasjonens krav.
- Utvikle designløsninger.
- Evaluere design mot krav.

Som figuren viser bør denne prosessen starte allerede på et tidlig stadium i prosjektet. Det er nemlig viktig å forstå og identifisere brukskonteksten for å kunne ta tidlige designvalg, og for å danne et grunnlag for evaluering [1]. Designaktivitetene imøtekommer en rekke utfordringer:



Figur 3.1: Den gjensidige avhengigheten i brukersentrerte designaktiviteter, ISO 9241-210 [1].

- Det er ofte mange forskjellige brukergrupper og interesser som må tas i betraktning.
- Bruksteksten kan være svært forskjellige fra brukergruppe til brukergruppe.
- I begynnelsen av et prosjekt er det svært usannsynlig at kravene er fullstendige.
- Noen krav dukker bare opp idet en løsning foreslås.
- Brukerkrav kan være mangfoldige og motsi hverandre.
- Innledende designløsninger tilfredsstiller sjelden brukernes behov.
- Det er vanskelig å forsikre seg om at alle delene av et system er tatt i betraktning på en integrert måte.

Brukerne har i denne modellen en aktiv del i utviklingen. På et detaljert nivå kan disse aktivitetene bli brukt til å få feedback på tidlige designkonsepter før kravene er utarbeidet. Primitive prototyper og mockups vil bidra til å få en dypere forståelse av brukerkravene. Designaktivitetene er en iterativ prosess, der det kan itereres mellom de forskjellige aktivitetene etter behov. Hver aktivitet kan bruke output fra en annen.

Årsaken til at brukerne bør få en så sentral posisjon i designprosessen er at det er de som sitter på domenekunnskapen, kjenner til brukskonteksten og til syvende og sist skal bruke systemet i fremtiden. Involveringen av brukere i utviklingsprosessen er en verdifull kilde til kunnskap om brukskontekst og oppgavene systemet skal løse. Dessuten gir det innsikt i hvordan brukerne sannsynligvis vil komme til å bruke det fremtidige systemet [1]. ISO 9241-210 påpeker også at fordeling av funksjoner mellom bruker og teknologi er et av de viktigste prinsippene innenfor brukersentrert design. Dette dreier seg om hvilke funksjoner som skal tilordnes brukeren og hvilke som skal automatiseres av teknologien.

Et viktig punkt når man skal designe og utvikle systemer er å sørge for at systemet stemmer overens med brukerens modell av det. Som forklart i Seksjon 2.3.3 er målet å få brukernes mentale modell av systemet til å stemme overens med systemets konseptuelle modell for å få brukeren i stand til å løse de gitte oppgavene på en mer effektiv måte [50]. Nettopp derfor bør brukerne involveres i systemutviklingsprosessen slik at deres mentale modeller kommer til syne for designerne.

3.2 Hvem skal involveres?

En utfordring når man står overfor store brukergrupper er hvordan man skal gjøre et representativt utvalg av brukermassen, siden det som regel er svært vanskelig (om ikke umulig) å inkludere alle potensielle brukere av et system. I artikkelen “*From small scale to large scale user participation: a case study of participatory design in e-government systems*” [54] av Oostveen og Besselar blir det forklart hvordan Participatory Design (PD) kan benyttes i prosjekter med flere brukere og ikke minst flere brukergrupper. Her presenteres et prosjekt kalt FASME¹ som involverer mange ulike brukere, institusjoner og forskrifter. Aktørene er spredt over et stort geografisk område og danner store brukergrupper med et stort spekter av ulike roller innenfor systemet. Artikkelen nevner blant annet borgere og administrative roller, og vi ser at

¹Facilitating Administrative Services for Mobile Europeans

det er mange brukergrupper å ta hensyn til. Vi ser også at forholdet mellom brukerne og utviklerne blir løsere enn i de mindre prosjektene. For å involvere de ulike brukerne ble det benyttet en rekke teknikker og verktøy som intervjuer, gruppediskusjoner og designworkshops. FASME prosjektet viser hvordan man kan gjøre et representativt utvalg brukere når det ikke er mulig å involvere alle, både på grunn av at den geografiske spredningen er for stor og antallet brukere er for mange.

For å praktisere brukersentrert design trenger man et spekter av egenskaper. Dette kan man oppnå gjennom å danne *multidisiplinære team*. Multidisiplinære team vil styrke produktets integritet siden det vil vurderes ut ifra forskjellige interesseområder. Slike team trenger ikke være store, men de bør være tilstrekkelig variert til å kunne foreta seg hensiktsmessige avveinger og kompromisser i designet. Rollene kan eksempelvis inkludere sluttbrukere, spesialister på MMI og tekniske- og økonomiske ledere hos kunden (hvis systemet skreddersys). Det er ikke med det sagt at alle rollene bør være tilstede i ethvert design team. Det er snarere en oppfordring til å ta i betraktning det brede spekteret av interesser.

3.3 Hvordan kan man involvere brukerne?

Det finnes mange nivåer av brukerinvolvering, hvor både typen og nivået kan variere. Mumford [43] presenterer en medvirkningsstige (ladder of citizen participation) som illustrerer en gradering av ulike typer brukermedvirkninger. På toppen av stigen er den høyeste graden av brukermedvirkning plassert. Her blir brukerne tillagt den fulle kontrollen og utviklerne fungerer kun som en ressurs i utviklingen av systemet. Nederst på stigen plasseres den laveste graden av brukermedvirkning, manipulering. Dette er typen medvirkning som kun eksisterer på papiret, men uten at brukernes synspunkt har blitt tatt i betraktning. Det som kommer frem av Mumfords medvirkningsstige er at det ikke er en direkte sammenheng mellom det å bli involvert i utviklingsprosessen og det å ha en reell påvirkning. Hvis systemet som utvikles skal være basert på brukernes behov og preferanser må de involveres på et høyt nivå på denne stigen. Det er derfor viktig å inkorporere de tidlig i prosjektfasen og sørge for at de blir hørt.

Medvirkning legger til rette for en mekanisme som brukere kan benytte for å redesigne sine arbeidsvaner og jobbmønstendigheter om det nye systemet [5]. Dette leder til spørsmålet; hvordan skal man involvere brukerne? Gjennom intervjuer, spørreundersøkelser, observasjoner, etnografiske studier og fokus-

grupper hentes det ikke bare inn beskrivelser av gjeldende arbeidssituasjon, krav og ønsker for et nytt system. Det skapes også langsiktige arbeidsrelasjoner mellom brukere og systemutviklere [30]. Preece et. al [62] peker på at brukere som har fått bidra til utviklingen av et produkt som regel vil føle en større grad av eierskap for nettopp dette produktet. De vil da være mer kjent med hvordan produktet fungerer og være mer mottakelige når det kommer ut på markedet. Dette vil igjen føre til økt kundetilfredshet hos brukerne.

3.4 Spesifisere bruker og organisasjonens krav: Designworkshop

Designworkshop (heretter kalt workshop) er en mye brukt metode i dagens it-industri og er spesielt god på brukerinvolvering. Det finnes ingen entydig definisjon av hva en workshop er, men kort sagt kan man si at det er en samling hvor utviklere, representanter fra kunden og brukere jobber sammen for å designe en løsning. Metoden er godt egnet til å bekrefte overensstemmelser, belyse konflikter og diskutere uenigheter, slik at man kan komme til en felles enighet rundt designet av et system. Styrker ved workshop som en metode innenfor PD er [19]:

- Gir brukerne mulighet til å høres i designprosessen, noe som øker sannsynligheten for et brukervennlig design.
- Gjør det mulig for både tekniske og ikke-tekniske deltakere å delta på lik linje.
- Gjør det mulig for utviklere å møte, arbeide med og forstå brukerne.
- Skaper et forum for identifiserte problemer.
- Metoden ansees som svært produktiv.
- Benytter seg av teknikker som er lett å lære og kan anvendes i senere aktiviteter.

Når man skal gjennomføre en workshop er det viktig å gjøre et bevisst utvalg av deltakere. Blant annet kan personlige relasjoner mellom deltakerne påvirke utfallet av workshopen, for eksempel kan noen personer vise seg å være mer dominerende og opptatte av å ta kontroll, eller at det forekommer

et statusskille [62]. Først og fremst workshopen inkludere brukere, siden det er de som er domeneekspertene og verdien av deres bidrag dermed er uvurderlig. I tillegg er det aktuelt å ha med representanter fra kunden, utviklere og noen til å fasilitere. For å oppnå best mulig gruppedynamikk er det hensiktsmessig å ikke overstige ni deltakere, samtidig som man ønsker å ivareta de ulike rollene. Det finnes ingen fastsatte rammer på hvordan en workshop skal gjennomføres – en agenda vil variere avhengig av problemet, deltakerne og tiden som er tilgjengelig. *Information & Design*, et firma som er ekspert på å fasilitere designworkshops legger vekt på noen punkter som kan være aktuelle å ha med i agendaen. Dette innebærer blant annet: introduksjon av deltakerne, informasjon om workshopens hensikt og aktuelle scenarier [19].

3.5 Utvikle designløsninger: prototyper

Når utviklere har kommet fram til systemdesignets krav sammen med brukerne, må det utarbeides designløsninger som kan evalueres av brukerne (“Utvikle designløsninger” i Figur 3.1). En utbredt måte å formidle designløsninger på er å utvikle prototyper. En prototype kan defineres som en representasjon av alle eller deler av et produkt, eller interaktivt system som kan benyttes i analyse, design og evaluering, selv om den innehar visse begrensninger [1]. I IT-sammenheng brukes prototyper til å formidle ideer om hvordan systemet bør se ut og fungere. De fører utviklerens idé over i fysisk form og består av brukergrensesnitt som oppfører seg som et fungerende system til tross for at det er en imitasjon [66].

Brukere er ikke designere, og det er dermed ikke rimelig å forvente at de skal komme opp med designforslag helt fra grunnen av [47]. For å få fullt utbytte av brukerinvolvering er det nødvendig å presentere det foreslåtte designet på en måte som brukerne forstår. Siden brukere og designere skal kunne jobbe tett sammen og fatte forståelse for hverandre er det viktig at de kan kommunisere gjennom et felles språk. Ved at designerne lager konkrete eksempler som modeller, simuleringer, papirprototyper eller programvareprototyper kan brukerne få et visuelt bilde av hvordan et endelig produkt vil se ut. Dette vil bidra til at designerne kommuniserer mer effektivt med brukere og reduserer behovet for kostnad og omarbeiding som lett kan oppstå senere i utviklingsfasen.

3.5.1 Low-fi kontra hi-fi prototyper

Det finnes ulike former for prototyper. Det kan være alt fra enkle skisser med penn og papir til komplisert datamaskinbasert simulering. Hvilken fase utviklingen befinner seg i vil være med på å avgjøre hvor virkelighetsnær prototypen skal være i forhold til et endelig produkt. Tidlig i systemutviklingsfasen er det typisk å benytte seg av såkalte *low-fi* (low fidelity) prototyper. Slike prototyper skal være billige og lite tidkrevende å utvikle. Etter hvert som prototyper testes og raffineres vil de kunne øke i “troverdighet”. De vil da ligne mer på et fungerende system og kan ofte lure brukeren til å tro det er et fullverdig produkt. Dette kalles *hi-fi* (high fidelity) prototyper.

3.5.2 Low-fi prototyper

Skisser

I boka “*Sketching User Experiences*” [7] legger Bill Buxton vekt på nytteverdien av å skissere et design i idéfasen av et prosjekt. Siden investeringene er lave i begynnelsen av prosjektet bør det lages flere enkle skisser av mulige design. Ved å feile tidlig, feile ofte og lære av disse feilene vil utviklere spare tid og penger ved at disse feilene unngås i senere faser.

Skisser kan lages av det materialet man har tilgjengelig. Poenget er å formidle en designidé på en rask måte. En kan typisk tegne med blyant på ark eller tusj på whiteboard-tavle, men også bruke andre materialer på kreative måter. Et illustrerende eksempel på sistnevnte er en fungerende skisse Buxton lagde for å vise en type interaksjon for kartnavigasjon på en Tablet-PC. Måten han gjorde det på var å besøke en karttjeneste på Internett via den bærbare PC-en hans og deretter vende skjermen bakover så langt at den lå flat med baksiden av tastaturet. Til slutt festet han en datamus til siden av skjermen og holdt museknappen nede over kartet. Når denne anretningen ble flyttet rundt på et bord beveget kartet seg tilsvarende (se Figur 3.2).

Papirprototyper

En mye brukt low-fi metode er *papirprototyping*. Papirprototyper skiller seg fra skisser ved at prosjektet har gått videre fra begynnerfasen og utvikleren har en klarere idé om hvordan designet skal være [7]. Papirprototyping innebærer å illustrere et brukergrensesnitt ved hjelp av konvensjonelle midler som papir, blyant, saks og lim/teip [67]. Dette er en billig og lite tidkrevende



Figur 3.2: Buxtons skisse av kartnavigasjon på Tablet-PC. Faksimile fra “*Sketching User Experiences*” [7].

form å prototype på, og gir maksimal feedback fra brukerne ut av minimal anstrengelse. En papirprototype skal være enkel å endre på og primitiv nok til å kastes etter bruk [62]. Poenget med papirprototyping er ikke å lage en papirversjon av brukergrensesnittet som ser estetisk fint og realistisk ut. Papirprototypen skal kommunisere høynivå aspekter ved grensesnittet som eksempelvis hvordan flyten i systemet skal være, hvordan informasjonen skal organiseres og hvor knapper skal være plassert. Det er snarere positivt å la utseendet bære preg av ujevne linjer og håndskrift. Terskelen for å komme med kritikk til designet vil være mindre for testbrukere når brukergrensesnittet ser uferdig ut. [67]

Mockups

En mockup er en full-skala modell av et produkt eller system. Typiske materialer som brukes for å lage mockups er papir, papp, treverk og plast, men skiller seg fra papirprototyper ved å ligne mer på et realistisk produkt. Eksempler på mockups er:

- Papirutskrifter av GUI-elementer som settes sammen til forskjellige skjermbilder.

- En trekloss som er skåret som en mobiltelefon med post-its på som simulerer skjermbilder.
- En pappeske formet som en panteautomat der testbrukere kan legge inn flasker og motta pantelapp.
- En plastinnredning med automatutseende der testbrukere kan hente togbilletter ved å skrive inn et referansenummer på et ordentlig tastatur.

Wizard-of-Oz

Low-fi prototyper er primitive modeller uten interaktiv funksjonalitet. Derfor brukes det en teknikk som kalles *wizard-of-Oz* når slike prototyper skal brukbarhetstestes. Wizard-of-Oz fungerer ved at et menneske opptrer som mellomledd mellom testbrukeren og datamaskinen [47]. Wizarden oppfatter hvilken input testbrukeren gir til prototypen og manipulerer den slik at den gir riktig output. Denne personen – *wizarden* – skal ideelt sett være skjult for brukeren. Hvis prototypen som testes består av en ordentlig datamaskin vil en teknikk være å koble seg til denne via en skjult maskin. I eksemplene på mockups ble det nevnt en billettautomat. Denne kan brukbarhetstestes ved hjelp av wizard-of-Oz på følgende måte: en person kan gjemme seg inne i plastautomaten og følge med på en skjerm hvilket referansenummer testbrukeren skriver inn. Han kan så printe ut en billett på innsiden og levere den ut av en luke i automaten.

Selv om wizard-of-Oz-tester tradisjonelt sett innebærer en skjult wizard er ikke dette nødvendig for å demonstrere brukergrensesnittet til en prototype [67]. Under testing av primitive papirprototyper er det enklere å la wizarden sitte ovenfor testpersonen og flytte på papirelementene når testbrukeren gjør en handling. Når testbrukeren eksempelvis trykker på bildet av en knapp kan wizarden legge på et nytt ark som representerer det nye skjermbildet som dukker opp.

Smoke-and-mirrors

Buxton [7] bruker begrepet *Smoke-and-Mirrors* om en teknikk som har til felles med wizard-of-Oz at en prototype testes ved å gi testpersonen en illusjon av et fungerende system. Forskjellen er at smoke-and-mirrors utelater bruken av en mellomperson. Ved å anvende andre kreative teknikker kan testpersonen lures til å føle at et brukergrensesnitt responderer til handlingene

hans. Smoke-and-mirrors er et godt alternativ til wizard-of-Oz for prototyping av systemer med interaksjon som er vanskelig å papirprototype og/eller imiteres av en mellomperson.

Eksempelet med kartnavigasjon på en tablet-PC (Figur 3.2) viser bruk av smoke-and-mirrors for å demonstrere interaksjonen med skissen Buxton lagde.

3.5.3 Hi-fi prototyper

I senere faser av et utviklingsprosjekt blir bruk av hi-fi prototype relevant. Slike prototyper kjennetegnes ved at de ligner mer på et endelig produkt og er mer ressurskrevende å lage. I IT-bransjen blir det brukt programvareverktøy og enkle programmeringsspråk til å lage prototypene.

Når hi-fi prototyper benyttes i brukbarhetstester får testpersonen en klarere *følelse* av hvordan det vil være å bruke det endelige systemet. Brukergrensesnittets utseende, navigering og funksjonalitet vil kunne utforskes av testbrukeren uten å benytte illusjoner som i wizard-of-Oz eller smoke-and-mirrors.

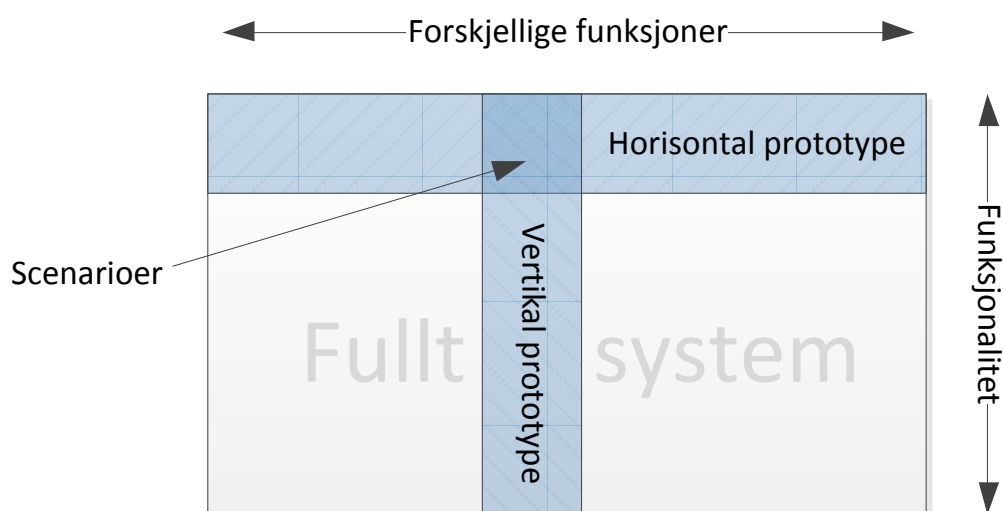
3.5.4 Horisontale kontra vertikale prototyper

Hele ideen bak prototyping er å spare tid og penger på å utvikle noe som kan testes på brukere [47]. For å spare disse ressursene kan man begrense antall funksjoner prototypen inneholder eller dybden av funksjonene. En skiller her mellom horisontale og vertikale prototyper (se Figur 3.3).

Horisontale prototyper viser alle funksjonene som det endelige systemet skal inneholde, men inneholder ingen underliggende funksjonalitet. Prototypen virker snarere som et skall som viser strukturen og navigasjonen i systemet. I motsetning til dette viser vertikale prototyper bare en liten, smal del av systemet, men følger disse få funksjonene i dybden. Prototyper kan også være en kombinasjon mellom horisontal og vertikal ved å brukbarhetsteste forhåndsdefinerte scenarioer.

3.5.5 Tre dimensjoner av prototyping

Houde og Hill presenterer en modell for tre dimensjoner av prototyping i artikkelen “*What do Prototypes Prototype?*” [32]. De poengterer at ikke alle prototyper har de samme hensiktene, men at de søker å få svar på forskjellige sider ved et foreslått, interaktivt produkt. Modellen er et triangel med tre



Figur 3.3: To dimensjoner av prototyping. Gjengitt fra [47].

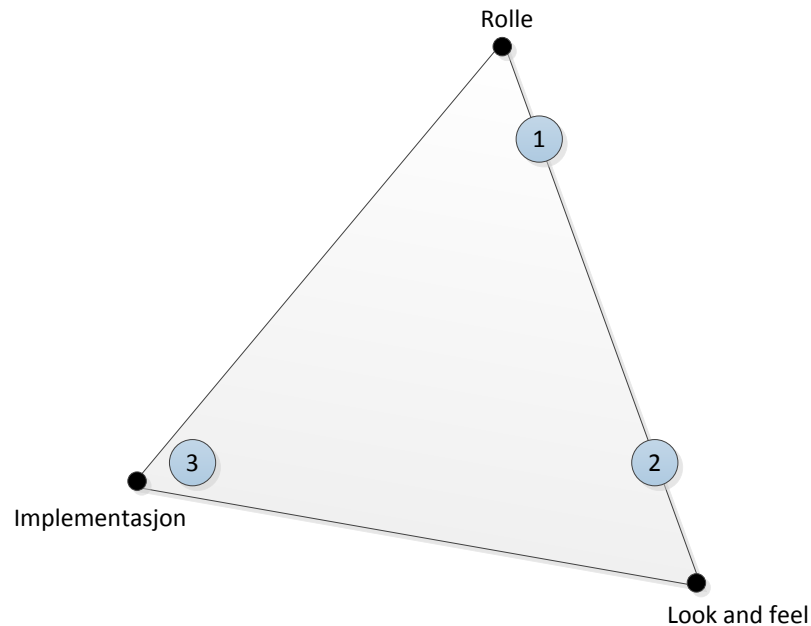
ulike dimensjoner i hjørnene. Disse er *rolle, look and feel* og *implementasjon*. Alle prototyper vil befinne seg et sted inne i dette triangelet, avhengig av hvor mye de forskjellige dimensjonene prototypen har til hensikt å få svar på.

- **Rolle** henviser til spørsmål angående hvilken *funksjon* produktet vil tjene i brukerens liv (hva er det som gjør dette produktet nyttig for brukeren?).
- **Look and feel** angår opplevelsen brukeren har av å bruke produktet (hva brukeren ser på, føler og hører mens han bruker det).
- **Implementasjon** dreier seg om teknikken og komponentene som skal til for at produktet skal utgjøre sin funksjon.

Målet med modellen er å la utviklere skille problemstillinger rundt designet i disse tre klassene av spørsmål, som hver krever ulike tilnærminger til utviklingen av en prototype. En prototype som skal utforske *look and feel* ved et produkt bør eksempelvis simulere den konkrete brukeropplevelsen. Hvis *rollen* til produktet er i fokus bør prototypen etablere en kontekst for bruken av prototypen.

Figur 3.4 viser Houde og Hills modell. Punktene merket med tallet 1, 2 og 3 representerer tre prototyper med ulikt fokus. Prototype 1 fokuserer noe

på look and feel, men aller mest på rollen produktet skal tjene. Prototype 2 fokuserer på det motsatte. Den tredje prototypen utforsker implementeringspotensialet til et produkt, mens den samtidig fokuserer noe på rollen og look and feel.



Figur 3.4: Modell av hva prototyper prototyper. Gjengitt fra [32].

3.6 Evaluering av design mot krav: brukbarhetstester

Ettersom det i brukersentrert utvikling er et stadig ønske om å holde brukeren i fokus er det også svært viktig å la brukerne gjøre gjentatte evalueringer av systemet (“Evaluere design mot krav” i Figur 3.1). Brukbarhetstesting er en mye anvendt form for evaluering, som benyttes av utviklere for å teste hvorvidt produktet de utvikler er brukbart for den tiltenkte brukergruppen [62]. Metoden baserer seg på en kombinasjon av teknikker som inkluderer observasjon, spørreskjema og intervju. Kort sagt går metoden ut på å la potensielle brukere teste det fullstendige systemet eller en prototype av det mens de observeres, samt utspørres i etterkant. Det overordnede målet med brukerevaluering er å måle produktets brukstilfredshet. Får brukeren til å

bruke systemet? Hvor effektivt lar oppgavene seg løse? Og hva er den enkelte brukers opplevelse av systemet? Brukertesting med reelle brukere er en av de mest fundamentale metodene innenfor brukbarhet. Nielsen [47] hevder at dette er en metode som på mange måter er uerstattelig.

3.6.1 Testdeltakere

For å velge ut testdeltakere som er representative er det viktig å finne ut hvem målgruppen er. Å teste mennesker med ulike nivå av teknisk erfaring kan være fornuftig for å avdekke hvilke feil som tar utgangspunkt i selve systemet, og hvilke som skyldes problemer med teknologien i seg selv. Faktorer som alder, kjønn og personlighetstyper er også aktuelle å ta til etterretning når man skal velge ut testdeltakere. Et spørsmål som dukker opp i forbindelse med brukbarhetstesting er hvor mange testpersoner man trenger. Nielsen [48] argumenterer for at det er tilstrekkelig å teste med kun fem personer dersom man har begrensede ressurser. Dette fordi de ytterligere funnene man får ved å øke antall testpersoner, er små i forhold til økningen i ressursbruk. Fem testpersoner vil gjennomsnittlig føre til at rundt 85% av brukbarhetsproblemene blir identifisert. Dette vises i Figur 3.5.

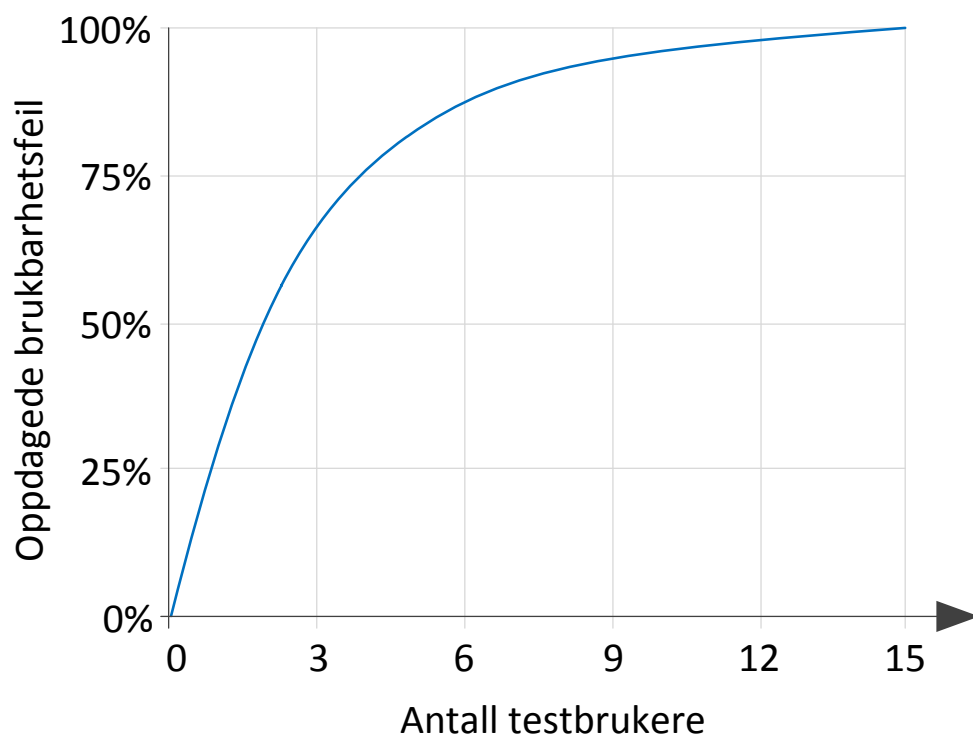
3.6.2 Gjennomføring

Det er mange måter å gjennomføre en brukbarhetstest på og en må dermed tilpasse testene til det produktet man ønsker å teste. Det skilles mellom ulike typer brukbarhetstester som alle benyttes til hvert sitt formål. *Utforskende tester* samler inn kvalitativ data og brukes på et tidlig stadium for å vurdere effektiviteten av tidligere modeller. *Vurderingstester* samler inn både kvalitativ og kvantitativ data og gjennomføres enten tidlig eller midtveis i designprosessen. Den siste typen det skilles mellom er *verifiseringstester* som baserer seg på kvantitativ datainnsamling og er en objektiv sertifisering av produktets brukbarhet.

DECIDE er et mye brukt rammeverk for å planlegge brukbarhetstester og baserer seg på følgende punkter [62]:

1. **Bestemme målene for evalueringen**

- For å få fullt utbytte av testingen er det hensiktsmessig å vite hva de overordnede målene med testen er, slik at man unngår å gå glipp av viktig informasjon underveis.



Figur 3.5: Brukbarhetsfeil som avdekkes i forhold til antall testbrukere.

2. Velg teknikker for å besvare spørsmålene

- Hvordan skal testene gjennomføres slik at man kan besvare spørsmålene? Det blir ofte benyttet video, spørreundersøkelser og intervjuer i forbindelse med testene. En kjent teknikk er å bruke et såkalt SUS-skjema (se Seksjon 3.6.3).

3. Forberede oppgaver som skal besvares under testen

- Det er viktig å avgjøre hvilke oppgaver som skal benyttes i testen. Som regel utarbeides det et sett med definerte oppgaver som skal være med på å dekke målene som er satt for testen. Her er det spesielt viktig å tenke på at oppgavene skal være så enkle at de tolkes likt, uavhengig av testperson.

4. Valg av testdeltakere

- I utvelgelsen av testdeltakere må man sørge for at man gjør et bredt utvalg deltakere for å sikre at de mest aktuelle målgruppene er representert.

5. Forbered testforholdene

- Når man skal gjennomføre brukbarhetstester er det vanlig at disse finner sted på et brukbarhetslaboratorium fordi man gjerne ønsker å benytte seg av videoopptak og skjulte observatører.

6. Planlegge gjennomføringen

- For at en brukbarhetstest skal gjennomføres på beste vis er det en nødvendighet å være godt forberedt. Utstyr skal settes opp, teststegene skal gjennomgås og instruksjonene bør være klare i forkant av testen.

7. Behandle etiske tema

- Man bør være oppmerksom på at en brukbarhetstest kan oppleves som belastende på deltakerne og at man dermed må være klar over at deltakerne kan velge å avbryte når de vil. Deltakerne skal også informeres om at de observeres og om eventuelle opptak som blir gjort.

8. Evaluer, analyser og presenter data

- For å evaluere brukbarhetstestene kan man generere grafer ut fra resultatene man oppnår gjennom testene, som antall utførte tester, hvor lang tid testene tok og resultater fra spørreundersøkelser. Videomateriale er også svært nyttig til å foreta analyser av testene i etterkant.

3.6.3 System Usability Scale

System Usability Scale (SUS) ble utviklet av John Brooke med formål om å måle brukbarheten til et system ut fra en subjektiv bedømmelse [6]. Et SUS-skjema er et spørreskjema bestående av 10 påstander som vurderes ut fra en Likert-skala [78] der testpersonene skal angi hvor enige de er i påstandene på en skala fra 1 (sterkt uenig) til 5 (sterkt enig). Til sammen vil dette utgjøre en sluttsum som gir en indikasjon på hvorvidt hver testperson anser systemet som brukervennlig eller ikke.

Skjemaet skal i følge Brooke fylles ut av testdeltakerne etter at testen er gjennomført og deltakerne har fått testet systemet som evalueres, men før debriefing eller diskusjon finner sted. Et problem med denne typen vurdering er at ettersom den fullstendige poengsummen regnes ut fra alle de ti punktene vil ethvert ubesvart eller feilbesvart spørsmål ikke bare påvirke det enkelte spørsmålet, men også skalaen som en helhet [22].

Kapittel 4

Interaksjonsteknologier

Dette kapitlet gir en innføring i interaksjonsteknologier, med hovedfokus på multi-touch. Innledningsvis gis det en innføring i evolusjonen til interaksjonsteknologier. Det fokuseres videre på hva multi-touch innebærer i Seksjon 4.3 og hvordan det implementeres på ulike måter. Seksjon 4.4 tar så for seg multi-touchbord generelt og Surface spesielt.

4.1 Fra CLI til NUI

Siden datamaskinen ble oppfunnet har fokuset på brukerens interaksjon med maskinen blitt stadig større og vi kan se tilbake på en evolusjon av forskjellige interaksjonsparadigmer.

4.1.1 CLI

CLI (Command-line Interface) er en måte å kommunisere med datamaskinen på der brukeren skriver tekstlige kommandoer til maskinen ved å ta i bruk et tilkoblet tastatur. CLI oppstod på 50-tallet da man begynte å koble fjernskrivere til datamaskiner for at brukeren kunne ha en interaktiv dialog med datamaskinen fremfor batch-prosessering med hullkort [77]. Det som kjenner seg ut for CLI er at brukeren må lære seg og memorere tekstlige kommandoer som må skrives for at maskinen skal utføre operasjoner. Dette medfører en høy læringskurve og gjør interaksjonen vanskelig for ikke-tekniske personer.

4.1.2 GUI

Douglas Englebart holdt i 1968 en demonstrasjon for over tusen dataprofesjonelle som senere har gitt ham tilnavnet “the father of GUI (Graphical User Interface)” [59]. Englebart viste en dataskjerm med vektorgrafikk som han opererte på med et tastatur, datamus og ytterligere fem knapper. Ideene til Englebart ble videreført av Xerox PARC som i 1973 utviklet datamaskinen *Alto* som for første gang demonstrerte bruken av skrivebordsmetaforen og et grafisk brukergrensesnitt. Det grafiske brukergrensesnittet Xerox PARC utviklet bestod av vinduer, ikoner, menyer og en peke-enhet; populært forkortet WIMP (window, icon, menu, pointing device). WIMP gjorde interaksjonen enklere for brukeren ved å redusere den kognitive lasten det var å kunne tekstlige kommandoer. Ved å bruke en datamus kunne brukeren bevege en peker fritt på skjermen og klikke på ikoner som representerte filer og utføre kommandoer på disse fra menyer. *Alto* var ikke et kommersielt produkt, men ga inspirasjon til design av senere datamaskiner. Da Apple lanserte *Lisa* i 1983 og *Macintosh* året etter ble GUI for alvor kommersialisert som interaksjonsmåte. GUI-et ble videreutviklet av en rekke senere datamaskiner og ble vanlig i operativsystemer etter hvert som datamaskinene ble dagligdagse i husholdninger og kontorer. GUI er i dag en del av de aller fleste operativsystemer, men tekniske brukere kan fortsatt kommunisere med maskinen via CLI.

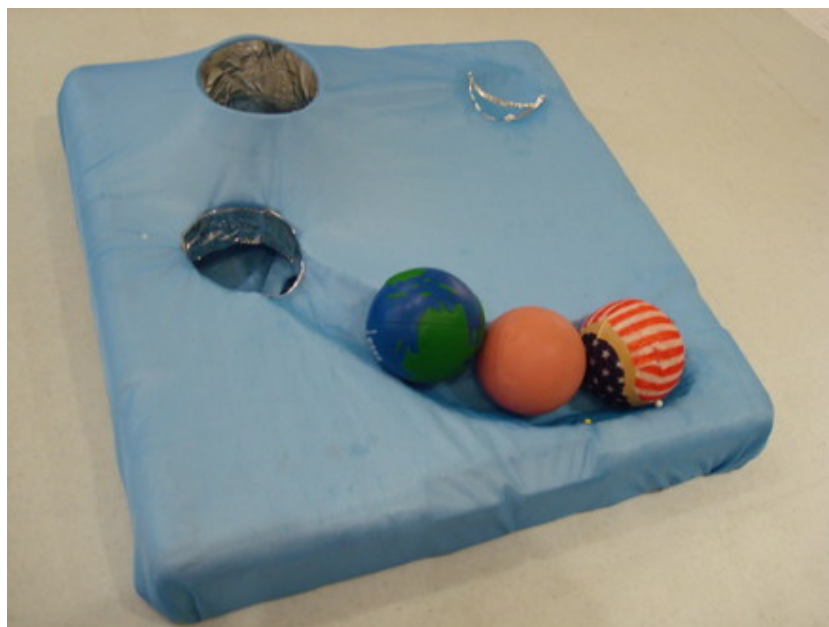
4.1.3 NUI

De siste årene har NUI (Natural User Interface) blitt et stadig mer aktuelt begrep. NUI er brukergrensesnitt som utnytter menneskelige evner som berøring, bevegelser, syn, stemme og hørsel til å kommunisere med datamaskiner på [28]. Essensen er at brukergrensesnittet skal la brukeren gjøre handlinger med kroppen som er naturlige for oppgaven. I stedet for å bruke tradisjonelle verktøy som tastatur og datamus, som må læres før de benyttes, kan brukeren eksempelvis velge elementer på skjermen ved å ta direkte på disse med fingeren (direkte manipulasjon). Andre brukergrensesnitt lar brukeren styre maskinen utelukkende ved å bruke stemmen, som eksempelvis *Windows Speech Recognition* [15]. *Xbox Kinect* [83] viser en annen innovativ interaksjonsmetode der personer bruker hele kroppen til å spille videospill. Det finnes en rekke andre måter å interagere med datamaskiner på som går under betegnelsen NUI. Fellesnevneren er at brukeren skal kunne gå fra nybegynner til ekspert på kort tid.

Med introduseringen av iPhone, Surface og andre små og store enheter med multi-touch, har NUI i form av direkte manipulasjon fått vidstrakt fokus de siste årene. Hensikten med multi-touch er å gi brukeren en naturlig og direkte måte å interagere på. Gestures som “pinch”, “drag”, “flick” og “tap” [74] skal gjøre det intuitivt for brukeren hvordan elementer på touchskjermen kan behandles. Eksempler på den korte læringskurven for slike brukergrensesnitt kommer til syne i flere youtube-videoer der barn ned i ettårsalderen anvender iPhone [9][11].

4.2 TUI

Surface og nyere multi-touchbord har muligheten til å interagere med fysiske objekter. Denne formen for interaksjon kalles Tangible User Interface (TUI). TUI er et brukergrensesnitt der en person kan interagere med digital informasjon gjennom et fysisk miljø [81]. Et av de mest kjente og interessante eksemplene på et TUI er Durell Bishops *Marble Answering Machine* (se Figur 4.1) [65]. Dette var et brukergrensesnitt for en telefonsvarer der hver innkommende talemelding var representert av en fysisk klinkekule som trillet ut av maskinen. For å høre talemeldingen måtte brukeren plassere klinkekulen i en grop på maskinen.



Figur 4.1: Durell Bishops *Marble Answering Machine*.

Et åpent rammeverk som definerer et felles API for multi-touchskjermer med TUI er TUIO (Tangible User Interface Objects) [72]. TUIO er en enkel, men samtidig allsidig, protokoll som definerer felles egenskaper ved objekter på multi-touchflater. Protokollen oversetter det som fanges opp på skjermen og sender dette til klientapplikasjoner som kan dekode protokollen.

4.3 Multi-touch

Begrepet *multi-touch* innebærer en overflates evne til å gjenkjenne to eller flere kontaktpunkter samtidig [80]. Personer kan på en multi-touchoverflate bruke flere fingre simultant til å interagere med et system. Overflaten kan være en *trackpad* eller *touchskjerm*. En trackpad brukes som en inputenhet til datamaskiner og oversetter kontakt og bevegelse på overflaten til en relativ posisjon på skjermen. Slike er vanlige å finne på laptopen som et alternativ til datamus. Touchskjermer skiller seg fra trackpads ved at detekteringen av kontakt og bevegelse skjer direkte på displayet. På denne måten kan brukere interagere med et brukergrensesnitt ved ta på GUI-elementene på skjermen.

Multi-touch legger til rette for bruk av *gestures*. Gestures er finger- og håndbevegelser på touch-overflaten som gjør at brukeren kan gjøre mer avanserte funksjoner enn simple “trykk” i brukergrensesnittet. En komplett guide til kjente typer gestures kan studeres i Wroblewski et al. sin “*Touch Gesture Reference Guide*” [74]. Gesturen “pinch” kan eksempelvis brukes til å skalere bilder eller zoome et kart inn og ut i en applikasjon.

Når multi-touch åpner for at brukergrensesnitt kan motta mange kontaktpunkter samtidig betyr det også at flere brukere kan interagere på samme skjerm. Dette er mest praktisk på større skjermer som har plass til flere personer omkring den. Multi-touchbord (se Seksjon 4.4) og multi-touchvegger er spesielt egnet for bruk av flere personer samtidig. De største tabletene, som Apple iPad [36], HP TouchPad [13] og Motorola Xoom [39], har også et potensial for dette.

4.3.1 Teknologier

Selv om det finnes flere ulike teknikker for å oppnå touchsensitive skjermer er det hovedsakelig to metoder som implementerer ekte multi-touch [51].

En *prosjektet kapazitiv* skjerm utnytter den elektriske egenskapen *kapasitans* til å måle hvor elektriske ledere, som eksempelvis fingre, berører skjermen ved

hjelp av et elektrisk ledende lag under displayet som er oppdelt i form av et koordinatnett. Prosjekterte kapasitive skjermer blir brukt i flere håndholdte enheter, deriblant nyere smarttelefoner, som eksempelvis Apple iPhone.

Optiske touchskjermer bruker en eller flere kameraer til å observere berøring på skjermen, som er en delvis gjennomsiktig flate. Kameraet kan være plassert over eller under skjermen. For at skjermen ikke skal la seg påvirke av omgivende lys bruker flere optiske touchskjermer infrarøde lyskilder og kameraer. Disse lyskildene og kameraene kan organiseres på ulike måter. Én slik måte er å plassere både lyskilden og kameraene på baksiden av skjermen vendt mot den (se Figur 4.2). Det infrarøde lyset vil da rømme gjennom skjermen. Når fingre og andre objekter berører skjermen vil kameraene fange opp det reflekterte lyset fra kontaktpunktene. En prosjektør eller LCD-skjerm på baksiden av den delvis gjennomsiktige flaten skaper skjermbildet for brukeren. For at kameraene skal kunne se hele kontaktflaten trenger optiske touchskjermer større plass enn kapasitive skjermer. Teknologien er derfor ofte å finne i form av touchbord og -vegger.

4.3.2 Forskning og utvikling av multi-touchbord

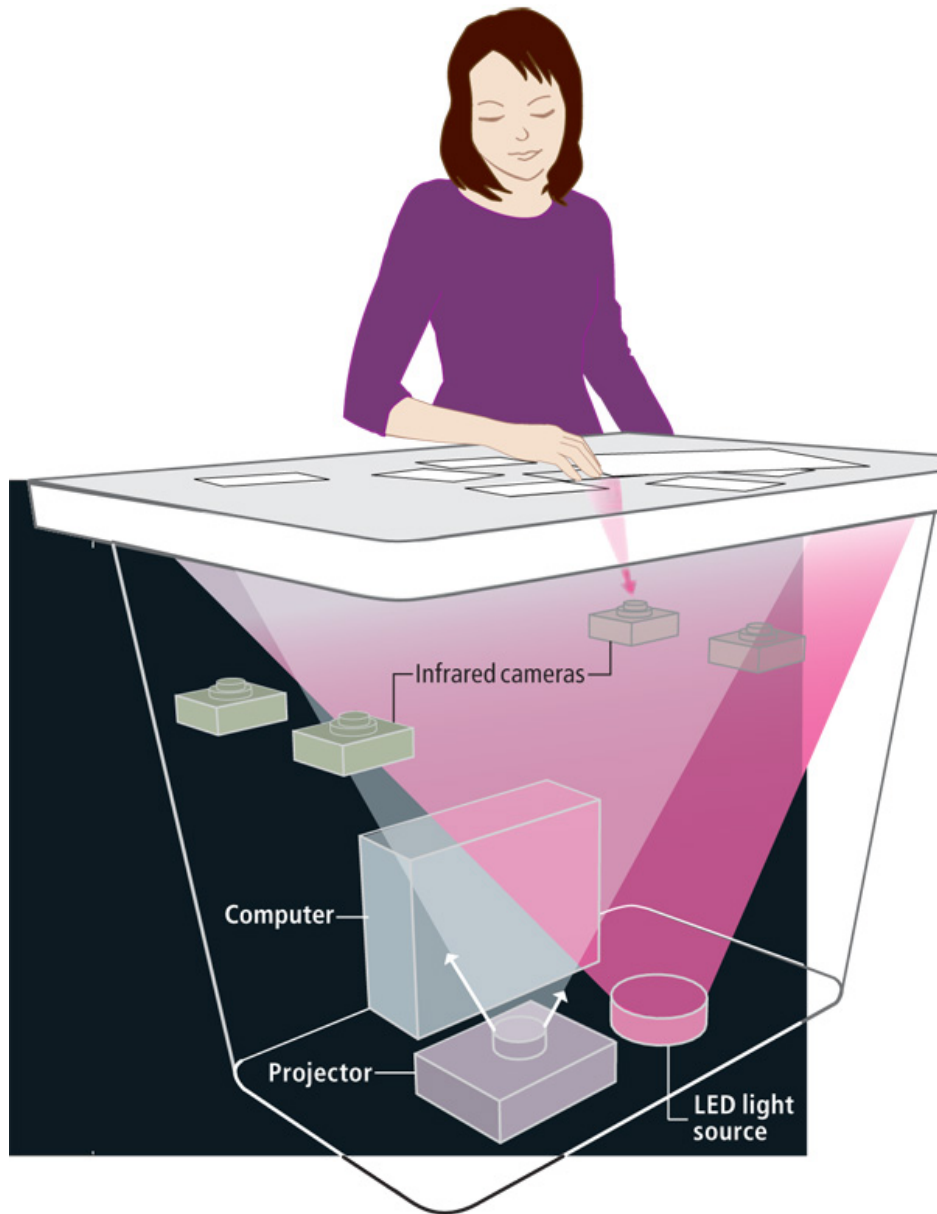
Arbeidet med forskning og design av produkter med multi-touch har pågått siden begynnelsen av 80-tallet. Som nevnt innledningsvis (se Seksjon 1.1.1) regnes Bob Boie ved Bell labs for å være den første som utviklet en skjerm med multi-touch [8]. Denne skjermen brukte et gjennomsiktig, kapasitivt nett av sensorer som var lagt over et CRT¹-display. Skjermen lot brukeren manipulere grafiske objekter ved bruk av fingrene med en utmerket responstid.

Pierre Wellner publiserte i 1991 paperet "*Digital Desk*" som presenterte et frontprosjektert multi-touchbord som brukte optiske og akustiske teknikker til å oppdage hender/fingre, samt fysiske papirbaserte kontrollere [75]. Dette har blitt et klassisk paper innen fagfeltet.

The Input Research Group fra University of Toronto demonstrerte i 1995 konseptet med, og senere implementasjonen av, et digitalt PC-display som kunne sanse identiteten, plasseringen og retningen til fysiske objekter på skjermen. Den resulterende artikkelen introduserte forestillingen av tangible multi-touchskjermer [24].

Senere kommersielle multi-touchbord er basert på forskningen som ble gjort på 80- og 90-tallet og har mange likhetstrekk.

¹Cathode Ray Tube



Figur 4.2: Optisk multi-touchskjerm. Illustrasjon hentet fra Scientific American [3].

4.4 Kommersielle multi-touchbord

Da Microsoft kunngjorde multi-touchbordet Surface i 2007 (se Seksjon 1.1.1) markerte dette starten på de kommersielle multi-touchbordene. Det finnes ingen entydig definisjon på hva et multi-touchbord er, men det ligger i begrepet at det innebærer en multi-touchskjerm som er plassert på ben eller en kropp slik at det opptrer som et bord. Brukere kan sitte eller stå rundt bordet og interagere på den samme skjermen. Flere av produsentene, deriblant Microsoft, oppfordrer indirekte gjennom promosjonsvideoene sine til å bruke multi-touchbord som et vanlig bord. Eksempelvis kan det brukes til å hvile drinken sin på i en bar, samtidig som mulighetene er der for å bruke applikasjoner på bordet. På denne måten hevder de å inkorporere PC-en i hverdagsbordet og bidra til såkalt “Ubiquitous computing” [82]. Som de sier i en promosjonsvideo for Surface: *“The possibilities seem endless, as the line between the virtual world and the physical world becomes increasingly thin”*.

Andre produsenter fulgte i Microsofts fotspor og lanserte egne multi-touchbord etter introduksjonen av Surface, deriblant Ideum MT-50 [33] og Evolve One [21]. Disse bordene lignet svært mye på Surface og brukte samme teknologi for å oppnå multi-touch.

Microsoft, Ideum og Evolve har nylig kunngjort oppdaterte generasjoner av multi-touchbordene sine, henholdsvis Surface 2.0, MT-55 og Evolve Two. Trenden er at skjermene blir større og tynnere, samtidig som kroppen erstattes av ben slik at skjermen også kan demonteres og henges på veggen.

4.4.1 Microsoft Surface

I denne oppgaven fokuseres det på utforskning av brukersentrerte metoder for multi-touchbordet Surface (se Figur 4.3). Surface ble kunngjort i mai 2007 og lansert 17. april året etter [79]. Den har en 30" overflate med en XGA DLP prosjektør på innsiden av kroppen som projiserer skjermbildet mot undersiden av flaten slik at brukeren ser det fra oversiden. En LED-lyskilde kaster infrarødt lys mot skjermflaten og fem kameraer fanger opp refleksjon fra fingre og objekter som berører den. Multi-touchbordet klarer å reagere på 52 slike kontaktpunkter samtidig.



Figur 4.3: Microsoft Surface.

Superrealisme

Med Surface prøver Microsoft å oppnå noe de kaller *superrealisme* [14]. Superrealisme forsøker å etterligne interaksjonsmønstre fra virkeligheten, og samtidig drive interaksjonen et steg lenger enn hva som er fysisk mulig. Eksempelvis kan du flytte rundt på et bilde både i virkeligheten og på Surface ved å ta direkte på bildet og skyve (se Figur 4.4). På Surface kan du imidlertid i *motsetning* til virkeligheten skalere bildet opp og ned ved å bruke to fingre på bildet og skyve fra eller mot hverandre (se Figur 4.5). På denne måten kan brukeren oppleve en “better than real” kvalitet.



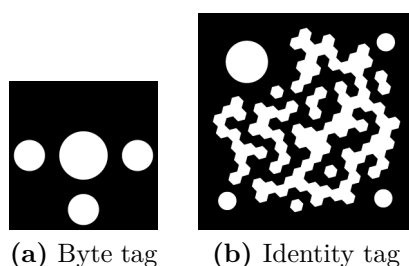
Figur 4.4: En bruker tar på et bilde og flytter det rundt. Illustrasjon hentet fra Surface User Experience Guidelines [14].



Figur 4.5: En bruker skalerer bildet. Illustrasjon hentet fra Surface User Experience Guidelines [14].

Surface tags

Surface kan også gjenkjenne fysiske objekter. Til dette bruker man såkalte “*Surface tags*” (heretter kalt tags). Tags er spesiellagede strekkoder som Surface klarer å gjenkjenne når de kommer i kontakt med skjermflaten. Det finnes to typer tags; *byte tags* (Figur 4.6a) og *identity tags* (Figur 4.6b). Byte tags lagrer 8 bits med data (1 byte) som betyr at det finnes 256 unike varianter [42]. De er svarte kvadrater på ca 2x2 cm fylt med et mønster av hvite prikker som danner byte-verdien. Identity tags lagrer 128 bits (2 x 64), som praktisk talt gir et uendelig utvalg av verdier. De er større enn byte tagene og inneholder en annen mønsterform. Tagene kan skrives ut på vanlig papir. Når de legges på skjermflaten reflekterer hvitfargen infrarød stråling og Surface vil gjenkjenne mønsteret. Den vil også vite hvor på skjermen tagen befinner seg og hvilken retning den peker i. Utviklere kan anvende tags ved å feste de på fysiske objekter og fortelle applikasjonen hva som skal skje når den tagen oppdages på skjermen.



Figur 4.6: Surface tags i riktig størrelsesforhold.

Tags kan brukes til en rekke forskjellige formål i en Surface-applikasjon. Noen eksempler på virkelige applikasjoner som utforsker bruken av tags er:

- Tagging av bøker hos en bokhandler. Når ei bok plasseres på skjermen

vil informasjon om den være tilgjengelig for brukeren [17].

- Tagging av visittkort. Kortet bringer blant annet frem eierens personalia og flickr²-konto [57].
- Et spill der konkurrenter kaster små puter med tags på mot en blink på skjermen [58].
- Stempling av forskjellige symboler på bilder ved å bruke fysiske stempler med tags [29].

Denne egenskapen ved Surface gjør det, ved siden av å være en multi-touchskjerm, også til et Tangible User Interface (se Seksjon 4.2).

Andre tekniske potensialer

Ved siden av å være et multi-touchbord fungerer Surface som en ordinær PC med de samme mulighetene man kan forvente av nyere datamaskiner. Den har innebygget høyttalere, Bluetooth 2.0, USB-porter og trådløst Internett. Utviklere kan utnytte disse egenskapene til å skape rikere opplevelser for brukeren i Surface-applikasjoner. Bluetooth kan eksempelvis brukes til å sende innhold mellom Surface og eksterne enheter som mobiltelefon og PC. Ved å koble en RFID-leser til Surface åpnes muligheten for å kommunisere med fysiske gjenstander som inneholder en RFID-sender (eksempelvis adgangskort og transportbevis).

Surface 2.0

6. januar 2011 kunngjorde og demonstrerte Microsoft den nye versjonen av Surface, kalt Surface 2.0. Gjennom den nye 40" store og 10 cm tykke skjermen Samsung SUR40 har den både større skjermflate og adskillig tynnere design, noe som gjør at den kan stå på bordben eller festes på veggen. Denne LCD-skjermen har dessuten 1080p oppløsning.

Det som gjør Surface-oppfølgeren så tynn er at den bruker en ny teknologi for å detektere kontakt med skjermen. Microsoft har kalt denne teknologien *PixelSense*. PixelSense er bygget på forskning rundt tynne, optiske multi-touchskjermer (blant annet ThinSight [31]) og fungerer ved at pikslene i skjermen har integrerte optiske sensorer [8]. Dette gjør at skjermen kan "se"

²<http://www.flickr.com>

alt som blir plassert på den (eksempelvis fingre, tegninger, strekkoder, tekst) tilsvarende en skanner.

Kapittel 5

Forskningsmetoder

Dette kapitlet tar for seg de ulike forskningsmetodene som er relevante for forskningen. Teorien er basert på Briony J. Oates sin bok “*Researching Information Systems and Computing*” [52]. Dersom ikke noe annet er sagt er alle siteringer og fakta hentet fra denne boken.

5.1 Kvalitative og kvantitative studier

Valget av forskningsmetoder er med på å påvirke hvordan forskere samler inn og evaluerer data. Forskningsmetoder kan klassifiseres som kvalitative og kvantitative, hvor hver metode er formålstjenlig i ulike situasjoner. Kvantitative forskningsmetoder har fokus på store mengder tallbaserte data. Dette er typisk data som kan være generert ut fra spørreundersøkelser, eksperimenter, matematiske modelleringer eller andre forskningsmetoder der dataene analyseres for å se etter mønster som kan føre til konklusjoner.

Kvalitative forskningsmetoder har til følge å benytte seg av kvalitative data, som betegnes som all data som ikke er tall og som gjerne kommer fra intervju, observasjon og dokumenter. Som oftest kommer disse dataene til syne gjennom *etnografi*, *action research* og *casestudier* [52][45]. Etnografi er en forskningsmetode som baserer seg på førstehåndsobservasjon og gjøres ved at forskeren tar del i en ukjent kultur over en lengre periode [64]. I Action Research er forskeren aktivt involvert, og må observere og delta i det gitte fenomenet under studiet. Casestudier skiller seg fra etnografiske studier og action research ved at forskeren baserer datainnsamlingen sin på intervjuer, supplert med dokumenter i stedet for observasjoner [44].

Når man forsker på informasjonssystemer og den konteksten de befinner seg i er det vanlig å bruke kvalitative metoder, nettopp fordi forskning på dette ofte innbefatter forståelsen rundt bruken av systemene. Michael D. Myers sier blant annet: *“qualitative research methods are designed to help researchers understand people and the social and cultural context within which they live”* [45]. Når man forsker på informasjonssystemer er det ofte forståelsen rundt bruken som er utgangspunktet for forskningen, og kvalitative metoder er derfor godt egnet.

5.2 Eksperiment

I følge Oates kan et eksperiment innenfor akademisk forskning defineres som en søken etter å bevise eller motbevise en kausal kobling mellom en faktor og et observert resultat [52]. Metoden går ut på at man starter med en hypotese og empirisk tester denne gjennom forsøk. Oates nevner fire design for eksperimenter:

- En gruppe, pre-test og post-test
- Statisk gruppesammenligning
- Pre-test/post-test kontrollgruppe
- Salomo fire-gruppes design

Statisk gruppesammenligning er den mest aktuelle tilnærmingen når man skal sammenligne forskjeller mellom grupper basert på ulik påvirkning. Her blir deltakerne delt inn i to grupper som utsettes for forskjellig stimuli. Videre måles oppførselen til de to gruppene. Når man benytter seg av statistisk gruppesammenligning kan ulikhetene i oppførselen mellom gruppene forklares ut fra stimuli. Man må likevel være oppmerksom på at dersom deltakerne ikke blir tilfeldig utvalgt til de to gruppene kan ulikhetene skyldes andre faktorer enn stimuli.

5.2.1 Variabler

Når man utfører et eksperiment er det viktig at man gjør seg bevisst på hva som er en avhengig og en uavhengig variabel, og hvilke av disse variablene som skal være konstante og hvilke som skal endres. Variabler som deltakere, kunnskapsnivå og tilgjengelig informasjon skal være konstant.

5.2.2 Måling

Før man foretar et eksperiment er man nødt til å vite hva man er ute etter å måle og hvordan. For å kunne avgjøre hva som er et resultat av eksperimentet må man gjøre målinger både på forhånd (pre-test) og i etterkant (post-test) av eksperimentet. Her er det typisk å bruke kvantitative data, siden man er nødt til å måle ut fra statistiske data.

5.2.3 Validitet

Når det gjelder eksperimentets validitet skilles det mellom intern og ekstern validitet. Intern validitet er i hvilken grad resultatene er gyldige for det utvalget og det fenomenet som er undersøkt. De mest vanlige fallgruvene er:

- **Ulikheter:** Dersom gruppene er ulike til å starte med vil ikke den målte forskjellen nødvendigvis kunne tillegges manipuleringen av den eksperimentelle gruppen.
- **Historie:** Hendelser mellom pre- og post-test kan ha innvirkning på deltakerne slik at resultatet påvirkes.
- **Modning:** Deltakerne modnes mellom testene.
- **Instrumentering:** Defekte instrumenter som medfører feil eller unøyaktige målinger.
- **Svikt:** Deltakere faller fra, noe som kan medføre at de gjenværende deltakerne blir påvirket i større eller mindre grad.
- **Påvirkning:** Endret oppførsel kan for mange være en reaksjon på å bli testet. For eksempel kan faktorer som alder, rase eller kjønn ha innflytelse på deltakerne. Blant annet kan de endre oppførsel fordi de vet at de blir testet.

Ekstern validitet er i hvilken grad resultatene kan være holdbare for situasjoner utover de som ble forsket på i den gitte undersøkelsen:

- **Avhengighet av en spesiell type deltakere:** Resultater som er funnet gjennom eksperimenter som involverer frivillige deltakere kan ikke nødvendigvis generaliseres til en større del av befolkningen.

- **Før få deltakere:** Dette kan i alle tilfeller føre til at det er umulig å bevise at et resultat er statistisk signifikant.
- **Ikke representative deltakere:** Det er viktig å sørge for at man får deltakere som er representative for den gruppen hypotesen er fremstilt for.
- **Ikke representative test case:** Casene som brukes i eksperimentet må være utformet på en slik måte at de kunne vært brukt i virkeligheten.

5.3 Intervju

Målet med intervju vil være å få dypere innsikt i situasjonene enn de man finner ved dokumentanalyse og observasjon. Det er også her et faktum at informasjonen man får gjennom et intervju vil være svært subjektiv. Den er basert på den enkeltes oppfatning av hvordan ting er, og det gjør at man skal tenke seg ekstra godt om hva denne informasjonen brukes til. Ved å intervju flere personer og sammenligne resultatene fra disse, vil det kunne trekkes mer valide slutninger.

Intervju deles inn i de tre typene *strukturert*, *semi-strukturert* og *ustrukturert*, der hver type brukes til ulike formål. Et *strukturert* intervju baserer seg på at det foreligger en rekke forhåndsdefinerte spørsmål som brukes for alle intervjuobjektene i en gitt rekkefølge. *Semi-strukturerte* intervjuer skiller seg fra strukturerte intervjuer ved at det er rom for endring. Rekkefølgen er ikke fastsatt og det er lov til å stille tilleggs spørsmål for å støtte opp under interessante påstander fra intervjuobjektet. I et *ustrukturert* intervju tar intervjueren opp tema som det skal snakkes rundt, noe som fører til at intervjuobjektet får snakke fritt i mye større grad enn dersom intervjuet hadde hatt fastsatte rammer.

Et semi-strukturert intervju bærer mer preg av en dynamisk interaksjon mellom intervjuer og den som blir intervjuet, hvor alle har anledning til å styre forløpet ut fra hva som fremtrer som interessant under samtaleprosessen. Fordelen med dette er at de som blir intervjuet får stor frihet til å presentere sin individualitet [84]. Svakheten er at det lett kan bli vanskelig å sammenligne og tolke ulikheter mellom flere intervjuobjekter ettersom man ikke har noen gitte føringer på hva som skal besvares. Oates peker på at kjønn, alder, etnisk opprinnelse, aksent og status kan være med på å influere respondentene når de avgjør hvilken informasjon de gir deg. Både Myers og Oates

peker på at ulik bakgrunn, alder og kunnskapsnivå mellom intervjueren og intervjuobjektet kan være med på å prege intervjuet.

Intervjuer kan foregå enten individuelt eller i grupper. I et gruppeintervju er det vanlig å ha mellom tre til seks deltakere. Gruppeintervju er preget av interaksjon mellom personene og kan dermed føre til en mer effektiv tilgang til flere aktørers synspunkter på et tema. Ved å gjennomføre gruppeintervju vil man få samlet inn data fra flere personer på en gang, noe som gjør det til en svært effektiv metode. Noen av ulempene relatert til gruppeintervjuer er at enkelte av intervjuobjektene kan være mer dominerende enn andre og forhindre at alle slipper til med sine synspunkter. Videre kan det lett forekomme at de svarene som kommer opp under intervjuet er svar som blir ansett å være “akseptable” innenfor gruppen. Gruppetenking er også en risiko ved gruppeintervjuer, at behovet for å oppnå enighet blir så stort at man overskygger motivasjonen til kritisk tenkning.

5.4 Observasjon

Målet med observasjon er å få et komplett og nøyaktig bilde av en situasjon, og ikke å teste hypoteser som er utformet på forhånd. En viktig fordel med observasjon er at det skaper en rik innsikt i hvordan ting foregår i sosiale settinger, uten å være avhengige av at personene er bevisste på sin egen praksis og kan rapportere denne [23]. Sammenlignet med intervjuer får forskeren her innblikk i hvordan mennesker *faktisk* handler og oppfører seg i bestemte situasjoner, og ikke bare hvordan de *beskriver* at de oppfører seg i disse situasjonene. Ved observasjon får man flere nyanser av det store bildet, som hvordan den fysiske situasjonen er, og alle detaljene en beskrivelse ikke vil kunne dekke.

Å observere noen fører også med seg en rekke utfordringer. Blant annet gjør man seg avhengig av en subjektiv oppfatning av det som skjer. Vi har alle en selektiv hukommelse som husker noe og glemmer andre ting. Alt dette er avhengig av hvilke erfaringer hver enkelt person har og hvilken tilstand man er i under observasjonen.

Observasjon er en særdeles billig metode i den grad at det krever få ressurser for å gjennomføre det. Det er derfor en enkel måte å finne nye data på og kan danne et grunnlag for senere forskning. Metoden egner seg godt når problemstillingen er knyttet til et mer avgrenset og overkommelig geografisk område. Effekten av å bli studert kan imidlertid påvirke utfallet hos personen som blir studert. Denne ulempen ved observasjon kalles “Hawthorn-effekten”

– mennesker endrer bevisst og ubevisst oppførsel fordi de vet at de blir observert.

En meget anerkjent forsker innenfor sosialpsykologien, Daryl Bem, mente at når vi observerer andres atferd gjør vi det på samme måte som når vi attribuerer vår egen atferd [4]. Dette er noe som lett kan få innvirkning på de resultatene man ender opp med etter en observasjon og som det er viktig å ta til etterretning.

5.5 Spørreskjema

Spørreskjema er en godt innarbeidet teknikk for å samle inn demografiske data og brukernes meninger på [62]. Oates definerer et spørreskjema som “et forhåndsdefinert sett spørsmål som er satt sammen i en bestemt rekkefølge”. Respondentene blir spurt om å svare på spørsmålene, noe som vil skaffe forskeren data som kan analyseres og tolkes. Metoden er derfor godt egnet i tilfeller hvor man ønsker å samle inn spesifikk informasjon [62].

Det er viktig at spørreskjemaet utarbeides nøye, på en slik måte at det er mulig å generere pålitelige data. Tolkningen av de ulike spørsmålene kan ha stor innvirkningene på resultatene av spørreundersøkelsen. Derfor er det viktig å formulere spørsmålene på en slik måte at de oppfattes likt. Metoden fører med seg ulemper som at man ikke kan korrigere eventuelle misforståelser som har oppstått hos respondentene.

Måten spørreskjemaene er utarbeidet på er altså svært viktig med tanke på de resultatene man får ut av det. Kvaliteten på informasjonen man oppnår fra et spørreskjema har direkte kobling med kvaliteten på selve spørreskjemaet [55]. Det er mange måter å designe et spørreskjema på. Oates lister blant annet opp følgende:

- Ja/nei-svar
- Kvantitative spørsmål
- Enig/uenig utsagn
- Grad av enig eller uenig ved bruk av en Likert-skala

5.6 Research through design

Zimmerman et al. [85] foreslår bruk av “Research through design” som metode for forskning på interaksjonsdesign innenfor MMI. Metoden lar interaksjonsdesignere bidra til forskningen ved å gjøre det de kan best; å studere verden, for så å designe ting som skaper positiv endring. Ved å designe artefakter (prototyper, mockups, produkter) i et forsøk på å gjøre de *riktige* løsningene kan en få viktig kunnskap om utfordringene ved design av lignende artefakter og hvordan de kan løses (designmønstre). Research through design er dermed et forsøk på fylle gapet mellom teori og praksis i interaksjonsdesign.

Zimmerman poengterer at når denne metoden skal brukes er det viktig å skille mellom *forskningsartefakter* og *designpraksisartefakter*. Hvis artefaktene som designes skal bidra til forskningsteori (forskningsartefakter) må (1) hensikten med artefakten være å gjøre designet riktig, og ikke en kommersiell suksess, og (2) artefakten må demonstrere nyskapning. Eksempelvis kan en prototype for en kopimaskin demonstrere hvordan det er mulig for novice brukere å reparere papirstopp i kopimaskiner og ikke bare fagfolk, og på denne måten bidra til teoretisk kunnskap rundt nye måter å interagere med maskiner på.

For å kunne bruke research through design som forskningsmetode er man nødt til å oppfylle følgende kriterier:

- **Prosess:** Studiet må oppgi nok detaljer rundt prosessen som ble brukt, slik at det er mulig å gjenskape den. Det må i tillegg oppgis en logisk forklaring for de spesifikke metodene som ble brukt.
- **Nyskapning:** Forskningsbidraget må utgjøre en signifikant nyskapning i fagfeltet. Dette må begrunnes gjennom et litteraturstudie.
- **Relevans:** Det må utledes hvordan den nåværende situasjonen er og hvordan det ønskes at den skal være. Det må også støttes opp under hvorfor fellesskapet skal ønske det samme.
- **Utvidningsevne:** Det må være mulig å bygge videre på resultatene fra forskningen; enten ved å bruke prosessen i et senere designproblem eller ved å forstå og øke kunnskapen som blir skapt gjennom de resulterende artefaktene.

5.7 Triangulering

Metodetriangulering innebærer at bestemte fenomener studeres fra ulike synsvinkler og synspunkter, og at problemstillingen belyses ved hjelp av forskjellige metoder [27]. Ved å kombinere kvalitative og kvantitative metoder vil man få et klarere bilde av realiteten og øke validiteten. Metoden kan avgjøre om det er noen korrelasjon mellom det som er observert og det de observerte personene selv uttrykker.

5.8 Refleksjon

Som forsker er det nødvendig å reflektere over de dataene som samles inn slik at man kan si noe om validiteten til dataene.

Klein og Myers redegjør for syv prinsipper for evaluering av fortolkende feltforskning i informasjonssystemer [40]. Prinsippene er ment å gjøre det enklere å forstå hvordan mennesker tenker og handler i sosiale og organisatoriske kontekster. Dette skal bidra til å gi en helhetlig forståelse av studiet, og på denne måten gjøre det enklere å kvalitetssikre forskningsresultater innen informasjonssystemer.

1. **Den hermeneutiske sirkel:** Prinsippet danner grunnlaget for de seks andre prinsippene. Ideen er at menneskelig forståelse oppnås ved å gjenta prosessen med å forstå betydningen av de gjensidig avhengige delene og helheten av disse.
2. **Kontekstualisering:** Det kreves at man reflekterer over den historiske og sosiale bakgrunn for konteksten som undersøkes.
3. **Samhandling mellom forsker og forskningsobjekt:** Det må gjøres en kritisk refleksjon rundt hvordan de innsamlede dataene kan være påvirket av samhandlingen mellom forskere og forskningsobjekt.
4. **Abstraksjon og generalisering:** Det må være mulig å relatere ideografiske detaljer som avdekkes i datatolkningen gjennom å bruke prinsippene 1 og 2.
5. **Dialogisk resonnement:** Krever at man er oppmerksom på at de faktiske funnene ikke nødvendigvis bekrefter forskningsteoriene.

6. **Mangfoldige fortolkninger:** Krever at forsker er oppmerksom på mulige forskjeller i tolkninger blant deltakere. Mennesker oppfatter situasjoner forskjellig, noe som igjen fører til at man får ulike versjoner av samme situasjonen.
7. **Mistanke:** Forsker bør være oppmerksom på mulige feilkilder. Deltakernes historier kan være delvise eller forvrengte.

Det er viktig å forstå at disse prinsippene ikke er en absolutt mal for hvordan fortolkende studier bør utføres, men er snarere ment som en rettledning hvor forskeren velger ut hvilke av prinsippene som er relevante for sin forskning.

Del II

Forskningsspørsmål 1

Kapittel 6

Forskningsdesign for FS1

Dette kapitlet presenterer og beskriver forskningsdesignet til FS1: *“I en brukerdeltakende designworkshop, hvilken merverdi gir domenespesifikke demonstrasjonseksempler fremfor abstrakte ved formidling av mulighetene og begrensningene til Surface?”*. Forskningsdesignet består av metodene som sammen vil danne grunnlaget for å svare på forskningsspørsmålet.

6.1 Forskningsstrategi

Vi kom tidlig i kontakt med NTNU Vitenskapsmuseet om å etablere et samarbeid i forbindelse med forskningen vår. Ettersom museet allerede hadde et Surface-bord var dette en type forskning som også de så for seg å ha et stort utbytte av. De ser på dette prosjektet som noe som vil inngå som en del av utviklingen av årets pedagogiske tilbud ved Vitenskapsmuseet. Vitenskapsmuseet er et universitetsmuseum og en enhet ved NTNU [69], og egner seg svært godt for å gjennomføre en slik type forskning.

FS1 vil bli besvart gjennom et eksperiment av typen statistisk gruppesammenligning (se Seksjon 5.2). Dette gjør at vi kan måle forskjeller basert på stimuli. Måten eksperimentet vil bli gjort på er å arrangere to workshops der vi formidler sentrale konsepter ved Surface som viser dets muligheter og begrensninger ved hjelp av ulike demonstrasjonseksempler. Eksempelene skal danne grunnlag for å komme opp med ideer til hva Surface kan brukes til på Vitenskapsmuseet. Deltakerne vil ikke få noen gitte rammer for ideene, men ideene skal ta utgangspunkt i å kunne brukes på Vitenskapsmuseets Surface. Vi ønsker å gjennomføre workshopene med et lite utvalg elever og lærere, der

vi på den ene workshopen presenterer deltakerne for abstrakte formidlingsdemoer mens vi på den andre bruker konkrete, domenespesifikke formidlingsdemoer. På denne måten kan vi måle og sammenligne *idérikdommen* på de to workshopene og dermed undersøke hvilken av de to formidlingsmetodene som fungerer best. Idérikdommen omfatter både antallet unike ideer som kommer opp under hver av de to workshopene, samt kvaliteten på ideene og hvor dyptgående de er. En høyere idérikdom indikere en bedre metode å fremlegge egenskapene til Surface på.

Vår hypotese er at det vil være en forskjell i idérikdommen mellom workshopene: workshopen som får demonstrert konkrete, domenespesifikke eksempler vil ha høyere idérikdom enn den med abstrakte.

Alternative forskningsstrategier kunne også vært aktuelle for å svare på dette forskningsspørsmålet. Det kunne vært testet individuelt på deltakerne der hver deltaker enten ble utsatt for abstrakte eller domenespesifikke formidlingsdemoer. Vi ønsket imidlertid å gjøre workshopene så realistiske som mulig ved å la deltakerne samarbeide. Når det gjelder idégenereringen ville en mulighet vært å gi deltakerne mer konkrete rammer rundt hva ideene skal inneholde, som for eksempel at ideen skal være linket opp mot en konkret avdeling på Vitenskapsmuseet. Alternativt kunne vi gitt de helt frie tøyler ved å la de komme opp med ideer til hva Surface kan brukes til også utenfor Vitenskapsmuseet. Hensikten var for øvrig å ha en utforskende workshop for et gitt domene, som ville være så tett som mulig opp mot en virkelig utviklingsprosess.

6.2 Deltakere

Som deltakere på workshopene er det ønskelig for oss å bruke to små grupper bestående av seks deltakere i hver. Årsaken til dette er at et lite antall deltakere vil føre til at vi får en tett og uformell dialog. For å oppnå det vi kaller et multidisiplinært team (se Seksjon 3.2) vil vi involvere en ansatt ved Vitenskapsmuseet, en lærer, samt fire elever. Dette gjør vi fordi vi ønsker både noen som er eksperter på domenet og noen som er i målgruppen for aktuelle applikasjoner for Surface på Vitenskapsmuseet. Lærerne og elevene rekrutteres fra en videregående skole i Trondheim.

6.3 Observasjon

Under workshopene kommer vi som forskere til å fasilitere og observere. Som nevnt i Seksjon 5.4, har observatører en selektiv hukommelse og fokuserer på ulike ting. Det vil derfor bli gjort video- og lydopptak under begge workshopene, slik at vi får audiovisuelle data som nøye kan analyseres i etterkant. Et alternativ ville være å rekruttere noen eksterne fasilitatorer slik at vi kunne konsentrert oss om å observere. Vi vurderte likevel at video- og lydopptakene ville være tilstrekkelig for å analysere sesjonene nærmere.

6.4 Intervju

I kombinasjon med observasjon vil intervju kunne utfylle de manglende årsaksspørsmålene som ikke kan besvares bare gjennom observasjon, og som dermed er med på å gi et mer helhetlig bilde av situasjonen. Vi vil altså da kunne få muligheten til å høre intervjuobjektens oppfatning av hvordan prosessen gikk. Gjennom å holde et semi-strukturert gruppeintervju, formet som en diskusjon, tror vi at vi vil få en mer dynamisk interaksjon mellom deltakerne. En ulempe er for øvrig at det kan føre til at det blir vanskelig å sammenligne svarene i de to workshopene. For å unngå at dette skal føre til at vi ikke får besvart de samme spørsmålene i begge workshopene vil vi ha forhåndsdefinerte temaer som vi ønsker å få svar på (se Tillegg A.6).

6.5 Spørreundersøkelse

Vi kommer ikke til å ha kjennskap til deltakerne i forkant av workshopene og ser derfor nytten av å starte med en spørreundersøkelse. Dette vil kartlegge hvor mye innsikt og kunnskap deltakerne har om teknologien multi-touch. Dette er nemlig faktorer som kan spille inn på deres forståelse av mulighetene og begrensningene ved Surface. Spørreskjemaet kan studeres i Tillegg A.3.

For å kunne oppsummere hver enkelt deltaker sitt inntrykk av workshopene vil det også være nødvendig å gjennomføre en spørreundersøkelse etter workshopene. Dataene som samles inn gjennom disse spørreundersøkelsene benyttes til å kartlegge hva som har fungert på workshopene, hva som kunne vært gjort bedre og hva som har påvirket de mest. Dette skjemaet struktureres som påstander der deltakerne vurderer hvor enige de er i påstandene på en Likert skala (se Tillegg A.4).

6.6 Variabler

I forbindelse med eksperimenter er det en rekke variabler som er vanskelig å kontrollere. Som Oates påpeker kan en ujevn fordeling av elever ha innvirkning på resultatet. Et eksempel på dette er at alle de “flinke” elevene er i en og samme gruppe. Dette er noe vi er nødt til å ta stilling til i vår analyse av resultatene.

Vi vil gjennom de to workshopene sørge for å gi eksakt samme informasjon, slik at dette ikke skal være en variabel som har innvirkning på resultatene. Det eneste som skiller opplegget mellom de to workshopene er hvilke formidlingsdemoer de får presentert; abstrakte på den første og domenespesifikke på den andre.

Omgivelser kan være en variabel som er sterkt knyttet opp mot kreativitet. Ettersom Surface-bordet i utgangspunktet er plassert i avdelingen med utstoppede dyr på museet er vi redde for at disse omgivelsene vil låse tankene. Vi vil derfor plassere Surface i et nøytralt rom under presentasjonen av formidlingsdemoene. Workshopene vil også finne sted på et skjermet rom, der vi er isolert fra resten av de besøkende på museet. Det vil også være nødvendig at dette rommet er nøytralt og ikke representerer et gitt domene.

6.7 Oppsett av workshop

Workshopene vil ha følgende oppsett:

1. **Introduksjon:** En introduksjon til opplegget og en kort innføring til Surface.
2. **Spørreundersøkelse, pre:** Spørreundersøkelse for å kartlegge forhåndskunnskapene.
3. **Omvisning:** En kort omvisning rundt på Vitenskapsmuseet.
4. **Demonstrasjon av Surface:** Her blir formidlingsdemoene vist og utforsket på Surface.
5. **Individuell idégenerering:** Her blir deltakerne satt til å jobbe alene med å uttrykke tanker og ideer de har på post-its.

6. **Parvis diskutere seg frem til mulige ideer:** Deltakerne blir satt sammen i par og formulerer ulike ideer med bakgrunn i den individuelle seansen. For å gi de en struktur på ideene vil vi dele ut forhåndsleste kort som de skal notere på. Kortene inneholder tre felter som ba dem definere *hvem applikasjonen er ment for, hva hensikten med den er og en forklaring av hva applikasjonen går ut på*. Malen for disse kortene kan studeres i Tillegg A.5.
7. **Presentere ideene i plenum:** Presentasjon og diskusjon av ideene i plenum.
8. **Spørreundersøkelse, post:** Spørreundersøkelse for å oppsummere deltakernes inntrykk av workshopen og demoene.
9. **Semi-strukturert gruppeintervju:** Intervju for å samle inn kvalitativ data omkring deltakernes syn på workshopen.

6.8 Validitet

Som forsker er det viktig at man reflekterer over de dataene som er samlet inn, eksempler på dette er hvordan forskningen ble utført og hvordan dette kan ha øvd innflytelse på resultatet. Robson mener at analyse i alle faser er en nødvendighet, fordi rådata stort sett ikke sier noe i seg selv [61]. Vi vil derfor forsøke å bearbeide og analysere de innsamlede dataene så godt som mulig under hele prosessen, slik at vi kan basere videre forskning på disse resultatene.

Når det gjelder interpretative forskningsmetoder har Klein og Myers utarbeidet et sett prinsipper rundt disse metodene. Prinsippene har til formål å forbedre metodene ved å skape en felles standard, og er forklart i Seksjon 5.8. Vi finner en delmengde av disse relevante for vår forskning; *interaksjon mellom forsker og forskningsobjekt, dialogisk resonnement og mistanke*.

Prinsipp 3, *interaksjon mellom forsker og forskningsobjekt*, understreker spesielt at forskeren bør gjøre en kritisk refleksjon av hvordan dataene som er samlet inn kan ha vært påvirket av sosiale aspekter. I vårt tilfelle bør vi være observante på at resultatene fra forskningen kan være påvirket av hvordan skoleelevene, lærerne og de ansatte på Vitenskapsmuseet forholder seg til oss som studenter. Det er dessuten like viktig å være oppmerksom på at forholdet mellom lærer og elev kan påvirke resultatene. Elever ser på lærere som autoritære personer, noe som kan prege væremåten deres og arbeidet de gjør

under workshopene. For å skape trygge og mer naturlige rammer for både elevene og læreren ønsker vi derfor at læreren og den ansatte ved museet utgjør et par og elevene danner to par under den parvise idégenereringen. Vi vil også sørge for at guttene utgjør et par og jentene et annet, for å øke sannsynligheten for at de klarer å være seg selv.

Prinsipp 5, *dialogisk resonnement*, er et spesielt viktig prinsipp å ta stilling til i vår forskning. Blant annet kan hypotesen vår føre til at vi har et snevert syn på hva som er viktig å fokusere på, og dermed ikke være sensitive for mulige motsigelser mellom forskningsteorier og faktiske funn. For å unngå at hypotesen skygger for de faktiske funnene vil vi ha et bredt fokus og åpent sinn når vi observerer deltakerne under workshopene, slik at vi vil få med oss viktige sammenhenger. I tillegg vil video- og lydopptakene gjøre det mulig å gå gjennom seansen gjentatte ganger for å styrke funnene.

Prinsipp 7, *mistanke*, krever oppmerksomhet rundt mulige feilkilder. En aktuell feilkilde i sammenligningen mellom de to workshopene er deltakernes forkunnskaper og forutseneringer. Det vil være naivt å ikke kartlegge disse variablene i forkant av gjennomføringen. Pre-spørreundersøkelsen har til formål å avdekke forkunnskapene, og resultatene vil danne grunnlag for en refleksjon rundt validiteten.

Kapittel 7

Formidlingsdemoer

I dette kapitlet presenteres formidlingsdemoene som blir brukt i workshopene. Formidlingsdemoene er applikasjoner som er utviklet for- og vil kjøres på Surface. Formålet med demoene er å vise deltakerne på workshopen hvilke tekniske muligheter Surface har.

7.1 Konseptene

Gjennom en analyse av Surface har vi valgt å generalisere egenskapene ved plattformen gjennom fire konsepter. Disse er:

- Direkte manipulasjon (Seksjon 7.3)
- Interaksjon med fysiske objekter (Seksjon 7.4)
- Samhandling (Seksjon 7.5)
- Kombinasjon med andre teknologier (Seksjon 7.6)

Noen av disse konseptene, spesielt *Direkte manipulasjon*, vil være kjent for deltakere som har erfaring med multi-touchenheter fra før, som eksempelvis iPhone. *Interaksjon med fysiske objekter* og *Samhandling* er på den andre siden mer spesifikt for Surface og vil trolig være ukjent for deltakere uten erfaring eller kjennskap til multi-touchbord fra før.

For hvert av de fire konseptene utvikles det én abstrakt og én domenespesifikk applikasjon. I de følgende seksjonene vil konseptene forklares i nærmere detalj og de to demoene for konseptet vil beskrives.

7.2 Abstrakt kontra domenespesifikt

For å kunne utvikle abstrakte og domenespesifikke demoer er det viktig å redegjøre for hva som kjennetegner de to motpartene.

At noe er *abstrakt* innebærer at det ikke hører hjemme i verken tid eller sted [73]. Det er begreper og konsepter som eksisterer i tankene våre. Eksempler på abstrakte konsepter er tid, rom, farger og geometri. Noe domenespesifikt står i kontrast til dette ved at tingene hører hjemme i en særskilt kontekst. Eksempler på domenespesifikke artefakter ved Vitenskapsmuseet er oldtidsgjenstander, utstoppede dyr og tekniske maskiner. De abstrakte demoene vil i større grad fokusere på interaksjonsmulighetene til Surface uten å pakke det inn i et domenespesifikt eksempel.

Formidlingsdemoene vil demonstrere både look and feel og rolle for Surface (se Seksjon 3.5.5), men vil ha ulikt fokus på de to. I de abstrakte demoene har vi valgt å bruke geometriske figurer og farger til å representere ulike elementer. Fokuset i demoene vil være hva som er mulig å gjøre med disse elementene med tanke på interaksjon, fremfor hva de representerer. Demoene vil dermed befinne seg i nærheten av punkt 2 i Figur 3.4. I kontrast til dette vil de domenespesifikke demoene bruke bilder og konsepter som hører hjemme på vitenskapsmuseet. Bildene som blir brukt i to av de domenespesifikke demoene er tatt av faktiske museumsgjenstander på museet. Det har også blitt brukt referanser til dyreriket og dyrenes relasjoner, siden dette har en sentral rolle i flere av utstillingene. Ved å knytte de domenespesifikke demoene til konkrete gjenstander og temaer vil de vise hva Surface faktisk kan brukes til på museet. Dette vil tilsvare en plassering i nærheten av punkt 1 i Figur 3.4.

7.3 Direkte manipulasjon

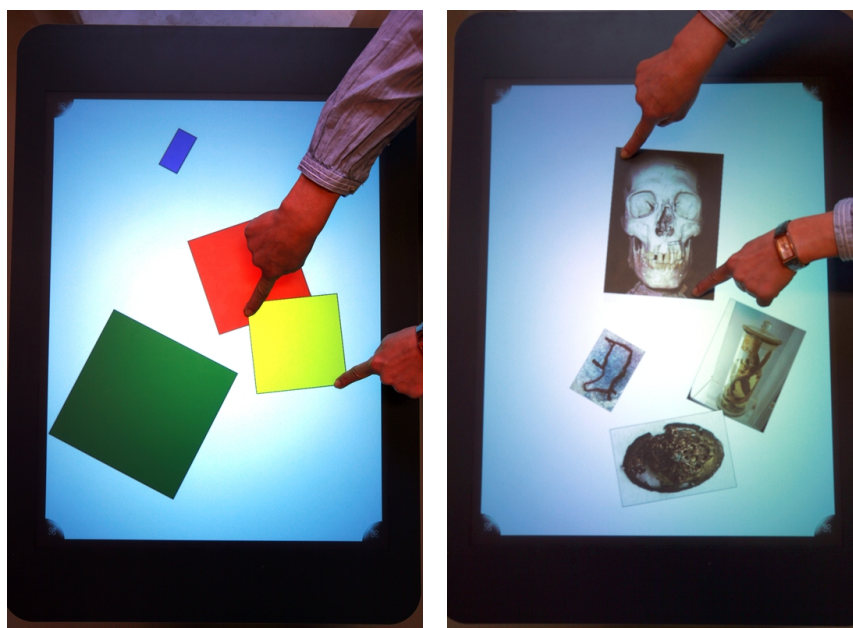
Dette konseptet innebærer noe av det mest grunnleggende innen multi-touch. Brukerne kan flytte, skalere og rotere elementer som vises på skjermen ved å bruke direkte kontakt med fingrene. Flere brukere kan manipulere forskjellige elementer på skjermen samtidig eller de kan samarbeide på samme element. Elementene kan være hva som helst; eksempelvis bilder, videoer, puslespillbrikker, menyer og tekst. Det er viktig at demoapplikasjonen klarer å formidle til workshopdeltakerne hvilke muligheter denne typen interaksjon åpner for og at det kan brukes for mange ulike formål.

7.3.1 Abstrakt demo

Den abstrakte demoen for direkte manipulasjon består av en enkel, hvit arbeidsflate med fire fargede rektangler på (se Figur 7.1a). Brukerne kan flytte, skalere og rotere rektanglene. Ved å bruke kun fargede rektangler som elementer er tanken at demoen blir tilstrekkelig abstrakt for å vise at det spesielle er *manipulasjonen* av disse elementene.

7.3.2 Domenespesifikk demo

Den domenespesifikke demoen skiller seg fra den abstrakte ved at fargerekteklene er byttet ut med bilder av museumsgjenstander (se Figur 7.1b). Bildene kan på tilsvarende måte flyttes, skaleres og roteres.



(a) Abstrakt demo

(b) Domenespesifikk demo

Figur 7.1: Formidlingsdemoer for direkte manipulasjon.

7.4 Interaksjon med fysiske objekter

Som forklart i Seksjon 4.4.1, kan Surface gjenkjenne såkalte *tags*. Når disse tagene festes på fysiske objekter vil Surface kunne interagere med objektene.

For å formidle bruken av tags valgte vi å designe demoene minimalistisk. Eksempelet med stempeling av symboler på bilder (se Seksjon 4.4.1) ble brukt som inspirasjon siden dette kommuniserer det mest grunnleggende – det fysiske objektet (stempellet) utløser en særskilt handling på skjermen (etterlater et digitalt avtrykk).

7.4.1 Abstrakt demo

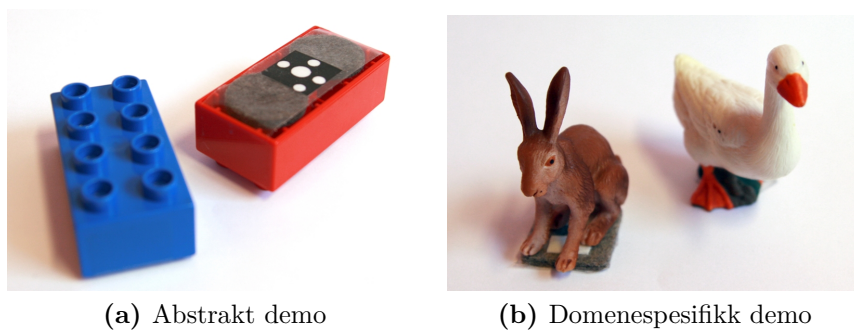
For den abstrakte demoen er det viktig å finne fysiske gjenstander som ikke innehar for mye domenekarakteristikk. Det er likevel av betydning at gjenstandene er noe de fleste kjenner til og har sett før, så de lettere forstår at det er tagen under, og ikke gjenstanden alene, som Surface gjenkjenner. Dessuten bør demoen få frem at flere ulike objekter kan skape forskjellig oppførsel på skjermen når de plasseres på den. Basert på disse kriteriene valgte vi å bruke to avlange Duploklosser¹ (se Figur 7.2a). Klossene hadde forskjellige farge; en rød og en blå. Fargeulikheten illustrerer at klossene er to forskjellige objekter og at de tilsvarende vil skape ulik feedback når de berører skjermen. Klossenes avlange form bidrar til å gi de en retning.

Når en person legger klossen på skjermen reagerer demoen med å vise en geometrisk figur rett under klossen (se Figur 7.3a). Den blå klossen frembringer en blå firkant, mens den røde viser en rød trekant. Trekanten er likebeint med en spiss som hele tiden peker i samme retning som klossens ene kortside. Hvis klossene blir flyttet på bordet uten at de løftes vil figuren følge etter. Dersom klossen løftes fra skjermen vil figuren bli stående og ikke forsvinne før applikasjonen avsluttes. Neste gang klossen berører skjermen vil det da dukke opp en ny, lik figur.

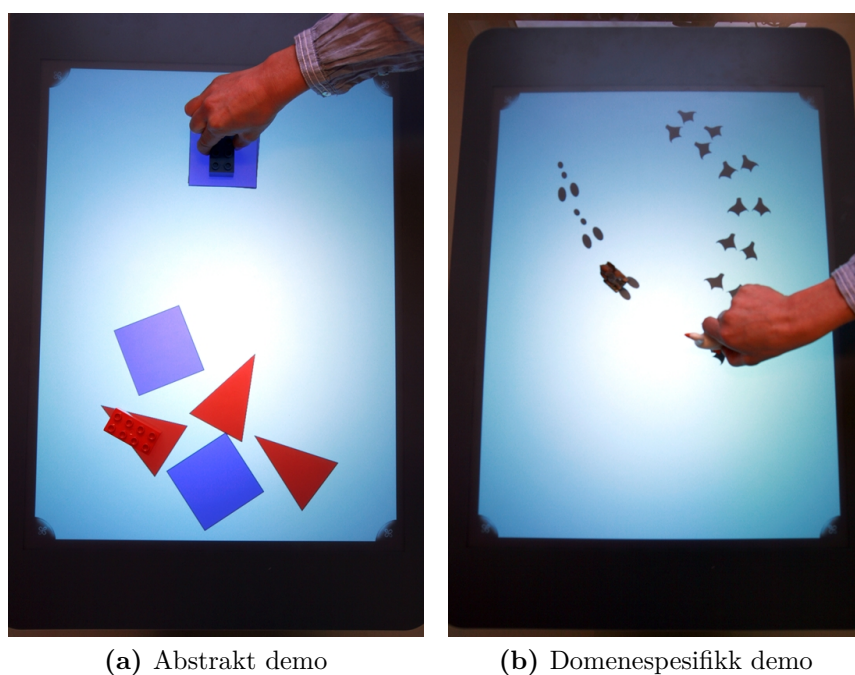
7.4.2 Domenespesifikk demo

I den domenespesifikke demoen er klossene byttet ut med miniatyrmodeller av dyr – en and og en hare (se Figur 7.2b). Når et dyr kommer i kontakt med skjermen blir dyrets sporavtrykk lagt igjen i svart farge på skjermen (se Figur 7.3b). På tilsvarende måte som med Duplo-klossene, peker dyresporet i samme retning som dyret.

¹Lekeklosser av plast, <http://duplo.lego.com>



Figur 7.2: Fysiske objekter med tags festet på undersiden.



Figur 7.3: Formidlingsdemoer for interaksjon med fysiske objekter.

7.5 Samhandling

Som nevnt i Seksjon 4.4 er en av fordelene med den store skjermstørrelsen på multi-touchbord at det er plass til flere brukere rundt bordet samtidig. Fire brukere kan eksempelvis plassere seg slik at de står ved hver sin kant av bordet. Bordet innbyr på denne måten til samhandling. I Microsoft Surface User Experience Guidelines [14] blir det oppfordret til å utvikle GUI-er på en slik måte at det ikke favoriserer en av sidene. Det bør ikke være noe opp eller ned. Elementene i applikasjonen bør flyte slik fysiske objekter kan gjøre på et fysisk bord. Ved å utvikle applikasjoner som er sosiale vil styrken til multi-touchbord komme klarere frem. Samhandlingsprinsippet er nemlig et av hovedområdene som skiller slike bord fra mindre touch-flater som smarttelefoner og touch-brett. For å formidle til workshopdeltakerne hvordan bordet kan brukes som en sosial aktivitet valgte vi å designe disse demoene på en slik måte at brukerne intuitivt ville forstå at man kan samarbeide. Dersom applikasjonen krever at flere enn to elementer må beveges jevnlig på skjermen for at den skal gi mening er tanken at personer vil samarbeide siden de naturligvis bare har to hender. Hensikten med demoene er å se hvorvidt deltakerne vil oppfatte Surface som et medium egnet for samhandling og om de vil kunne tenke ut andre scenarier for sosial bruk av Surface på Vitenskapsmuseet.

7.5.1 Abstrakt demo

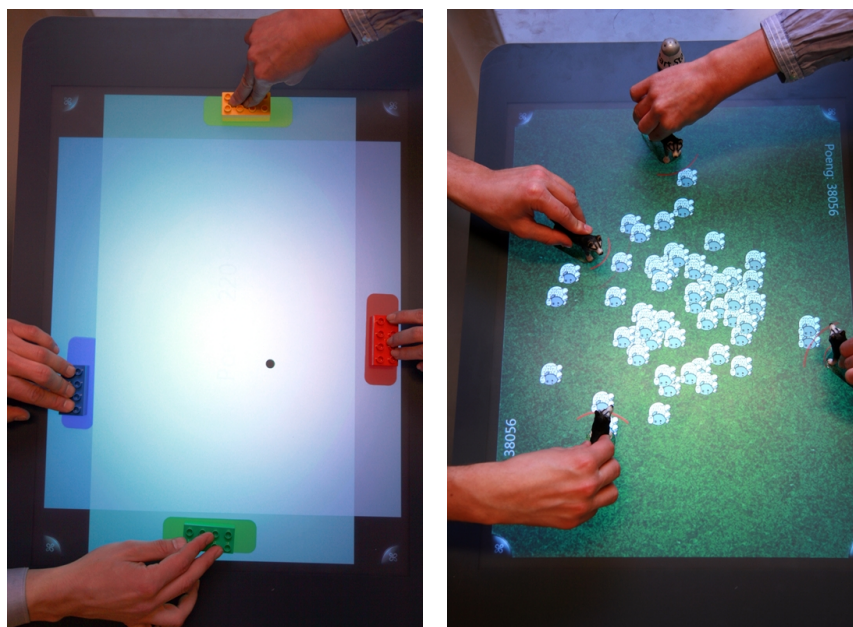
En av applikasjonene som følger med Surface er *“Paddleball”*. Paddleball er en variant av det klassiske spillet *“Pong”* [76] og går ut på at fire spillere prøver å holde en ball unna sin side av bordet. For å hindre at ballen skal treffe kanten flytter spillerne på hvert sitt rektangel som *“slår”* ballen vekk fra sitt område. Denne applikasjonen er et godt eksempel på en abstrakt demo for samhandling på Surface. Vi ønsket for øvrig at demoen skulle fokusere på samarbeid fremfor konkurranse. Det ble derfor utviklet et nytt spill basert på konseptet til Paddleball. Dette ga oss dessuten mulighet til å inkludere tags i interaksjonen, noe som viser deltakerne andre måter tags kan benyttes på.

Spillet består av fire områder på hver kant av bordet i fargene rød, blå, gul og grønn (se Figur 7.4a). Fire spillere eier sitt område og har en Duplokloss (se Figur 7.5a) i samme farge som området sitt. Når spillet starter vil en eller flere baller (svarte sirkler) bevege seg ut fra sentrum av bordet i en vilkårlig retning mot en av de fire områdene. Spilleren som kontrollerer området må

da bruke Duploklossen til å blokkere for ballen så den ikke forsvinner ut. Klossen kan ikke bevege seg utenfor sin sone, den vil ikke da ha noen effekt på ballen. I stedet for å konkurrere mot hverandre fungerer spillerne som et lag. Hver gang en spiller treffer en ball med klossen sin vil laget få poeng. Hvis en ball når ytterkanten av spillet har laget tapt.

7.5.2 Domenespesifikk demo

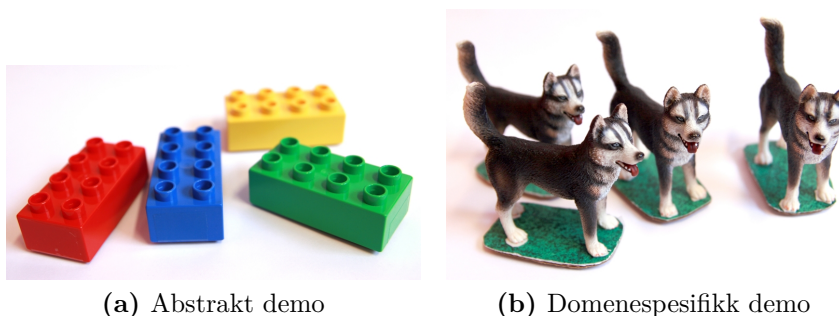
Den domenespesifikke demoen har likhetstrekk til den abstrakte. Den er et spill som går ut på at deltakerne må samarbeide for å holde et stort antall sauer innenfor skjermens grenser. Sauene forsøker å rømme ut fra sentrum av bordet og det er spillernes oppgave å holde de innenfor. Til rådighet har spillerne fysiske miniatyrhunder (se Figur 7.5b) som, når de kommer i kontakt med skjermen, skremmer sauene vekk fra seg (se Figur 7.4b). Spillerne fungerer som et lag og poengene øker jo lenger laget klarer å forhindre at en av sauene rømmer. Domenet i denne demoen er dyreriket og dyrenes relasjoner til hverandre, noe som er budskapet i flere av utstillingene på museet.



(a) Abstrakt demo

(b) Domenespesifikk demo

Figur 7.4: Formidlingsdemoer for samhandling.



Figur 7.5: Fysiske objekter med tags festet på undersiden.

7.6 Kombinasjon med andre teknologier

Det siste konseptet som skal formidles på workshopene er at Surface kan kombineres med andre teknologier (se Seksjon 4.4.1). Både den abstrakte og domenespesifikke demoen viser hvordan bluetooth kan brukes til å overføre elementer fra Surface til en mobiltelefon. Begge demoene er basert på en applikasjon utviklet av Microsoft som ble publisert i form av åpen kildekode [41]. En svakhet ved de to demoene er for øvrig at de kun vil fungere på bluetooth-enheter som støtter OBEX-protokollen² for filoverføring over bluetooth. Det betyr at blant annet Apple iPhone ekskluderes. Under demonstrasjonene vil filoverføringen demonstreres på mobiltelefonene *Samsung SGH-E730* og *Sony Ericsson C702* i tillegg til eventuelle mobiltelefoner deltakerne måtte ha med seg.

7.6.1 Abstrakt demo

Når demoen starter opp vises fire kvadrater i forskjellige farger som kan beveges fritt på skjermen, samt et ikon som forteller at Surface søker etter bluetooth-enheter i nærheten (se Figur 7.6a). For hver enhet som oppdages vil et ikon som representerer enheten legges til i en liste på høyre side på skjermen. Ved å føre et av de fargede kvadratene over et enhets-ikon vil Surface forsøke å overføre kvadratet i form av et bilde til enheten. Etter overføringen kan brukeren se bildet på sin enhet (se Figur 7.7a).

²<http://www.bluetooth.com/English/Technology/Works/pages/obex.aspx>

7.6.2 Domenespesifikk demo

For den domenespesifikke demoen er de fargede kvadratene byttet ut med bilder av museumsgjenstander, hvorav to av bildene er de samme som i demoen for direkte manipulasjon (se Figur 7.6b). Brukerne kan på denne måten føre over bilder av museumsgjenstander til mobiltelefonen sin (se Figur 7.7b).



(a) Abstrakt demo



(b) Domenespesifikk demo

Figur 7.6: Formidlingsdemoer for Surface i kombinasjon med andre teknologier.



(a) Abstrakt demo



(b) Domenespesifikk demo

Figur 7.7: Mobiltelefoner som har overført et bilde fra Surface.

Kapittel 8

Gjennomføring av Workshops

I dette kapitlet vises det hvordan de to workshopene ble gjennomført og hvilke resultater de førte til. Kapitlet er delt opp i to seksjoner, én for hver workshop. Første seksjon omhandler workshopen der abstrakte formidlingsdemoer ble benyttet, mens den andre skildrer workshopen med domenespesifikke formidlingsdemoer på tilsvarende måte. Seksjonene starter med å introdusere deltakerne og bakgrunnen deres. Videre fremstilles forkortede versjoner av ideene de kom opp med under idédugnaden før seksjonene avsluttes med kommentarer fra deltakerne.

Til hver av workshopene rekrutterte vi fire elever og én lærer fra Trondheim Katedralskole. Elevene gikk i første klasse, var 16 år og hadde blant annet naturfag i lærerplanen sin. Læreren som deltok var naturfaglærer. Årsaken til at vi etterlyste elever og lærere med innsikt i naturfag var at fagets innhold korrelerer i stor grad med Vitenskapsmuseets budskap. Deltakerne har på den måten domenekunnskap om museets innhold. I tillegg til representantene fra skolen deltok én ansatt fra Vitenskapsmuseet.

8.1 Gjennomføring

Til workshopene disponerte vi et grupperom på Vitenskapsmuseet som ble brukt til oppholdsrom under introduksjonen og idédugnaden. Workshopene startet med en introduksjon av prosjektet, en spørreundersøkelse for å kartlegge forkunnskapene og en kort presentasjon av Surface (se Figur 8.1). Deltakerne fikk deretter en omvisning på museet som varte i 20 minutter.

Etter omvisningen samlet deltakerne seg rundt Surface-bordet i et nøytralt



Figur 8.1: Introduksjon.

rom som var skjermet fra museets utstillingsområder. Her fikk deltakerne se formidlingsdemoene i aksjon (se Figur 8.2). Først viste vi frem en demo og forklarte hvilket konsept den skulle formidle, før deltakerne fikk utforske demoen hands-on.



Figur 8.2: Demonstrasjon av formidlingsdemoene.

Etter demonstrasjonen av Surface samlet alle deltakerne seg på kontoret igjen for en idédugnad. Idédugnaden bestod av 10 minutter individuell jobbing før deltakerne jobbet parvis i 30 minutter (se Figur 8.3). De ble bedt om å komme opp med ideer til hva Surface kan brukes til på Vitenskapsmuseet.

For å gi deltakerne en struktur for ideene fikk de utdelt forhåndslagde kort som de skulle notere på (se Seksjon 6.7). Figur 8.4 viser et av kortene under workshopene.



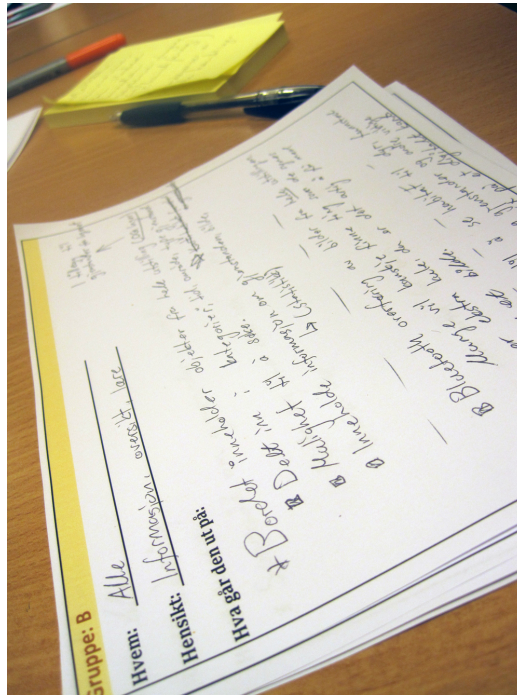
Figur 8.3: Idédugnad i par.

Etter idédugnaden gikk vi gjennom alle ideene i felleskap (se Figur 8.5). Dette ble gjort for at alle deltakerne skulle få forklare ideene sine og hva som hadde inspirert dem.

Avslutningsvis gjennomførte deltakerne en spørreundersøkelse som inneholdt spørsmål omkring hvordan de hadde opplevd workshopen, med hovedvekt på formidlingsdemoene. En ustrukturert diskusjon for å få mer kvalitative svar avrundet til slutt workshopene.

8.2 Workshop 1 - abstrakt formidling

Deltakerne som var med på den første workshopen fikk prøve abstrakte formidlingsdemoer på Surface.



Figur 8.4: Kort for gruppe B.



Figur 8.5: Presentasjon av ideene.

8.2.1 Deltakerne

De fire elevene som deltok bestod av to jenter og to gutter som alle gikk i samme klasse. Læreren var naturfaglæreren deres. Den ansatte fra Vitenskapsmuseet jobbet som avdelingsingeniør ved seksjon for formidling og hadde god kjennskap til museets budskap, innhold og daglige virke.

Spørreundersøkelsen som ble gjennomført innledningsvis i workshopen avslørte at deltakerne hadde lite til ingen kjennskap til Surface fra tidligere. Kun én av seks hadde hørt om teknologien og prøvd et eksemplar. Likevel eide halvparten av deltakerne en enhet med multi-touch og av de resterende hadde to deltakere prøvd et eksemplar. Deltakerne var med andre ord kjent med multi-touchteknologi, men ukjent overfor Surface. Ut ifra dette kan man anta at de kjente til grunnleggende prinsipper for gestures på multi-touchflater.

Når det gjelder Vitenskapsmuseet hadde samtlige deltakere besøkt museet tidligere, hvorav tre hadde besøkt det i løpet av det siste året. Elevene hadde stort sett besøkt museet i forbindelse med undervisning på barne- og ungdomsskolen.

Av hensyn til deltakernes anonymitet har de blitt gitt pseudonymene Per, Kari, Ola, Anne, lærer Hermansen og ingeniør Berg.

8.2.2 Ideer

Det ble utarbeidet totalt 18 ideer under denne workshopen. Dette utgjorde 14 unike konsepter, etter å ha slått sammen ideer som lignet veldig på hverandre. Dette er konseptene som kom frem (noe omskrevet for bedre lesbarhet):

Objekter med informasjon

Et sett med figurer/bilder med tags på baksiden. Besøkende kan finne objektene på steder der det er lite tekst og kan ta de med til bordet, legge objektet på skjermen og få mer informasjon.

Endring av historie

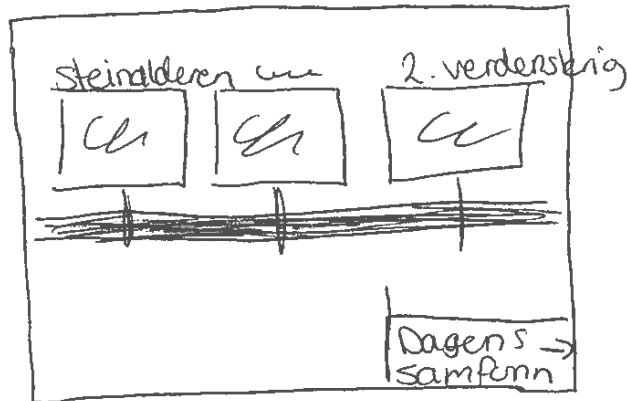
I en tidslinje skal man kunne velge å ta bort en historisk hendelse, som for eksempel 2. verdenskrig, og se hvordan noen mener verden ville sett ut i dag.

Dyrs oppførsel

Plassere dyrearter i to forskjellige grupper (eller flere). Når man drar et dyr til en gruppe vil man kunne se væremåten til dyret endres i forhold til hvilken gruppe den er i. Gruppene kan for eksempel være "i frihet" og "i fangenskap/dyrehage".

Historiske hendelser

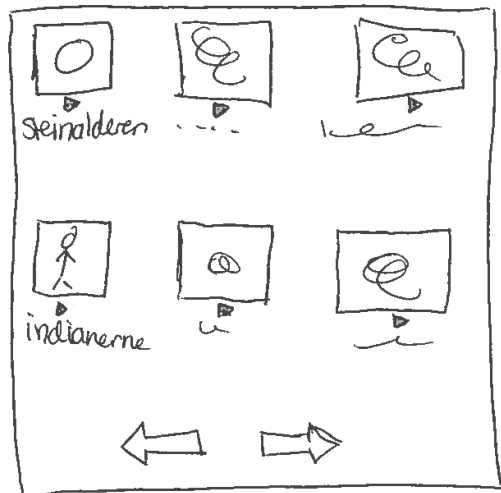
Surface kan brukes til å bla tilbake i tid (tidslinje) hvor man kan trykke på forskjellige historiske hendelser og gjøre en aktivitet som er relevant for den tidsepoken. Figur 8.6 viser illustrasjonen gruppen laget på idékortet.



Figur 8.6: Illustrasjon av ideen *Historiske hendelser*.

Videoavspilling

Barn kan spille av korte videosnutter om en hendelse, et dyr etc. som de vil lære mer om. Passer bra for de som ikke kan lese og for de som lærer best gjennom det visuelle. Figur 8.7 viser illustrasjonen gruppen laget på idékortet.



Figur 8.7: Illustrasjon av ideen *Videoavspilling*.

Relasjoner i universet

Man får et stort bilde av verdensrommet og får endre og plassere planeter på en annen plass enn der de allerede står. Man kan så zoome inn på en planet og se om plasseringen i verdensrommet har påvirket den i form av f.eks. klima eller om det faktisk er mulig at planeten blir plassert på en annen plass.

Kategorisering av attraksjoner

Vitenskapelig spill for barn og unge. Dra og plassere forskjellige dyr, figurer, personer etc. til en rute merket med tidsepoke, geografisk sted, sjanger e.l. Figur 8.8 viser illustrasjonen gruppen laget på idékortet.



Figur 8.8: Illustrasjon av ideen *Kategorisering av attraksjoner*.

Letelek

En skoleklasse blir delt opp i grupper og får et oppdrag. Oppdraget går ut på at de skal finne noe på museet, ta bildet av det og sende det til Surface via bluetooth. Den raskeste gruppa får poeng. Poengscoren holdes på Surface. Oppdraget kan også være motsatt – at man sender et bilde fra Surface til mobiltelefonen og man skal da finne gjenstanden på bildet.

Musikk

Ved å bruke et virtuelt piano på Surface kan man produsere musikk, akkorder o.l.

Spill

Bruk av Surface til å spille multi-touchbaserte spill som barn og unge kjenner til fra andre medier som iPhone og Android-baserte telefoner.

Bygge by

Bygge sin egen by ved hjelp av tags. Hus, veier, elver osv. kan være fysiske objekter med en tag under. Disse kan stemples inn i byen. Objektene skal kunne finnes fritt på museet.

Bildemanipulering

Laste inn bilde av seg selv som kan settes inn i en sammenheng, for eksempel et bilde fra steinaldersamfunn. Dette bildet skal det være mulig å sende på e-post/bluetooth.

Puslespill

Puslespill med utdrag fra ulike deler av utstillingen.

Økosystem

Oversiktsbilde av for eksempel et økosystem hvor man kan blåse opp og lese om de ulike artene.

8.2.3 Kommentarer fra deltakerne

Under diskusjonen i etterkant av idédugnaden fikk deltakerne mulighet til å uttrykke hva de mente om opplegget under workshopen. På spørsmål om det var omvisningen på museet eller demonstrasjonen av Surface som inspirerte de mest til ideene svarte de *“omvisningen”*. Per mente at dette blant annet var fordi demoene stort sett viste ting de kunne om multi-touch fra før, med unntak av tags. Omvisningen på sin side viste mer av hva som rørte seg på museet og meningen med utstillingene.

Når vi spurte om de kunne se for seg at det hadde vært lettere å komme frem til ideer hvis de hadde fått mer konkrete eksempler på hva Surface kunne brukes til enn abstrakte demoer, svarte de *“ja”*. Ingeniør Berg mente at det nok ville være mer inspirerende med eksempler på hva Surface faktisk kunne vært brukt til på museet. Per var enig og supplerte med at demoene var *“litt simple”*.

8.3 Workshop 2 - domenespesifikk formidling

Deltakerne som var med på den andre workshopen fikk prøve de domenespesifikke demoene.

8.3.1 Deltakerne

Elevene som deltok på den andre workshopen hadde også en jevn fordeling mellom gutter og jenter. De fire gikk i samme klasse og ble undervist i naturfag av læreren som deltok på denne workshopen. Representanten fra Vitenskapsmuseet var stipendiat ved seksjon for formidling og hadde god kjennskap til museet.

Til forskjell fra Workshop 1 (heretter forkortet W1) hadde flere av deltakerne på denne workshopen erfaring med Surface fra tidligere. Fire av seks deltakere hadde hørt om Surface og av disse fire hadde tre prøvd et eksemplar. Erfaringen med øvrige multi-touchenheter var imidlertid svært lik den av deltakerne på W1. Halvparten av deltakerne eide en enhet med multi-touch og den andre halvparten hadde prøvd et eksemplar. Dette betyr at alle deltakerne var kjent med multi-touchteknologien fra før av.

På lik linje med W1 hadde samtlige av deltakerne besøkt Vitenskapsmuseet tidligere. Fire av de seks deltakerne hadde dessuten besøkt museet i løpet av det siste året. Også her hadde elevene besøkt museet i forbindelse med skolen.

Av hensyn til deltakernes anonymitet har de blitt gitt pseudonymene Nils, Line, Erik, Trine, lærerinne Fredriksen og arkeolog Pedersen.

8.3.2 Ideer

Under den andre workshopen kom deltakerne frem til totalt 18 ideer. Dette var like mange som W1 kom frem til, men etter sammenslåing av lignende ideer utgjorde det 16 unike konsepter, to mer enn W1. Konseptene var (noe omskrevet for bedre lesbarhet):

Dyrenes relasjoner

Spill som lar barn "ta del" i samspillet i naturen, med for eksempel rovdyr-bytte-situasjoner. Ved å bruke artene som spiser hverandre i naturen, kan barn få et bilde av næringskjeder. Her kan en bruke tags som i sauespillet og feste en tag under hvert dyr.

Sammenheng i naturen

Spill der en kan krysse forskjellige gener med hverandre og se hvordan utviklingen blir ut i fra hva en gjør. Se hvilke spor ulike dyr lager ved å feste tags under. Se sammenhenger i økosystemer når man tar bort eller innfører dyr i etablerte samspill, eller forandrer på abiotiske faktorer osv. En kan også se hva som skjer i kroppen når en f.eks. spiser forskjellige ting.

Virtuelle reiser i naturen

Lage tredimensjonale rom en kan bevege seg innover i ved hjelp av multi-touch. En kan eksempelvis reise gjennom verdensrommet, gå i skogen og se på samspillet her eller svømme som fisk i havet. En kan se på hendelser, trykke videre på spesielle arter og se mer, lese om planetene og se avstander mellom dem e.l.

Objekter med informasjon

Besøkende tar med objekter med tags festet under fra tilhørende utstillinger til bordet og får opp utdypende info, bilder, filmer, spill osv. De kan selv velge hva de vil se mer på, slik at dette passer for alle aldre. Når en legger objektet på bordet kan en f.eks. få opp en meny med de tilgjengelige elementer som er relevant for det objektet representerer i utstillingen. Det at bordet har multi-touch gjør at flere kan se på innhold fra forskjellige deler i utstillinger samtidig (eksempelvis kan noen se en dokumentar mens andre spiller). Objektene kan være papirlapper som kan kastes etter bruk eller spesifikke gjenstander som kan sorteres tilbake til rett plass i utstillingen.

Fargelegging

Et tegnebrett der barn kan fargelegge eksempelvis en tiger. Barna kan velge hva de vil fargelegge ved å ha en hovedmeny med bilde av forskjellige elementer i utstillingen. Tegnebrettet bør kunne forminskes og forstørres slik at man kan gjøre andre ting samtidig. Figur 8.9 viser illustrasjonen gruppen laget på idékortet.



Figur 8.9: Illustrasjon til ideen *Fargelegging*.

Sammenhengen mellom dyr og habitat

Hva skjer om du plasserer et dyr i et habitat? Hva skjer om du plasserer flere dyr i et habitat? Hva skjer om du plasserer dyr i galt habitat? Hvordan oppfører dyr seg mot andre dyr? Barn vil vite hva som skjer om du gjør noe galt, og ikke bare få en tilbakemelding.

Dyrenes anatomi

Besøkende kan søke opp dyr og se animasjoner av dyrets anatomi. F.eks. hvordan fordøyelsessystemet til en løve fungerer eller hva som skjer med muskler og ben i kroppen når den løper.

Informasjon om museumsgjenstander

Bordet inneholder objekter fra hele utstillingen (og på lageret). Disse er delt inn i kategorier som tid, område, type gjenstand o.l. Det skal være mulig å søke og få frem informasjon om gjenstanden og bilde. Denne informasjonen skal det kunne gå an å overføre til mobiltelefonen via bluetooth. Det skal være mulighet til å se habitat til dyr, funnsted til gjenstander og andre viktige geografiske punkt på et digitalt kart.

Energi fra kraftverket til brødristeren

En besøkende skal kunne velge en type energikilde (eksempelvis vindkraft). Deretter skal han kunne sette sammen et slikt anlegg (dra ledningene fra anlegget til brødristeren i en husholdning). En kobler sammen og setter opp gjenstander på skjermen.

Dyrespor

Tags til forskjellige dyr i arken på museet. Når et objekt som representerer et dyr legges på bordet vises fotsporet. Dette må ikke nødvendigvis gjøres med objekter med tags. En samling med masse dyrespor kan være en del av et informasjonssystem om dyrene.

Quiz

Spørsmål om gjenstander i utstillingen. Besøkende kan på denne måten teste kunnskap.

Menneskekroppen

Starter med et "helt menneske". En kan trykke på ulike steder for å gå inn i kroppen. På denne måten kan en lære om organer og deres funksjon (hjertet, lunger, øyet, øret, fordøyelse osv.)

Systematikk

Sortere levende organismer inn i riktig kategori. Kan evt. sortere i rekke/divisjon som underoppgaver.

Bilder med informasjon

En kan ha fotografier på bordet på tilsvarende måte som i første demo. Fotografiene må ha informasjon, slik at man kan lære hva det er, alder, geografiske data o.l.

Brikke for å hente opp informasjonen hjemme

Man samler både bilder og info som man vil ha mer informasjon om i en egen "brikke" med en tag under som man kan ta med seg. Når man senere logger på museets nettside, kan man skrive inn nummeret på brikken og laste ned materialet til egen PC.

Arkeologisk utgraving

En kan feste en tag under en "graveskje" som kan skrelle av lag for lag med jord. Når en graver vil det dukke opp arkeologiske funn. En kan da lære noe om gjenstanden og de menneskene som brukte den.

8.3.3 Kommentarer fra deltakerne

I likhet med deltakerne på W1 mente også denne gruppen at det ikke var vanskelig å forstå formidlingsdemoene. Nils begrunnet dette med at *"de har vært borti multi-touchskjermer før på mobiltelefonen og lignende"*. Konseptet med å bruke tags var for øvrig nytt for dem.

Deltakerne syntes at både omvisningen og demoene ga god inspirasjon til idédugnaden. De innrømmet også at de trolig hadde blitt påvirket en del av innholdet i demoene. Eksempelvis hadde samtlige deltakere tenkt en del i retning av dyreriket under idédugnaden, noe som kunne skyldes at de hadde demoen med dyrespor og sauespillet i tankene samtidig som mye av innholdet på museet er dyr.

På spørsmål om hva de tror utfallet av idédugnaden hadde vært dersom de fikk presentert abstrakte formidlingsdemoer med farger og geometriske figurer fremfor domenespesifikke svarte de at det trolig hadde vært vanskeligere å komme opp med bruk av tags.

Kapittel 9

Resultater fra Workshops

I dette kapitlet blir resultatene fra workshopene analysert. Seksjon 9.1 presenterer funnene som ble gjort under workshopene. Funnene er basert på ideene som ble generert og spørreundersøkelsen som ble gjennomført på slutten av hver workshop. I Seksjon 9.2 analyseres funnene. Analysen er støttet opp av diskusjonen vi hadde avslutningsvis under hver workshop og øvrige utsagn underveis.

9.1 Funn

Som nevnt i Seksjon 6.1 måler vi resultatene fra workshopene i iderikdom. Det er interessant å merke seg at antall ideer som totalt ble generert var det samme for begge workshopene. Når vi likevel betrakter ideene nærmere og ser at noen av disse ligner svært mye på hverandre kommer det frem at det ble generert flere unike konsepter på Workshop 2 (heretter forkortet W2) enn W1. W2 resulterte i 16 unike konsepter mot W1 som produserte 14.

Når det gjelder bredden av innholdet i ideene viser resultatene imidlertid en større forskjell. Ideene fra W2 bærer et større preg av å utnytte Surface sitt fulle potensial. Blant annet inneholder seks av ideene som ble generert under W2 bruk av tags, mens det tilsvarende bare er to ideer fra W1 som utnytter denne muligheten. Også samhandlingsprinsippet er mer til stede i ideene fra W2. To av ideene her foreslår å kunne skalere ned en applikasjon så andre besøkende kan bruke deler av bordet samtidig. Likevel kan man se av ideene at W1 har utnyttet bruk av bluetooth-teknologi, noe som er fraværende på W2.

Deltakerne under begge workshopene var flinke til å bruke varierte temaer i ideene. Vi ser imidlertid en tendens at W2 bruker dyreriket og biologi som tema i applikasjonene i større grad enn W1. 6 av ideene fra W2 dreier seg direkte om dyr, noe vi kun ser tre tilfeller av i W1. Dessuten blir dyr brukt som eksempler i flere av ideene, blant annet fra ideen *Fargelegging*: “Et tegnebrett der barn kan fargelegge eksempelvis en tiger...”.

Spørreundersøkelsen som ble gjennomført avslutningsvis avslører tegn til at deltakerne oppfattet W2 som mer inspirerende enn W1. Selv om svarene avviker marginalt mellom de to workshopene, var deltakerne på W2 mer enig enn motparten i påstanden om at det var lett å komme med ideer til hva Surface kunne brukes til på Vitenskapsmuseet. De syntes dessuten i større grad at formidlingsdemoene ga god inspirasjon til idédugnaden, spesielt når det gjaldt demoen for samhandling (se Seksjon 7.5).

9.2 Analyse

Dog marginalt, kan vi se en positiv økning i kvantiteten på unike ideer når deltakerne blir presentert for domenespesifikke formidlingsdemoer fremfor abstrakte. Likevel tilsier den knappe forskjellen, kombinert med det faktum at begge workshopene faktisk resulterte i nøyaktig samme antall ideer totalt, at det ikke er den kvantitative idérikdommen som skiller mellom de to. Det er med andre ord grunn til å tro at deltakere som blir presentert for domenespesifikke formidlingsdemoer ikke vil produsere *flere* ideer enn hvis de fikk prøve abstrakte demoer. Forskjellen ligger snarere i *typen* ideer som blir generert.

9.2.1 Fordeler med domenespesifikke demoer

Funnene avdekket at det var en merkbar forskjell i bredden av innhold i ideene mellom de to workshopene. Spesielt hadde deltakerne på W2 tenkt mer på bruken av tags. Grunnen til dette kan være at de konkrete, domenespesifikke eksemplene som ble brukt til å demonstrere interaksjon med fysiske objekter og samhandling viste bruk av tags som er relevant for innholdet på Vitenskapsmuseet. Denne relevansen kan ha påvirket deltakerne på W2 til i større grad å tenke på bruk av tags under idédugnaden. Dette stemmer godt overens med resultatene fra spørreundersøkelsen der deltakerne på W2 er mer enig i at demoene var inspirerende og da spesielt demoen for interaksjon med fysiske objekter og demoen for samhandling. Per som deltok på

W1 kommenterte som nevnt i Seksjon 8.2.3 at formidlingsdemoene “viste ting de stort sett kunne fra før”. Dette beviser at de abstrakte demoene ikke klarte å illustrere hva som skiller Surface fra andre mer kommersielle multi-touchenheter som eksempelvis smarttelefoner. Dette kan bero på at ensfargede, geometriske figurer ikke fremhever egenskapene ved Surface som differensierer den fra andre multi-touchenheter; den store skjermen og tolking av tags. Noe vi observerte under demonstrasjonen var blant annet at deltakerne som forstørret et blått rektangel ikke opplevde samme nytteverdien som deltakerne som fikk forstørre et bilde av en hodeskalle og kunne studere detaljene i bildet.

Spørreundersøkelsen og kommentarene kunne fortelle at begge gruppene av deltakere forstod sine formidlingsdemoer godt og at de følte at de fikk et tydelig bilde av mulighetene ved Surface. Likevel kunne deltakerne på både W1 og W2 se for seg at det er bedre med konkrete eksempler på hva Surface kan brukes til på museet enn abstrakte eksempler som ikke har noe med settingen å gjøre.

9.2.2 Fordeler med abstrakte demoer

Selv om domenespesifikke formidlingsdemoer virker å ha en positiv innvirkning på utnyttelsen av Surface sitt potensial, viser funnene at deltakerne som prøvde disse kan ha blitt påvirket av eksemplene. Ideene utviklet under W2 bærer preg av temaer fra dyr og anatomi i så stor grad i forhold til W1 at det er grunn til å tro at deltakerne under W2 har latt seg prege av temaene som ble brukt i formidlingsdemoene. Både demoen for interaksjon med fysiske objekter (dyrespor) og samhandling (sauespill) hadde temaer fra dyreriket. Flere av ideene fra W2 var direkte videreføringer av dyresporideen, mens andre ideer bestod av dyr som påvirket hverandre. Når deltakerne på W2 fikk spørsmål om hvorfor mange av ideene dreide seg om dyr svarte lærerinnen Fredriksen blant annet: *“Det har kanskje en sammenheng med den der”*, mens vedkommende pekte på bildet av demoen med dyrespor. Videre var et av bildene i demoen for direkte manipulasjon og demoen for kombinasjon med andre teknologer en hodeskalle. Denne hodeskallen fant deltakerne svært fascinerende under demonstrasjonen. Det er nærliggende å tro at dette kan ha bidratt til ideen “Dyrenes anatomi” og “Menneskekroppen”.

Det er videre grunn til å tro at deltakerne under W2 kan ha blitt “blendet” av de konkrete eksemplene i formidlingsdemoene. Med dette menes det at deltakerne under W1 hadde mer ro til å kunne reflektere over hva det var som faktisk skjedde under demonstrasjonen i motsetning til W2 som hadde

en tendens til å la seg fascinere av motivet på bilder og utformingen av fysiske objekter. Et sitat under demonstrasjonen på W1 illustrerer hvordan deltakerne på denne workshopen reflekterte over hva demoen viste. Kari spurte under demoen for samhandling: *“Er det selve klossen som stopper ballen eller er det markeringa rundt?”*. Dette indikerer at vedkommende reflekterer over hvorvidt bordet “ser” klossen eller om den kun oppfatter den geometriske figuren som dannes på skjermen under den når tagen kommer i kontakt med bordet. Til sammenligning var det ingen som satte spørsmålsteget under W2 om hvorvidt det var hundemodellen eller bjeffeanimasjonen som påvirket sauene. Det er ikke dermed sagt at deltakerne under W2 trodde at Surface kunne oppfatte selve hundemodellene, men dersom det var tilfellet at deltakerne ikke forstod hva som faktisk skjedde ville dette kunne skape en mental modell (se Seksjon 2.3.3) hos deltakerne som ikke stemte overens med det som faktisk skjer. Det kunne i denne sammenhengen resultere i ideer som ikke baserte seg på det Surface er i stand til.

9.3 Konklusjon

Analysen viser at begge måtene å formidle mulighetene og begrensningene ved Surface på bringer med seg fordeler og ulemper. Mens domenespesifikke formidlingsdemoer er gunstig for å vise nytteverdi ved de forskjellige egenskapene ved bordet, forhindrer de abstrakte demoene workshopdeltakere mot å bli fastlåst i eksemplene. En måte å hente det beste fra begge konseptene vil være å kombinere demonstrasjon ved bruk av abstrakte og domenespesifikke eksempler.

9.3.1 Sammenligning med annen forskning

Klingsheim og Raae [56] påviste fordelene ved bruk av hands-on demonstrasjoner av ny teknologi for ikke-teknologiske brukere, og mente samtidig at domenespesifikke formidlingseksempler vil være bedre enn abstrakte (se Seksjon 1.2.1). Vi har gjennom våre resultater bekreftet fordelene med hands-on erfaring, men imidlertid kommet frem til at begge formene for eksempelbruk – domenespesifikk og abstrakt – har fordeler tilknyttet seg, og at de heller utfyller hverandre.

Del III

Forskningsspørsmål 2

Kapittel 10

Forskningsdesign for FS2

Dette kapitlet presenterer og beskriver forskningsdesignet til FS2: *“Hvilke metoder, verktøy og materialer bør benyttes ved prototyping av applikasjoner for Surface i de tidlige fasene av systemutviklingsløpet, og hvordan kan disse prototypene brukbarhetstestes ved hjelp av wizard-of-Oz?”*. Forskningsdesignet består av metodene som sammen vil danne grunnlaget for å svare på forskningsspørsmålet.

10.1 Forskningsstrategi

Formålet med denne delen av forskningen er å undersøke hvordan en applikasjon til Surface kan prototyper, og i tillegg se på hvordan brukbarhetstester av en slik prototype kan gjennomføres. Dette skal gjøres gjennom å se på hvordan et utvalg av brukere interagerer med en papirprototype av et multi-touchbord.

Vi vil med utgangspunkt i ideene fra workshopene under besvarelsen av FS1 utvikle et konsept som prototyper og deretter testes på et utvalg brukere. Under testene vil vi se på hvordan det vil fungere å brukbarhetsteste flere personer simultant på én skjerm. Vi avgjorde i samråd med veileder at det ville være tilstrekkelig å teste *to* testdeltakere samtidig for å kartlegge utfordringene rundt flerbrukertesting.

En alternativ strategi for å svare på forskningsspørsmålet kunne vært å inkludere brukerne i utviklingen av prototypen. Eksempelvis kunne vi bedt deltakere om å prototype en gitt applikasjon ut ifra et sett med materialer og verktøy vi gjorde tilgjengelig for dem. På den måten ville vi sett

hvordan brukerne ville angrepet problemene forbundet med prototyping for Surface. Vi tror imidlertid ikke dette ville gitt oss like dyp granskning av utfordringer knyttet til prototyping. Dessuten ville denne strategien forutsatt at vi allerede hadde visst at slik prototyping lar seg gjøre. Bakgrunnen for forskningsspørsmålet er at det er for tidlig å si noe om hvorvidt prototyping for Surface er mulig, og det er nettopp av denne årsaken vi utforsker materialer og teknikker selv.

10.2 Hva ønsker vi å finne ut?

Det vi ønsker å finne ut kan skilles i to deler:

1. hvordan Surface-applikasjoner kan low-fi prototypes
2. hvorvidt en wizard-of-Oz brukbarhetstest av en slik prototype vil avdekke brukbarhetsfeil

For å kunne svare på dette må vi (1) eksperimentere med ulike metoder og materialer for prototyping og (2) gjennomføre brukbarhetstestene som i en reell situasjon og lete etter brukbarhetsfeil. På denne måten vil vi “ha på oss brukbarhetstest-hatten” under gjennomføringen, og gå ut av denne når testseansene er ferdig. Vi ønsker dessuten å få svar på om testing av to personer parvis er en god måte å flerbrukerteste slike prototyper på. Testene vil dessuten gi svar på hvilke utfordringer testing av en low-fi prototype for Surface byr på i forhold til testing av prototyper for de mer tradisjonelle plattformene (som web-grensesnitt og PC-applikasjoner).

For å kunne gi svar på dette har vi utarbeidet en rekke konkrete spørsmål knyttet til prototypen og teststrategien (se Tabell 10.1). Disse er basert på de fire konseptene som sammen dekker egenskapene ved Surface (se Seksjon 7.1). Spørsmålene vil både besvares gjennom en spørreundersøkelse og brukes som temagrunnlag i et semi-strukturert gruppeintervju.

10.3 Research through design

Under arbeidet med prototypen vil det utforskes forskjellige kreative teknikker og materialer som kan brukes for å prototype en applikasjon for Surface. Denne prosessen vil beskrives i detalj og involverer teknikker og materialer

Tabell 10.1: Spørsmål vi ønsker å besvare gjennom brukbarhetstestene.

Prototypen	Simulering av multi-touch	Klarer prototypen å simulere multi-touch?
		Er det vanskelig å forstå hva de ulike elementene skal representere?
		Er det forståelig at man kan rotere, skalere og flytte på GUI-elementer?
	Fysisk kontra virtuelt	Er det lett å skille mellom fysiske og virtuelle elementer i prototypen?
		Er det forståelig at noen virtuelle elementer er direkte avhengig av et fysisk objekt?
		Er det lett å holde rede på hvilke virtuelle og fysiske elementer som hører sammen?
		Kommer det tydelig frem at dersom man fjerner det fysiske objektet fra bordet så vil det virtuelle som er avhengig av den forsvinne?
Teststrategi	Multi-user testing	Oppfatter brukerne det som forstyrrende at det er en annen testbruker som bruker prototypen samtidig?
		Føler testbrukerne at det er naturlig å bruke skjermen samtidig eller føler de at de bruker den på tur?
	Wizard-of-Oz	Vil brukerne føle at responstiden ødelegger mye for opplevelsen av systemet?

som prøves ut; både de som fungerer og de som ikke fungerer. På denne måten vil designprosessen inngå som en del av forskningen (se Seksjon 5.6). Både prosessen med å utvikle prototypen og resultatene fra testingen av den vil bidra til kunnskap rundt low-fi prototyping for Surface-applikasjoner.

Zimmermans fire kriterier for research through design vil oppfylles på følgende måte:

- **Prosess:** Prosessen rundt utprøving av forskjellige prototypingmetoder, -verktøy og -materialer vil dokumenteres i detalj. Planen for gjennomføring av brukbarhetstestene og resultatene fra disse vil også fremstilles grundig. På denne måten vil det være mulig å gjenskape arbeidet.
- **Nyskapning:** Produktet av prototypingeksperimenteringen og brukbarhetstestene vil resultere i nyskapende metoder for evaluering av designløsninger til Surface-applikasjoner.
- **Relevans:** I en industri der utviklingen i liten grad har involvert brukerne på grunn av interaksjonskompleksiteten ved multi-touch, vil resultatene våre kunne bidra til økt brukersentrering under utvikling av Surface-applikasjoner. Dette vil føre til økt brukertilfredsstillelse og tidligere avdekking av feil.
- **Utvidningsevne:** Resultatene skal abstraheres vekk fra den konkrete prototypen, slik at de kan anvendes i andre designproblemer i Surface-applikasjoner.

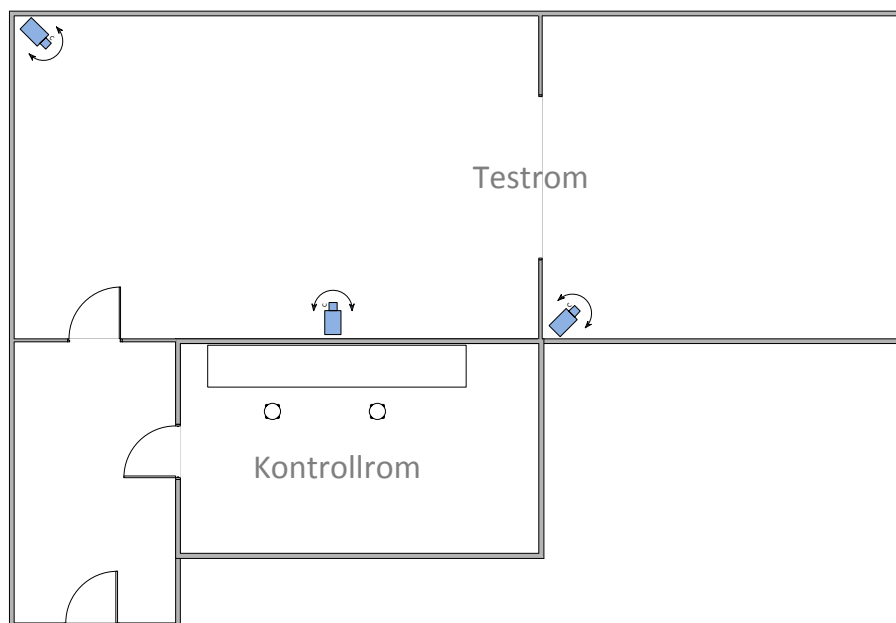
10.4 Deltakere

Når vi skal gjennomføre brukbarhetstestene ønsker vi å ha testbrukere som er i målgruppen for den applikasjonen vi skal teste. Samtidig bør de ha noe teknisk innsikt for å kunne teste en prototype for Surface uten at vi trenger å introdusere de for touchteknologien i samme grad som under workshopene (som ville vært en uønsket variabel i testene). For å finne brukere som fyller disse kravene ser vi for oss å gjøre et utvalg blant førsteårsstudentene ved informatikk, der begge kjønn er representert. Disse studentene antas å ha kjennskap til multi-touchteknologien, samtidig som de er potensielle museumsgjengere. I tillegg er dette er en type deltakere det vil være relativt lett å rekruttere.

Til tross for at studentene har noe teknologisk innsikt, vil dette kunne variere fra deltaker til deltaker. For at alle deltakerne skal ha tilnærmet likt grunnlag vil vi vise en kort demonstrasjon av Surface. Med en slik introduksjon vil deltakerne få større innsikt i egenskapene ved Surface og gjøre dem bedre rustet til å komme med konstruktiv tilbakemelding på prototypen, samtidig som vi forsikrer oss om at alle er på samme nivå.

10.5 Lokasjon

Brukbarhetstesten vil finne sted på brukbarhetslaboratoriet på NSEP. Brukbarhetslaboratoriet består av to områder: et testrom og et kontrollrom. Rommet har fastmonterte mikrofoner, samt tre styrbare kameraer i taket som kan stilles inn til å filme all interaksjon som skjer i testområdet (Figur 10.1). Kameraene kan styres og overvåkes fra kontrollrommet. Dette gjør det mulig å ta video- og lydopptak av hele seansen slik at vi får samlet inn data til videre analyse.



Figur 10.1: Skisse over brukbarhetslaboratoriet på NSEP.

Brukbarhetstestene kunne ha foregått på et ordentlig museum for å gi en mest mulig realistisk situasjon. Årsaken til at vi likevel velger å la testene finne sted på et brukbarhetslaboratorium i stedet for i en reell setting på

Vitenskapsmuseet er at vi er avhengige av å ta opptak av brukbarhetstestene for å kunne gjøre analyser av testene i etterkant. I tillegg til dette ønsker vi at testen skjer i kontrollerte omgivelser, skjermet fra andre besøkende.

For at testbrukerne skal oppleve den rette konteksten under testene, vil vi lage et fiktivt museum på brukbarhetslaboratoriet.

10.6 Vår rolle

Under gjennomføringen av brukbarhetstestene vil vi trenge en wizardrolle. Denne rollen vil utgjøres av oss. På den måten vil vi ta del i testseansene og ha en innflytelse på testpersonene. Faktorer som responstid, feilfeedback og stress kan spille inn på hvordan testpersonene oppfatter systemet. Det er derfor viktig at denne wizardrollen holder seg stabil for alle testene, slik at den ikke utgjør en variabel.

Når vi utgjør denne rollen vil det også være begrenset hvor mye vi kan observere og reflektere underveis. For å studere interaksjonen mellom testbrukere og wizarder, og få muligheten til å reflektere over vår egen påvirkning, vil vi derfor analysere videoopptakene fra brukbarhetstestene. På denne måten vil vi få avdekket hvorvidt vi utgjør en variabel.

10.7 Observasjon

Ettersom vi vil konsentrere oss om wizardrollen under testene kommer video- og lydopptakene til å være viktige for å observere deltakerne og deres interaksjon med prototypen i etterkant.

10.8 Spørreundersøkelse

For å kunne få en individuell tilbakemelding på hvordan hver enkelt deltaker oppfattet prototypens egenskaper og testopplegget vil vi gjennomføre en spørreundersøkelse i etterkant av testen. Denne spørreundersøkelsen er basert på temaene i Tabell 10.1 og er påstander som skal besvares på en Likert-skala. Dette vil gi kvantitative data på deltakernes inntrykk. Spørreundersøkelsen kan studeres i Tillegg C.4.

Det vil også gjennomføres en spørreundersøkelse som tar for seg brukbarheten til applikasjonen som prototypes. Dette vil gi tallbaserte data på eventuelle brukbarhetsfeil som prototypen klarer å avdekke. Denne vil være en variant av det standardiserte SUS-skjemaet (System Usability Scale) som vanligvis brukes for å evaluere brukbarhetstester. Denne spørreundersøkelsen presenteres i Seksjon 11.5.4 der oppsettet for brukbarhetstesten legges frem.

10.9 Intervju

I kombinasjon med spørreundersøkelsene ønsker vi å kjøre en form for semi-strukturert intervju slik at testpersonene får si hvordan de følte testene gikk, og komme med eventuelle innspill som spørreskjemaene ikke dekker. Dette gir oss mulighet til å få en dialog med deltakerne. Deltakerne får mulighet til å styre samtalen, samtidig som vi sørger for å holde den innenfor temaene. Disse temaene vil være de samme som spørsmålene i Tabell 10.1.

En ulempe med å intervjuer deltakerne i par er imidlertid at den ene parten kan være mer fremtredende enn den andre, og vi vil derfor risikere at vi ikke får alle deltakernes sanne meninger (se Seksjon 5.3). For å motvirke denne risikoen vil vi passe på å henvende oss direkte til begge deltakerne og sørge for at alle synspunkter kommer frem.

Hensikten med intervjuet er å kunne avdekke eventuelle problemer med prototypen og opplegget som spørreskjemaene ikke tar for seg. Svarene deltakerne gir gjennom spørreskjemaene avhenger i stor grad av hvordan de ulike utsagnene er formulert. Ved bruk av spørreskjemaer er det ikke mulig å korrigere eventuelle misforståelser som har oppstått hos respondentene. Etter som tolkningen av utsagnene kan variere fra person til person kommer vi til å gå nøye gjennom de ulike utsagnene i plenum slik at deltakerne har mulighet til å stille spørsmål underveis.

Kapittel 11

Prototyping og Testforberedelser

Dette kapitlet omhandler utvikling, prototyping og brukbarhetstesting av et applikasjonskonsept for Surface. Konseptet for applikasjonen er basert på ideer som ble fremstilt under workshopene på Vitenskapsmuseet. Følgelig er systemet tenkt å kunne implementeres på museer generelt og Vitenskapsmuseet spesielt.

Seksjon 11.1 presenterer konseptet for applikasjonen og hvordan dette ble utviklet. Seksjon 11.2 viser ulike teknikker vi vurderte for prototyping av multi-touchapplikasjonen. Siden forskningen omfatter gjennomføring av brukbarhetstesting på flere personer samtidig er Seksjon 11.3 viet til drøfting av oppsettet for testene. Kapitlet avsluttes med en forklaring av den endelige prototypen (Seksjon 11.4) og strukturen for testene (Seksjon 11.5).

11.1 Konseptutvikling

For å svare på FS2 ville vi trenge et konsept til en Surface-applikasjon som kunne prototypes og brukbarhetstestes. I et faktisk utviklingsprosjekt ville utfallet av workshops med brukere dannet grunnlaget for systemet som videre utvikles. For at denne oppgaven skal følge lik gang som reelle prosjekter ble derfor ideene som spirte under workshopene lagt til grunn for valget av konsept. Det var dessuten av betydning at konseptet skulle involvere de fire Surface-egenskapene *direkte manipulasjon*, *interaksjon med fysiske objekter*, *samhandling* og *kombinasjon med andre teknologier*. Prototypen ville

da måtte klare å simulere et bredt spekter av de egenskapene som skiller Surface-applikasjoner fra tradisjonelle PC-applikasjoner.

11.1.1 Valget av idé

Under begge workshopene kom ideen “*Objekter med informasjon*” frem (se Seksjon 8.2.2 og Seksjon 8.3.2). Ideen går ut på at besøkende kan finne objekter med tags på undersiden fra tilhørende utstillinger. Disse objektene kan tas med til Surface der de besøkende kan finne mer informasjon/media om utstillingsgjenstandene. Ideen egnet seg godt som grunnlag for konseptet og fascinerte dessuten workshopdeltakerne. Den inneholder interaksjon med fysiske objekter og direkte manipulasjon. For å inkludere samhandling utvidet vi konseptet med å la brukerne få vite interessante fakta om noe som objektene har til felles dersom de føres sammen på bordet. Hensikten med dette vil være at flere brukere som interagerer med skjermen samtidig kan sammenligne hverandres objekter. Vi la også til muligheten å føre over informasjon/media til mobiltelefonen for å inkludere kombinasjon med andre teknologier (se ideen “*Informasjon om museumsgjenstander*”, Seksjon 8.3.2).

11.1.2 MultiViten

Vi valgte å kalle konseptet for *MultiViten* (satt sammen av ordene multimedia og vitenskap). Tanken bak MultiViten er å la besøkende ved museer lære mer om utstillingsgjenstandene ved hjelp av Surface.

Museumskort

Konseptet innebærer at alle utstillingsgjenstandene på museet har en tilhørende bunke med *museumskort* som ligger rett utenfor monteret/sperrebåndet. Museumskortet viser navnet på gjenstanden, et bilde og kortfattet fakta på forsiden (se Figur 11.1). På baksiden av kortet er det en Surface-tag og en tekst som henter om at kortet kan brukes på museets Surface-bord. Besøkende som er interessert i å lære mer om utstillingsgjenstanden kan ta med dette kortet til Surface og plassere det på skjermen.

For å gjøre det lettere for brukeren å forstå at museumskort kan finnes omkring på museet, vil en tekst stå på Surface-skjermen når MultiViten ikke er i bruk: “let etter museumskort og plasser dem på multimediebordet”.



Figur 11.1: Museumskortet for utstillingsgjenstanden isbjørn.

Når brukeren er ferdig med et museumskort kan det tas med hjem eller legges i en oppsamlingsboks i nærheten av Surface-bordet.

Puslespillbrikke med mediaelementer

Når kortet legges på skjermflaten vil det dukke opp en puslespillbrikke på skjermen under museumskortet. Denne brikken inneholder ulike mediaelementer som relateres til utstillingsgjenstanden kortet representerer. Mediaelementene kan være bilde, video eller lyd. Disse vises som små rektangler på brikken og kan flyttes fritt rundt, roteres og skaleres. Puslespillbrikken kan ikke flyttes på eller manipuleres direkte, men vil følge etter det fysiske museumskortet sin plassering på skjermen. Brukeren må altså flytte på det fysiske kortet for å flytte den virtuelle puslespillbrikken. De mediaelementene som ligger oppå puslespillbrikken når museumskortet flyttes på vil følge etter, i kontrast til mediaelementene som har blitt flyttet vekk fra puslespillbrikken.

De elementene som er *bilder* kan bare flyttes, roteres og skaleres. Det vil ikke skje noe dersom brukeren trykker på bildet. Dette er for å unngå at brukeren oppfatter bildet som en knapp og snarere forstår at elementet skal behandles som et realistisk bilde, bortsett fra skaleringssegenskapen (se Seksjon 4.4.1 om

superrealisme). *Video* vil starte avspilling av et videoklipp dersom elementet trykkes på. Videoen kan også flyttes på, roteres og skaleres. På samme måte som video, avspilles det et lydklipp når *lydelementene* trykkes på. Lydelementene skiller seg fra video ved at de er representert med et høyttalerikon som ikke kan skaleres, men bare flyttes og roteres. Flere lydklipp kan spilles av samtidig og det er opp til brukerne av applikasjonen å sette lydklippet på pause hvis det forstyrrer for et annet.

De fire kategoriene av utstillingsgjenstander ble gitt hver sin farge; rød (geologi), blå (botanikk), gul (biologi) og lilla (Fridtjof Nansen). Fargen vises på toppen av museumskortene, dekker hele den tilhørende puslespillbrikken og vises som en ramme rundt mediaelementene. Eksempelvis fikk mediaelementene til Fridtjof Nansens fredsprismedalje en lilla ramme, i likhet med puslespillbrikken og øverste linje i museumskortet. Disse fargene ble brukt for å gjøre det lettere for brukerne å holde rede på hvilke museumskort som tilhører samme kategori når flere kort ligger på skjermen samtidig. Dessuten vil det gjøre det lettere å holde rede på hvilke puslespillbrikker mediaelementene tilhører i tilfeller der det ligger en stor mengde elementer på skjermen. Dette er spesielt nyttig siden mediaelementer og puslespillbrikken vil forsvinne når museumskortet fjernes, og mediaelementene ikke nødvendigvis ligger på puslespillbrikken når det skjer.

Dersom museumskortet fjernes fra bordet vil puslespillbrikken og alle mediaelementene som hørte til denne forsvinne, uavhengig av om elementet befinner seg oppå puslespillbrikken eller ikke.

Visste du at...?

Hvis en eller flere brukere har funnet museumskort og plassert de på Surface slik at det ligger flere av dem på skjermen samtidig skal de kunne oppdage noe nytt ved å kombinere kort. Når to eller flere kort flyttes mot hverandre slik at puslespillbrikkene under kobles sammen vil brukerne få vite interessante fakta relatert til det utstillingsgjenstandene har til felles. Tanken er at puslespillbrikkeformen vil være en affordance for brukerne av applikasjonen og dermed hjelpe de med å forstå at brikkene kan settes sammen. Når puslespillbrikker kobles sammen forsvinner de tilhørende mediaelementene og byttes ut med en snakkeboble som går ut fra koblingspunktet. Denne boblen inneholder et kort, skriftlig faktum, formulert på måten “visste du at...”. Hvis en konkret utstillingsgjenstand ikke har noe til felles med en annen vil faktumet baseres på høyere generaliseringer (utstilling eller fagfelt).

For å eksemplifisere en slik “visste du at...” ser vi for oss at Ola har funnet

et museumskort for utstillingsgjenstanden “Fridtjof Nansens fredsprismedalje”, mens Kari har fått tak i et kort for dyret isbjørn. Begge har lagt kortene på Surface og fører kortene mot hverandre så puslespillbrikkene kobles sammen. mediaelementene forsvinner og en snakkeboble dukker opp med teksten “visste du at Nansen og Johansen ble angrepet av en isbjørn på vei mot Nordpolen i 1896?”. Siden selve fredsprismedaljen til Nansen og en isbjørn ikke har noe til felles leter systemet i øvre abstraksjoner og finner noe felles mellom utstillingen om Fridtjof Nansen og dyret isbjørn.

Flere enn to museumskort kan kobles sammen og sjansen for at faktumet blir mer abstrakt i forhold til de konkrete gjenstandene øker.

For å lukke snakkeboblen og se museumskortenes mediaelementer igjen må brukerne føre kortene vekk fra hverandre. Museumskortene vil da initieres som om de var lagt ned på skjermen på nytt (alle mediaelementene nedskalert og plassert oppå puslespillbrikken).

Overføring til mobiltelefon

Dersom brukeren aktiverer bluetooth på mobiltelefonen sin vil et ikon som representerer denne dukke opp i en liste i kanten av skjermen. Ved å flytte et mediaelement over dette ikonet vil applikasjonen sende elementet til mobiltelefonen. Brukeren kan godta elementet og se bildet, spille av videoen eller høre på lydklippet på sin mobiltelefon.

11.1.3 Scenarier

For å konkretisere og eksemplifisere MultiViten ble det utviklet tre scenarier (se Tillegg B). Scenariene viser Petter og Trine i deres møte med systemet en dag de besøker Vitenskapsmuseet sammen. Første scenario viser Trine som finner et museumskort og tar det med til Surface (Tillegg B.1). I det andre scenarioet finner også Petter et museumskort som han tar med til bordet mens Trine står der (Tillegg B.2). De skyver kortene sine mot hverandre og oppdager at gjenstandene har noe til felles. I siste scenario aktiverer Petter bluetooth på mobiltelefonen sin og fører over et bilde fra Surface til den (Tillegg B.3).

11.2 Prototypingsteknikker

Uvikling av low-fi prototyper for multi-touchapplikasjoner, og da spesielt til Surface-plattformen, skiller seg fra prototyping av tradisjonelle PC-applikasjoner. Statistiske skjermbilder med knapper og menyer preger mange grafiske grensesnitt på tradisjonell PC og er derfor relativt ukompliserte å papirprototyper og brukbarhetsteste. Grensesnittet i applikasjoner på Surface baseres overveiende på direkte manipulasjon av dynamiske elementer og bruk av fysiske objekter på skjermen. Dette krever nye teknikker og metoder for prototyping enn tidligere brukt for prototyping av PC-applikasjoner.

Under arbeidet med å utvikle prototypen for konseptet ble vi oppmerksomme på at man hovedsakelig står overfor to utfordringer når Surface-applikasjoner skal prototyperes; (1) hvordan simulere multi-touch (Seksjon 11.2.1) og (2) hvordan skille mellom fysiske og virtuelle elementer (Seksjon 11.2.2).

11.2.1 Simulering av multi-touch

Simulering av multi-touch innebærer å utvikle en prototype som inviterer til bruk av gestures på GUI-elementer. Prototypen må dessuten støtte flere slike kontaktpunkter samtidig for å kunne simulere *multi-touch*.

I tilfellet av MultiViten vil prototypen være nødt til å formidle til testbrukerne at mediaelementene på puslespillbrikken kan:

- flyttes rundt på skjermen ved å ta direkte på dem og “skyve”,
- roteres ved å bruke to fingre på elementet og rotere det rundt sin egen akse,
- skaleres opp ved å bruke to fingre på elementet og dra fingrene lengre vekk fra hverandre,
- skaleres ned ved å bruke to fingre på elementet og dra fingrene nærmere hverandre.

Prototypen må også støtte at flere hender gjør disse gestikuleringene samtidig på skjermen. Det inkluderer også at flere museums kort kan legges ned og operere på skjermen til enhver tid.

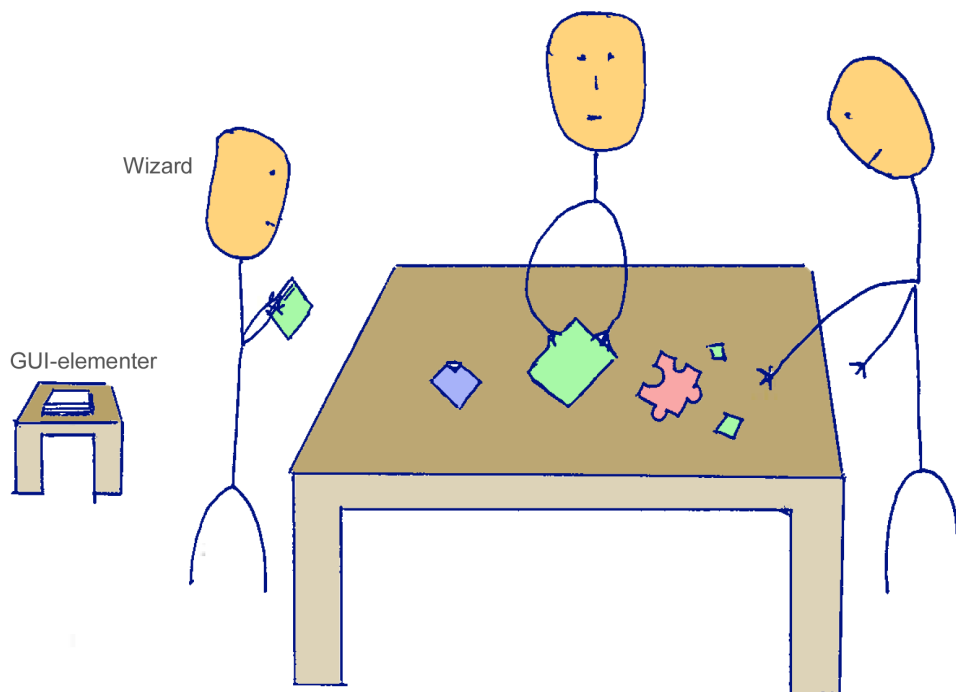
Papir

Hovedhensikten med Surface og ideen bak NUI generelt er å gjøre interaksjonen naturlig for brukeren ved å gjenspeile virkeligheten i brukergrensesnittet (se Seksjon 4.1.3). Gjennom direkte manipulasjon kan brukeren ta på, rotere og flytte rundt på elementer som vises på skjermen, tilsvarende slik fotografier og dokumenter kan organiseres på arbeidspulten i realiteten. Det var derfor naturlig å vurdere papir som materiale til prototypingen når det er nettopp dette Surface etterligner egenskapene til.

Ved å bruke papirlapper til å representere puslespillbrikkene og mediaelementene som hører til, vil brukerne kunne flytte og rotere disse fritt på et bord. Det papirlappene ikke direkte klarer å simulere er superrealisme (se Seksjon 4.4.1), nemlig egenskaper som går utover det som er realistisk for ikke-virtuelle materialer. Skalering er én slik egenskap. Papirlapper kan naturligvis ikke skaleres opp og ned i virkeligheten. En løsning på dette er å la testbrukere gestikulere tegn på papirlappen om den skal skaleres opp eller ned; to fingre som starter på lappen og går vekk fra hverandre betyr at lappen skal skaleres opp og to fingre som går fra ytterkantene av lappen mot sentrum betyr at lappen skal skaleres ned. Lappen kan så byttes ut med en større eller mindre versjon ved å bruke wizard-of-Oz-teknikken (se Seksjon 3.5.2).

Når man skal bruke papir i en prototype med direkte manipulasjonsinteraksjon må man også være klar over de affordancene papir har som virtuelle elementer savner. Papir kan løftes, brettes, vendes, krølles og rives. Det gjør at testbrukere eksempelvis vil føle det naturlig å flytte rundt på de forskjellige GUI-elementene ved å løfte de i stedet for å behandle de som virtuelle skjermbilder som bare kan flyttes i det todimensjonale skjermplanet. Om dette er en kritisk problemstilling eller ikke avhenger av hva prototypen skal teste. Dersom prototypens hensikt er å simulere en multi-touchapplikasjons look and feel er dette en hindring (se Seksjon 3.5.5). Demonstrerer derimot prototypen hvordan funksjonene i applikasjonen fungerer og dermed hvilken rolle applikasjonen skal spille i brukerens situasjon er det ikke kritisk hvorvidt brukeren løfter eller skyver på et element for å flytte på det. Det viktigste da er at brukeren forstår *at* elementet kan flyttes.

En skisse av bruken av papir for prototyping av multi-touch er gjengitt i Figur 11.2.



Figur 11.2: Skisse av papirprototyping for multi-touch.

Projisering

Et alternativ til den tradisjonelle papirprototypingen er å simulere applikasjonen på en ekstern datamaskin som er koblet til en projektor. Projektoren kan da monteres på undersiden av en delvis gjennomsiktig glassflate og projiseres mot denne slik at testbrukeren ser skjermbildet på glassflaten (se Figur 11.3). Dette er samme prinsipp som brukes i Surface. Den eksterne datamaskinen kan alternativt kobles til en skjerm som legges horisontalt og la brukerne interagere direkte på denne.

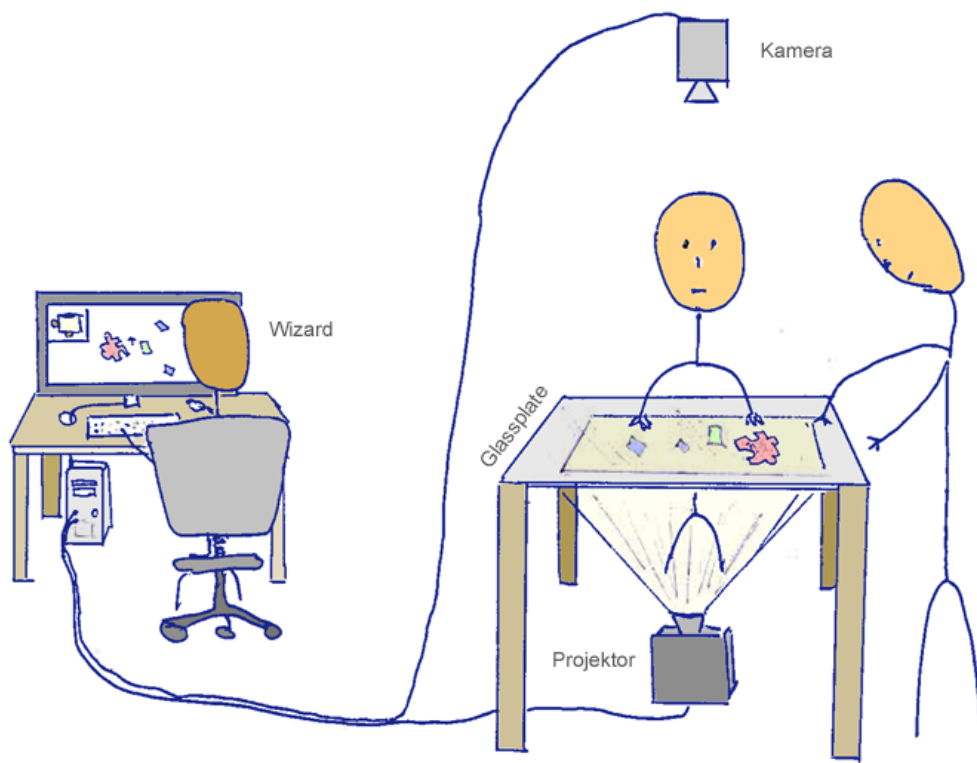
På samme måte som en papirprototype trengs det her en wizard som følger med på hva brukeren foretar seg på glassflaten/skjermen og endrer skjermbildet på datamaskinen i samsvar med dette. Wizarden kan enten ha denne flaten innen synsrekkevidde eller følge med via et kamera montert på oversiden. Det som vises på skjermen kan være GUI-elementer som flyttes rundt i et bilderedigeringsprogram.

Fordelene med å bruke projiseringsteknikken er at brukeren får en sterkere følelse av at prototypen er en dataapplikasjon og de uønskede affordancene med papir elimineres. Prototypen blir mer realistisk ved at GUI-elementene bare kan flyttes i det todimensjonale skjermplanet. Det er likevel ulemper relatert til denne teknikken;

- Den direkte manipulasjonen må simuleres av wizarden. Dette gjør at elementene vil “henge etter” når testbrukeren vil flytte og rotere dem. Det er også vanskelig for wizarden å beregne hvor GUI-elementer skal være på sin egen skjerm i forhold til hvor testbrukeren har fingrene på glassflaten/skjermen.
- Med én datamus vil ikke wizarden kunne manipulere mer enn ett GUI-element til enhver tid. Testbrukeren kan eksempelvis ikke føre to puslespillbrikker mot hverandre samtidig.
- Det blir raskt komplisert når man vil teste flere enn én testbruker samtidig.
- Det kreves teknisk utstyr utover det trivielle.

11.2.2 Skille mellom fysiske og virtuelle elementer

En stor utfordring ved prototyping for Surface er å skille mellom fysiske og virtuelle elementer. Når fysiske gjenstander med tags er en del av Surface-



Figur 11.3: Skisse av projiseringsteknikken ved prototyping for Surface.

applikasjonen som skal prototypes bør det til enhver tid være tydelig for testbrukeren hvilke av objektene som er slike. Hvis testbrukeren ikke forstår dette eller glemmer seg bort risikerer man at han behandler det fysiske objektet som noe virtuelt.

Å tydeliggjøre dette skillet blir spesielt viktig når de fysiske objektene er tynne papirstykker, som museumskortene i MultiViten. Dette er ikke et problem hvis projiserings-teknikken brukes for å simulere multi-touch, siden det da er tydelig for testbrukeren at det som ligger på skjermflaten er fysisk og det som vises på skjermen er virtuelt. Hvis det derimot er en papirprototype som testes, kan kortene lett oppfattes som en del av det virtuelle grensesnittet, siden papir brukes til å representere alle GUI-elementene. Ytterligere teknikker bør derfor brukes for å tydeliggjøre denne forskjellen.

For å redegjøre hvordan prototypen av MultiViten kunne papirprototypes, utforsket vi ulike ideer til teknikker og materialer for å skille mellom fysiske og virtuelle elementer.

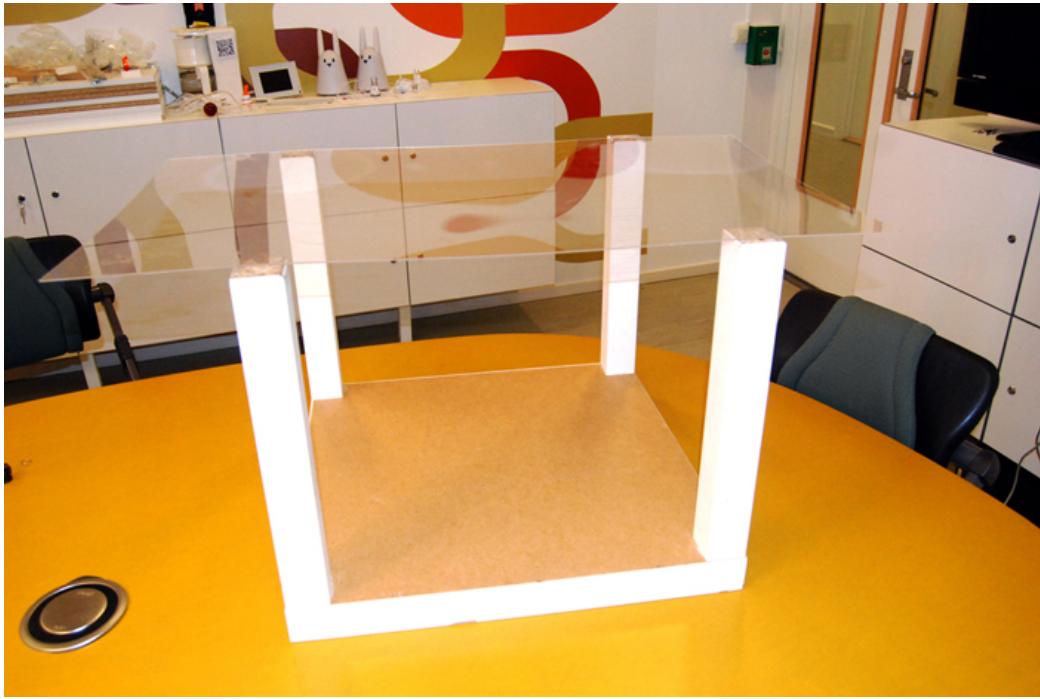
Glasskille

En tidlig idé vi hadde var inspirert av projisering og fremstår som en hybrid mellom nettopp det og papirprototyping. En glassplate ble brukt som Surface-skjerm. Denne ble plassert på bordbena til et IKEA-bord som var snudd på hodet (se Figur 11.4).

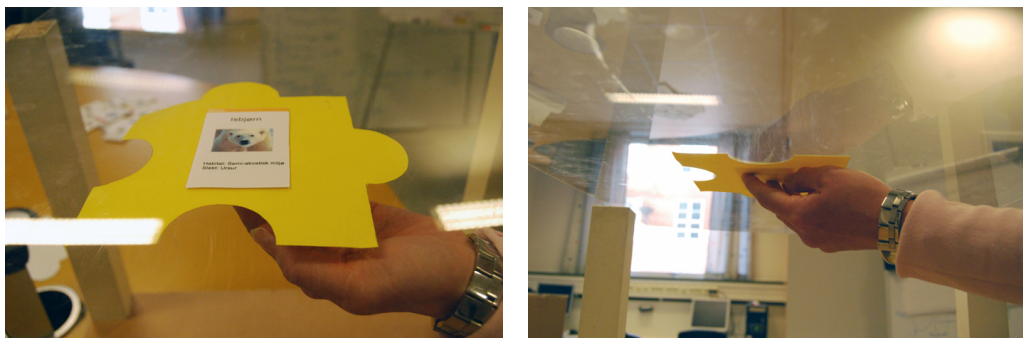
Testbrukeren kan legge museumskort på oversiden av glassplaten og bruke fingrene på glasset for å interagere med prototypen. På undersiden av glasset opererer en wizard. Wizarden presser GUI-elementer av papir mot undersiden av glasset slik at testbrukeren ser det på oversiden (se Figur 11.5). På denne måten har ikke testbrukeren direkte kontakt med papirbitene som representerer virtuelle elementer og kan utelukkende interagere med de ved å bruke glassplaten som grensesnitt. Når testbrukeren så manipulerer elementene med fingre eller fysiske objekter på glassplaten må wizarden sørge for at:

1. elementene følger testbrukerens gestures når elementer flyttes, roteres eller skaleres,
2. skjermbilder dukker opp og fjernes når de rette hendelsene inntreffer.

Selv om dette er en gjennomførbar måte å teste en papirprototype samtidig som skillet mellom det fysiske og virtuelle ivaretas, oppdaget vi at teknikken byr på problemer:

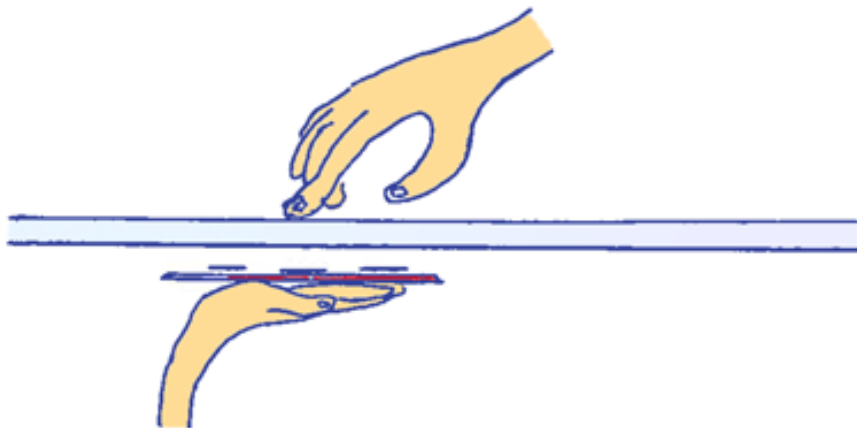


Figur 11.4: Glassplate på bordben.



Figur 11.5: Wizard som presser en virtuell puslespillbrikke mot undersiden av en glassplate når et fysisk museumskort ligger på oversiden.

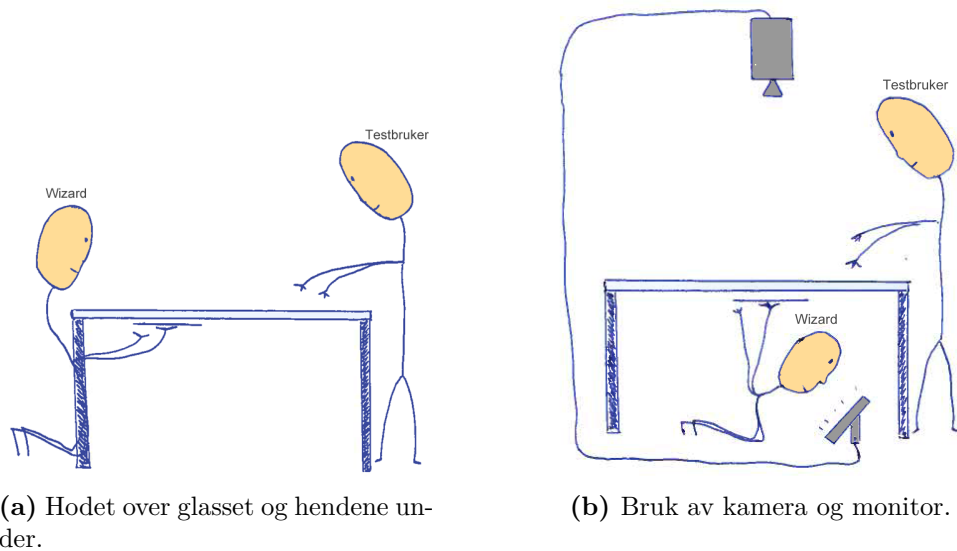
- Skjermbilder som består av flere lag gjør det vanskelig å behandle øvre lag. I MultiViten ligger mediaelementer oppå en puslespillbrikke. Der som testbrukeren ønsker å flytte et slikt mediaelement av puslespillbrikken kan ikke wizarden nå denne uten å senke papirlappene ned fra undersiden av skjermen (se Figur 11.6).
- Siden papirlappene må presses mot undersiden av skjermen begrenses prototypen av at alle elementene som vises på skjermen må kunne holdes oppe av to hender. Optimalt sett bør dessuten wizarden ha en hånd fri til eksempelvis å flytte mediaelementer vekk fra puslespillbrikken.
- Wizarden må se hva som skjer på oversiden av glassflaten. Dette kan gjøres ved å sitte i en stilling på siden av glassplaten der han både ser overflaten samtidig som armene rekker under bordet (se Figur 11.7a). Dette setter begrensninger på bordets størrelse. En annen mulighet er å filme oversiden av bordet og vise det direkte til wizarden på en monitor under glassflaten (se Figur 11.7b). Dette gjør at prototypen umiddelbart blir mer komplisert.



Figur 11.6: GUI-elementer som ligger oppå andre kan ikke flyttes uten å senke alle GUI-lagene.

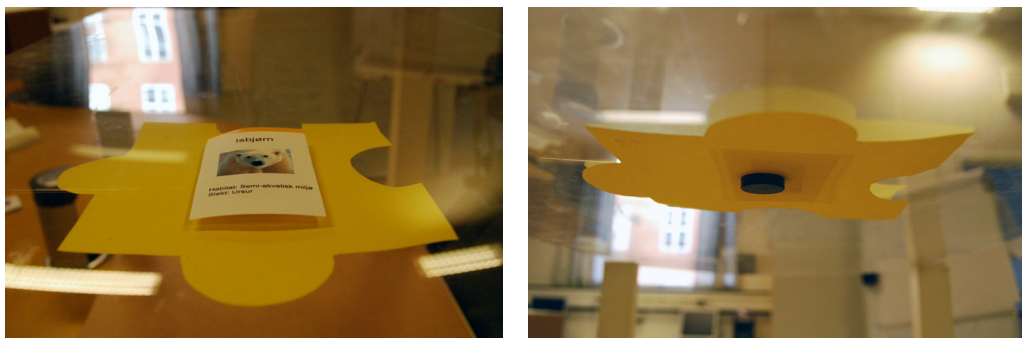
Magneter

Ulempene forbundet med glasskillet ga oss en idé som inkluderte bruk av magneter. For å teste ut ideen festet vi magneter på puslespillbrikkene og



Figur 11.7: Stillinger wizarden kan sitte i for å se hva som foregår på overflaten.

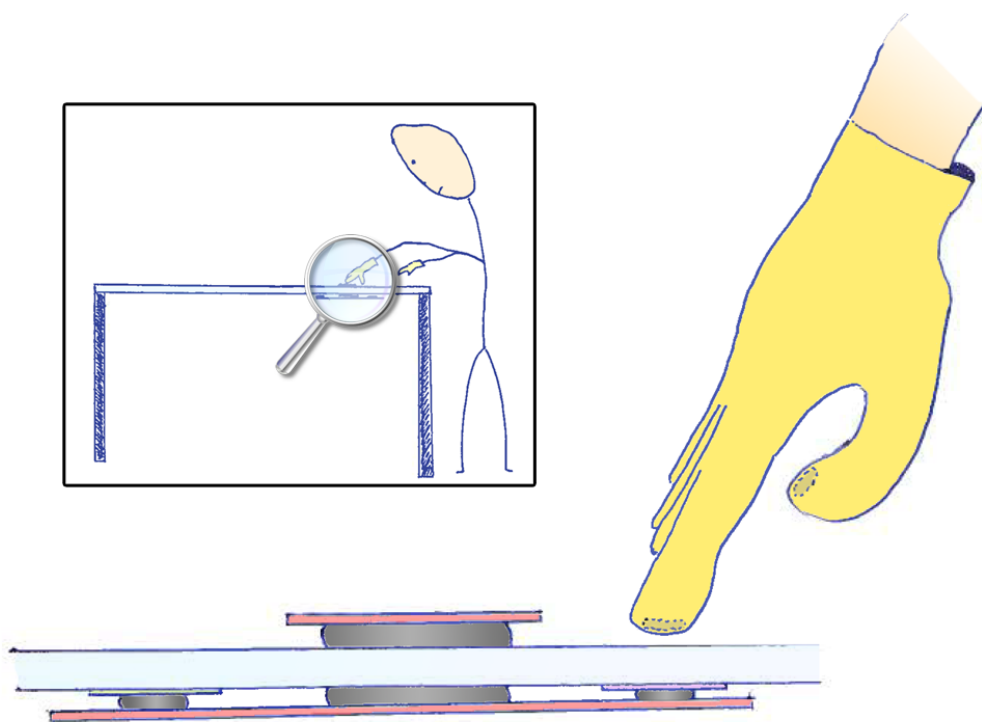
på museumskortene, og undersøkte om magnetkraften sørget for at museumskortet holdt den tilhørende puslespillbrikken oppe gjennom glassplaten. Dette fungerte bra når sterke nok magneter ble brukt (se Figur 11.8). Denne teknikken eliminerer problemet forbundet med at wizarden må presse puslespillbrikkene mot undersiden av glassplaten. Når testbrukeren beveger på museumskortet vil også puslespillbrikken følge etter uten wizardens hjelp, på grunn av magnetkraften.



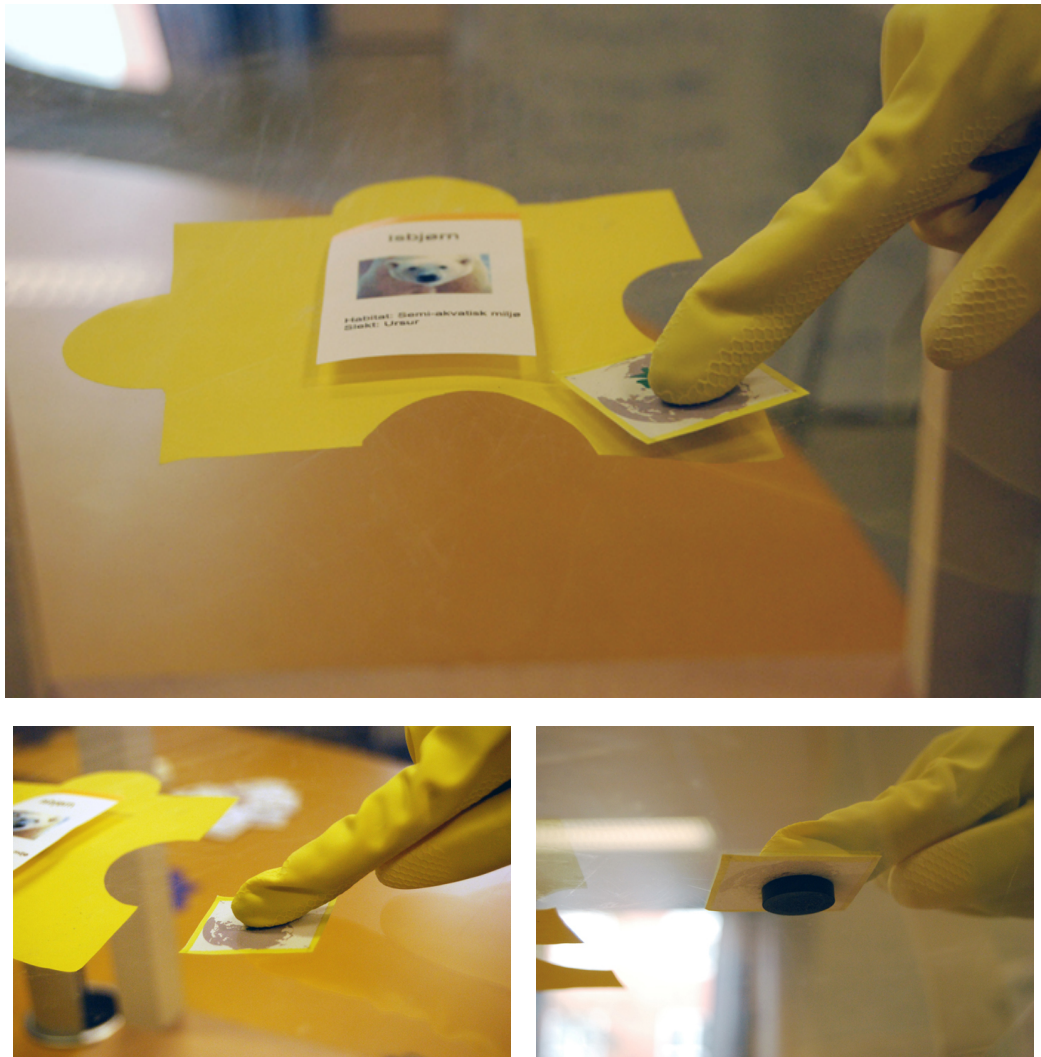
Figur 11.8: Bruk av magneter på museumskortet og puslespillbrikken.

Magnetteknikken løser likevel ikke problemet med lagvise GUI-elementer. Testbrukeren kan ikke flytte mediaelementene vekk fra puslespillbrikken uten

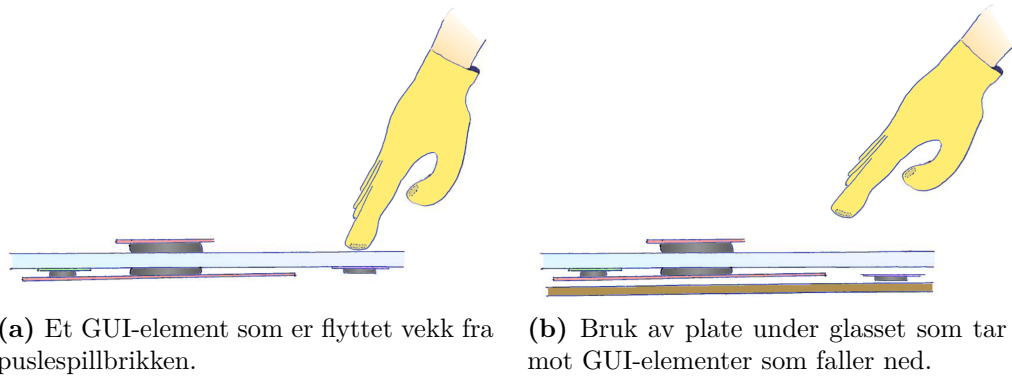
hjelp fra wizarden. Vi undersøkte derfor om det lot seg gjøre å feste magneter på alle mediaelementene, og samtidig utstyre testbrukeren med *magnetiske hansker*. Figur 11.9 viser en illustrert skisse av dette prinsippet, mens Figur 11.10 viser teorien i praksis. De magnetiske hanskene var konvensjonelle oppvaskhansker med magneter i fingertuppene. Tanken var at med disse hanskene kunne testbrukeren flytte og rotere mediaelementene vekk fra puslespillbrikken og rundt på glassflaten (se Figur 11.11a). Vi innså likevel raskt at dette ikke fungerte, siden glassplaten da ville trenge en ny plate på undersiden som støttet disse magnetiske papirlappene fra å falle ned når testbrukeren ikke berørte dem lenger (se Figur 11.11b). Denne støtteplaten under ville vanskeliggjøre arbeidet for wizarden med å “koble” på og fjerne de magnetiske papirlappene når museumskort ble lagt ned og tatt av glassplaten. Vi observerte også at de magnetiske GUI-elementene hadde en uforutsigbar oppførsel, spesielt når de skulle roteres.



Figur 11.9: Skisse av bruk av magneter på GUI-elementer og magnetiske hansker.



Figur 11.10: Hansker med magneter i fingertuppene interagerer med magnetiske GUI-elementer.



Figur 11.11: Når en testbruker løfter hånden vil magnetkraften mellom GUI-elementet og hansken opphøre.

Papp

Etter å ha erfart hvilke skyggesider glassplaten og magnetene bar på, valgte vi å fokusere på teknikker for å skille mellom fysiske og virtuelle elementer når prototypen bestod av GUI-elementer av papir *oppå* en flate. Med fokus på MultiViten så vi konkret på måter å skille museumskortene, som forestilt i konseptet ville være tynne papirkort i visittkortstørrelse, fra de virtuelle GUI-elementene i prototypen.

En distinkt forskjell mellom de fysiske og de virtuelle elementene i en Surface-applikasjon er at de *fysiske elementene kan gripes*. De kan løftes opp, legges ned, flyttes og roteres ved å gripe fatt i objektet. En god måte å skille de fysiske elementene fra de virtuelle er derfor å gi de en gripevennlig affordance. Tanken bak dette er å invitere testbrukeren til å skille mellom fysiske og virtuelle elementer uten å tenke over det. For å gi de fysiske museumskortene et mer gripevennlig preg enn de virtuelle puslespillbrikkene og medielementene brukte vi tynn bølgepapp (ca. 2 mm tykkelse). En for- og bakside for museumskortet ble limt på hver side av pappbiten. Til de virtuelle GUI-elementene ble det brukt vanlig ark. Figur 11.12 demonstrerer gripeeviden til museumskortet når det ble brukt papp.

Postit-lim

Den virtuelle puslespillbrikken er direkte avhengig av det fysiske museumskortet som hører til. Når museumskortet legges ned, dukker puslespillbrikken opp på skjermen under kortet og når kortet fjernes, forsvinner puslespill-



Figur 11.12: Museumskort i papptykkelse.

brikken. Puslespillbrikken vil alltid befinne seg under museumskortet når det ligger på skjermen, også når kortet flyttes. Puslespillbrikken må derfor følge etter kortet når det flyttes rundt på skjermen.

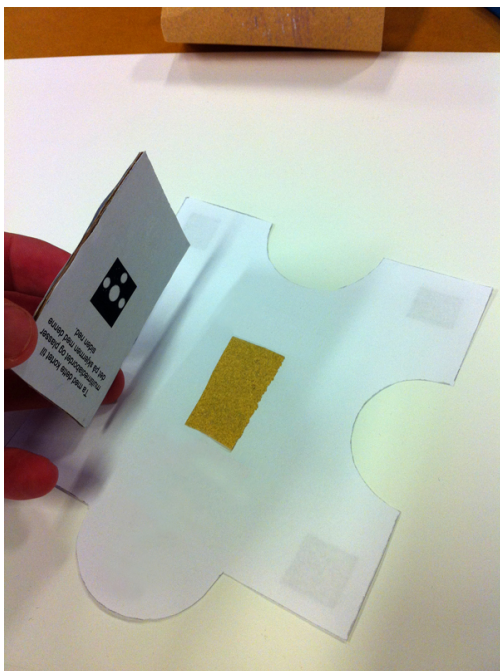
Hvis glasskille-teknikken brukes er wizarden nødt til å sørge for at puslespillbrikken følger etter. Dersom prototypen bruker papir og papp oppå en flate slipper man dette. Når museumskortet legges ned, løfter wizarden simpelthen kortet og legger puslespillbrikken under. Testbrukeren kan så skyve museumskortet rundt og puslespillbrikken vil følge skyvet. Vi erfarte imidlertid at det var for lite friksjon mellom de to elementene til at de fulgte hverandre. Et forsøk ble gjort med bruk av postit-lim (lim som ikke herdes) for å skape mer friksjon. Limet ble påført undersiden av museumskortet. Dette løste problemet, men skapte et nytt; puslespillbrikken klebet seg fast i museumskortet og fulgte etter når sistnevnte ble løftet vekk fra bordet.

Sandpapir

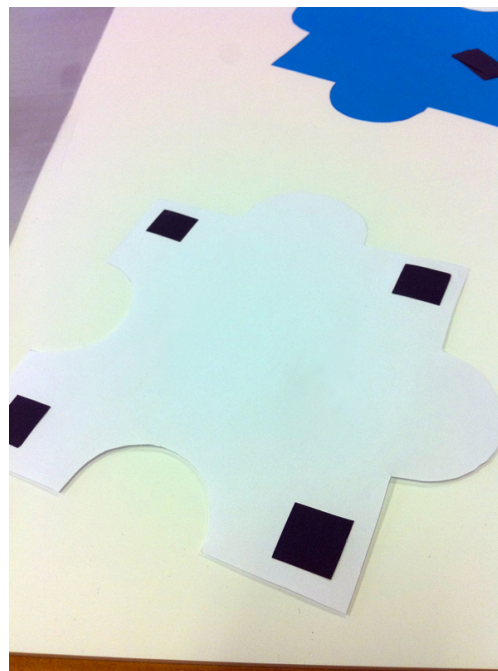
Veilederen vår rådet oss til å prøve ut sandpapir i stedet for postit-lim. Vi anskaffet sandpapir av tre forskjellige grovhetsgrader. Etter litt prøving og feiling fant vi en god løsning for bruken av sandpapir. Av det groveste sandpa-

piret klippet vi ut rektangler som var ca. halve størrelsen av museumskortene. Disse rektanglene ble limt fast på oversiden av puslespillbrikkene, der museumskortene ville ligge (se Figur 11.13a). Dette sandpapiret skapte tilstrekkelig friksjon mellom museumskortet og puslespillbrikken til at sistnevnte lot seg indirekte flytte av brukerens direkte kontakt med museumskortet.

Det ble også klipt ut små biter av det fineste sandpapiret som ble festet i hjørnene på undersiden av puslespillbrikken. Dette bidro til å skape den rette constrainten (se Seksjon 2.3.2) med tanke på at puslespillbrikkene ikke skal la seg flytte direkte ved å ta på dem, men *må* flyttes indirekte gjennom museumskortene (se Figur 11.13b). Dersom testbrukeren prøver å flytte puslespillbrikken med fingrene vil sandpapiret på undersiden “lugge” og gjøre flyttingen vanskelig. Siden det ikke er sandpapir under midten av puslespillbrikken glir den lettere når den flyttes av museumskortet.



(a) Grovt sandpapir for å skape friksjon mellom museumskort og puslespillbrikke.

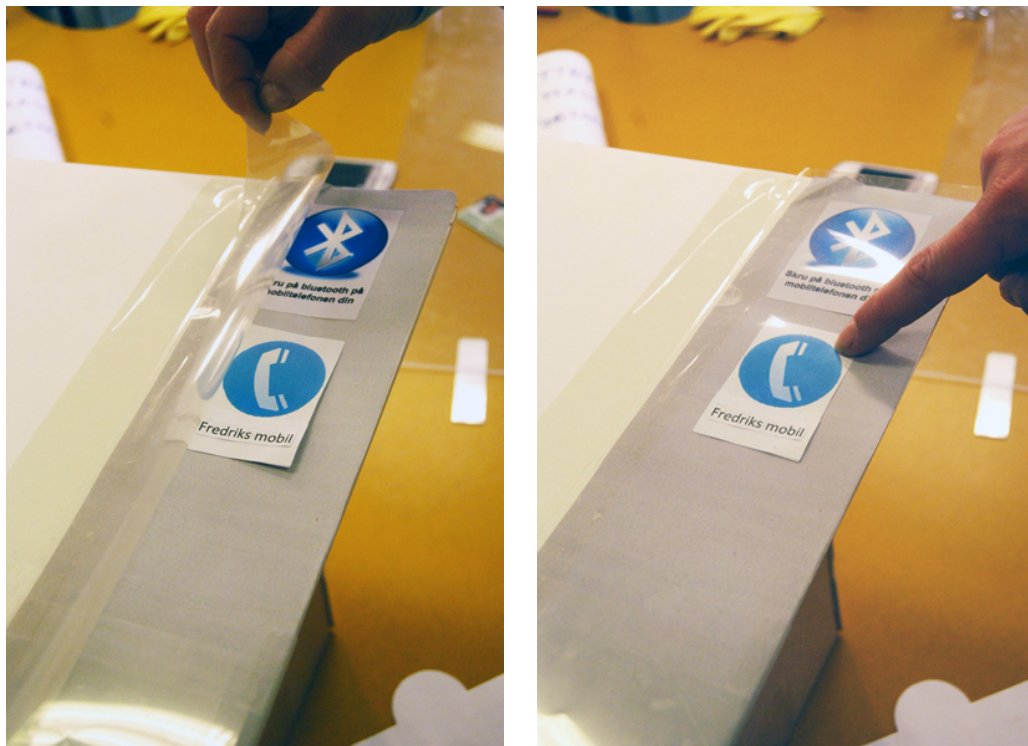


(b) Biter av fint sandpapir på undersiden for å skape friksjon mellom puslespillbrikke og underlag.

Figur 11.13: Bruk av sandpapir på puslespillbrikkene.

Transparente ark

Selv om brukergrensesnittet i MultiViten i hovedsak er dynamisk og består av elementer som kan flyttes, roteres og skaleres, inneholder den noen statiske GUI-elementer. Ei liste langs den ene kortsiden er dedikert til bluetooth-enheter. Den inneholder først og fremst et bluetooth-ikon med et hint om at brukerne kan aktivere bluetooth på mobiltelefonen sin. Når Surface oppdager bluetooth-enheter i nærheten blir disse enhetene representert med hvert sitt telefonikon med navn på enheten i lista. Verken bluetooth- eller telefonikonene vil reagere på direkte kontakt med brukernes fingre. Det er meningen å føre mediaelementer over et telefonikon. I prototypen kan denne constrainten skapes ved å bruke transparente ark over elementene (se Figur 11.14). Testbrukeren vil kunne se GUI-elementene, men ikke nå dem direkte med fingrene.



Figur 11.14: Bruk av transparenter for å hindre flytting av statiske GUI-elementer.

11.2.3 Hvilke teknikker ble brukt?

Tabell 11.1 oppsummerer teknikkene og materialene vi tok under vurdering. Den viser hvilke av teknikkene som ble brukt i den endelige prototypen og hvorfor.

11.3 Testteknikker

MultiViten er en applikasjon som skal kunne brukes av flere brukere samtidig. Det spesielle med applikasjoner til multitouch-bord er at brukere interagerer på samme skjerm. For å skape en realistisk testsetting bør derfor prototypen testes av flere enn én bruker samtidig. Utfordringen er å ha et testopplegg som koordinerer flere testbrukere på en og samme gang, samtidig som testene gransker prototypens brukbarhet.

11.3.1 Parallele oppgavesett

For å unngå å lage GUI-elementer for hele applikasjonen er det en god løsning å ha forhåndsdefinerte scenarier som testbrukeren skal gjennom (se Seksjon 3.5.4). Testbrukeren følger da et sett med oppgaver som vil teste et konkret tilfelle av bruk, men som likevel kan generaliseres til andre bruksscenarier. Ved testing av flere enn én testbruker samtidig må man dermed avgjøre om testbrukerne skal ha felles oppgaver som må løses eller adskilte oppgavesett.

Brukere av et Surface-bord i offentlige rom vil kunne være alt fra helt ukjente for hverandre til nære bekjente. Noen brukere vil komme og gå, mens andre blir stående ved bordet. Vi ser for oss en type bruk av MultiViten der brukere vil kunne interagere med applikasjonen for seg selv og i felleskap med andre dersom sjansen byr seg. For å teste begge typene av bruk vil en løsning være å gi testpersonene parallele oppgavesett der noen av oppgavene er helt individuelle, mens andre involverer den andre parten. Oppgavene vil være personlige for testbrukeren og snarere *oppfordre* ham til å interagere i felleskap, fremfor å tvinge frem dette samarbeidet gjennom felles oppgaver.

Tabell 11.1: Teknikker brukt til prototyping av MultiViten

Teknikk/materiale	Ble brukt i endelig prototype?	Begrunnelse
Papir	Ja	Enkel, rask og billig prototyping-metode. Får vist essensen i applikasjonen som prototypes.
Projisering	Nei	Dyr og komplisert. Avgrenses av wizardens begrensede evne til å manipulere GUI-elementene.
Glasskille	Nei	Lagvise GUI-elementer er vanskelig å håndtere. Begrenses til de elementene wizarden kan holde til enhver tid.
Magneter	Nei	Trenger sterke magneter. Lagvise GUI-elementer er vanskelig å håndtere.
Papp	Ja	Skaper gripeevne på fysiske elementer.
Postit-lim	Ja	Kleber for mye for bruk på museumskortene. Ble likevel brukt på mediaelementene for at de skulle sitte bedre på puslespillbrikken når den ble flyttet.
Sandpapir	Ja	Skaper ønsket friksjon.
Transparente ark	Ja	Innskrenker muligheten for direkte manipulasjon på statiske GUI-elementer.

11.3.2 Dedikerte wizards

Prototypen vil testes ved bruk av teknikken wizard-of-Oz (se Seksjon 3.5.2). Det å utspille rollen som datamaskin kan være krevende nok, og når flere enn én testbruker tester en papirprototype samtidig blir det raskt svært komplisert å holde oversikt. En måte å løse dette på er å dedikere en wizard til hver testbruker. Denne wizarden vil kun fokusere på sin testbruker og har muligheten til å forberede og spesialisere seg på de testoppgavene som hans testbruker vil gjennomgå.

11.3.3 Synkronisering

Ved bruk av parallelle oppgavesett under flerbrukertesting vil vi være nødt til å ha en strategi for synkronisering under testene. Testbrukerne vil ha noen oppgaver som skal gjennomføres alene, men samtidig på samme skjerm, og andre oppgaver der vi vil at de skal samarbeide. I MultiViten ønsker vi for øvrig at brukere skal samarbeide uten at de blir bedt om det; designet skal invitere til det av seg selv. Hvis eksempelvis de parallelle oppgavesettene består av 4 oppgaver hver, der de to første er individuelle mens de siste to inviterer til samarbeid mellom testbrukerne, må det sørges for at testbrukerne starter på oppgave 3 noenlunde samtidig.

En mulighet er aktivt å be testbruker1 om å vente dersom han er ferdig med en oppgave og testbruker2 henger etter. Wizarden til testbruker1 kan eksempelvis vente med å gi ham neste oppgave til testbruker2 er ferdig med sin.

En annen tilnærming er å la oppgavene være strukturert på en slik måte at de i seg selv klarer å synkronisere testbrukerne. Oppgavene til testbruker1 kan eksempelvis henvise til museumskort og mediaelementer som testbruker2 bruker eller har brukt til sine individuelle oppgaver og vice versa. De vil da være nødt til å koordinere seg imellom hvordan de kan komme tilbake i synkronisert tilstand igjen hvis en henger bak, ettersom den ene testbrukeren “står i veien” for den andres oppgaver.

11.4 Prototypen

Erfaringene fra eksperimenteringen med forskjellige prototypingsteknikker og -materialer bidro til å avgjøre hvordan MultiViten skulle prototypes. Bruken

av konvensjonell papirprototyping med papir og papp på oversiden av en flate hadde ulemper forbundet med seg; hovedsakelig når det gjaldt skillet mellom fysiske og virtuelle elementer. Enkelheten og de lave tids- og økonomiske kostnadene ved denne typen prototyping veide likevel mest. Dessuten ville ikke prototypen ha fullt fokus på look and feel, men også ta for seg rollen produktet vil tjene for museumsbesøkende. Den vil med andre ord befinne seg et sted mellom *rolle* og *look and feel* i Houde og Hills modell (se Seksjon 3.5.5). Vi valgte dermed å papirprototype MultiViten. Det interessante spørsmålet blir om en slik enkel prototype vil klare å formidle MultiVitens design og avdekke brukbarhetsmangler.

11.4.1 Museumskortene

Museumskortene ble laget i bildebehandlingsprogrammet Adobe Photoshop Elements 8.0¹. Kortene var på størrelse med gjennomsnittlige visittkort og inneholdt følgende informasjon om en utstillingsgjenstand: kategori, navn, et bilde og én til to korte fakta om gjenstanden. Kortets bakside bestod av en Surface-tag og teksten: “ta med dette kortet til multimediebordet og plasser det på skjermen med denne siden ned” (se Figur 11.1).

For- og baksiden ble skrevet ut i farger og limt på hver side av et 2 mm tykt pappstykke. Dette var for å gjøre museumskortet lettere å gripe og skille fra de virtuelle GUI-elementene som ble laget i arktykkelse (se Seksjon 11.2.2). 16 slike museumskort ble laget, fordelt over fire kategorier. Figur 11.15 og Figur 11.12 viser museumskort som ble laget. Tabell 11.2 viser en oversikt over hvilke gjenstander museumskortene representerte og hvilken kategori disse var knyttet til.

I konseptet MultiViten er tanken å ha flere eksemplarer av hvert unike museumskort og plassere bunker av slike i nærheten av hver utstillingsgjenstand. I prototypen vil det likevel kun være ett av hvert unike museumskort foran hver museumsgjenstand.

11.4.2 Puslespillbrikkene

Selve puslespillbrikkene ble klippet ut av farget ark. Størrelsen på puslespillbrikkene var ca. 15x15 cm (uten de utstikkende halvsirklene). De måtte være store nok til å kunne romme museumskortet og minst fire mediaelementer, men samtidig ikke bruke for mye av plassen på skjermen.

¹<http://www.adobe.com/no/products/photoshopel/>

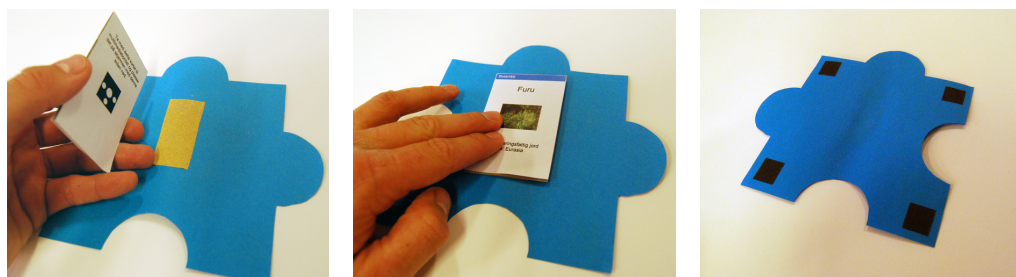


Figur 11.15: Museumskortene.

Tabell 11.2: Museumskortene som ble laget

Kategori	Gjenstand
Geologi	Granitt
	Marmor
	Amfibolitt
Botanikk	Gran
	Furu
	Eik
Biologi	Isbjørn
	Sjiraff
Fridtjof Nansen	Nansenpasset
	Boka "På ski over Grønland"
	Lue
	Polvotter
	Seddel
	Norsk flagg
	Kart over Grønland
Fredsprismedalje	

For å skape friksjon mellom museumskort og puslespillbrikke brukte vi et stykke grovt sandpapir som vi tidligere hadde eksperimentert med (se Seksjon 11.2.2). Dette sandpapirstykket var ca. halve størrelsen av museumskortet og ble limt fast på midten av puslespillbrikken. Små stykker av fint sandpapir ble også limt fast i hjørnene på baksiden av puslespillbrikken, for at brikken skulle “lugge” mot underlaget hvis testbrukeren prøvde å flytte på denne direkte fremfor via museumskortet. Som nevnt i Seksjon 11.2.2 er hensikten med sandpapirstykkene (både over og under puslespillbrikken) å bidra til at testbrukerne lettere vil skille mellom det fysiske (museumskortet) og det virtuelle (puslespillbrikken under). Figur 11.16 viser en blå puslespillbrikke (for museumskort i kategorien botanikk) med sandpapir på over- og undersiden.

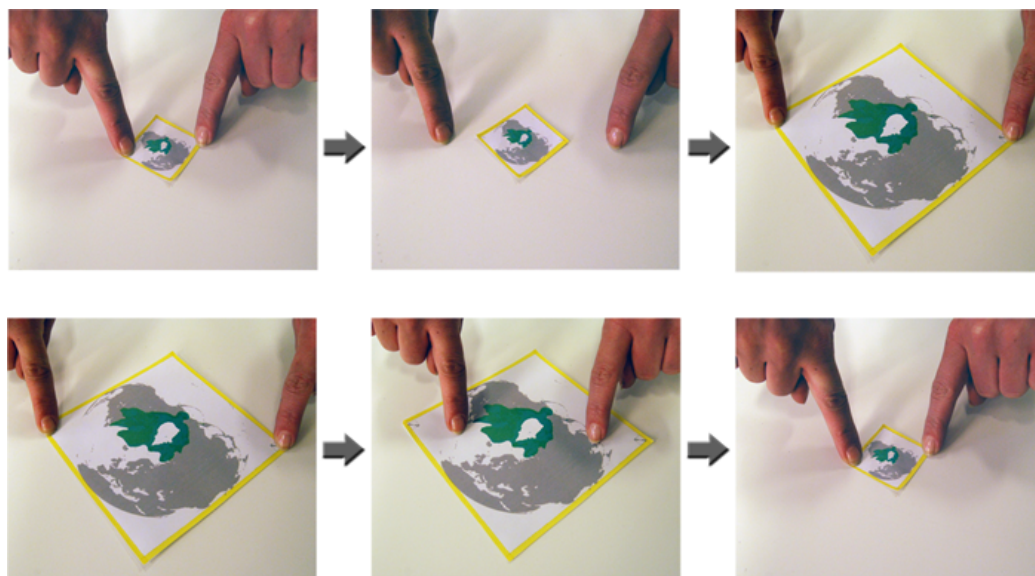


Figur 11.16: Puslespillbrikke med sandpapir for å skape friksjon mellom museumskort og underlag.

11.4.3 mediaelementene

mediaelementene som var bilder ble laget ved å skrive ut ekte bilder i liten versjon (ca. 2x3 cm) og stor versjon (ca. 12x15 cm). Bildene fikk også et ikon i to av hjørnene som indikerte at bildene kunne skaleres. Hvis en testbruker tok på det lille bildet med to fingre og dro fingrene fra hverandre, ville bildet bli byttet ut med den store versjonen. Det motsatte ville skje hvis to fingre beveget seg mot hverandre på det store bildet. Figur 11.17 illustrerer et eksempel der et bildeelement med kartmotiv blir skalert opp og ned.

mediaelementene som var videoklipp ble utviklet på samme måte som bildene, foruten et “play”-ikon midt på bildet som indikerte at elementet kunne trykkes på for å spille av en video. Hvis en testbruker trykket på dette elementet under testen ville det blitt byttet ut med en ny lapp som viste et annet, men lignende bilde og en tidslinje med halve tiden passert. Etter et par sekunder ville et siste bilde erstatte det forrige. Dette bildet ville



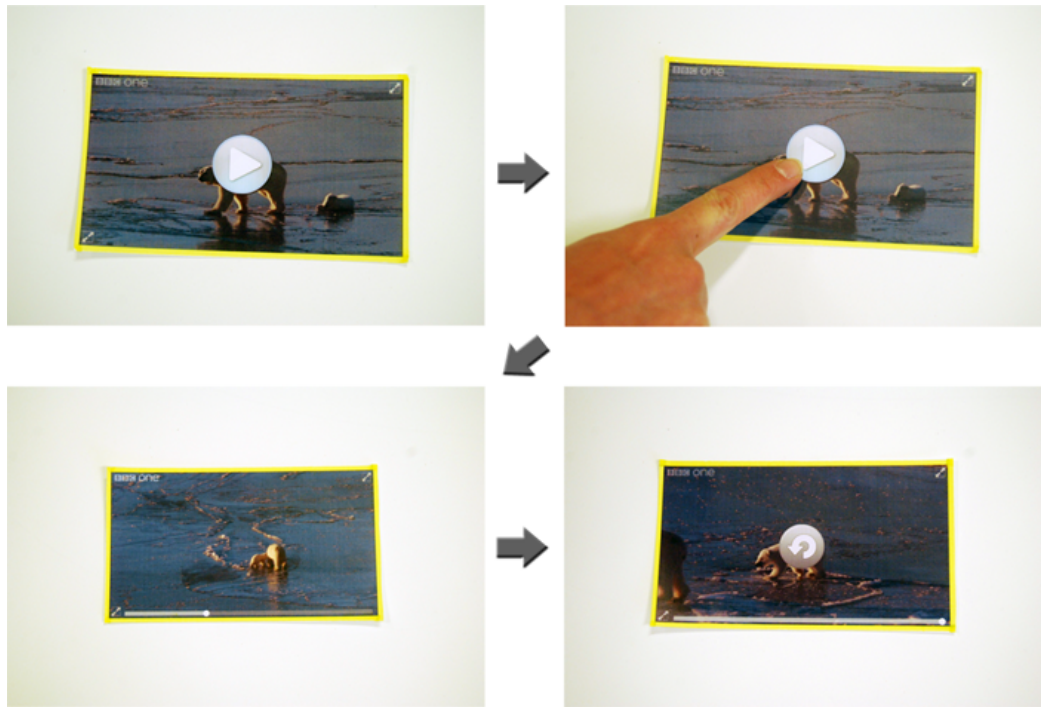
Figur 11.17: Skalering av et bildeelement.

vise et tilsvarende motiv som de to andre, med unntak av et ikon midt på som symboliserer “replay”. Dersom testbrukeren trykker på videoelementet igjen på dette tidspunktet vil klippet “spilles av” på nytt. Videoelementet kan også skaleres på samme vis som et bildeelement. Det ble derfor laget et sett med slike faser av videoklippet i stor versjon også. Figur 11.18 viser et videoelement som blir spilt av.

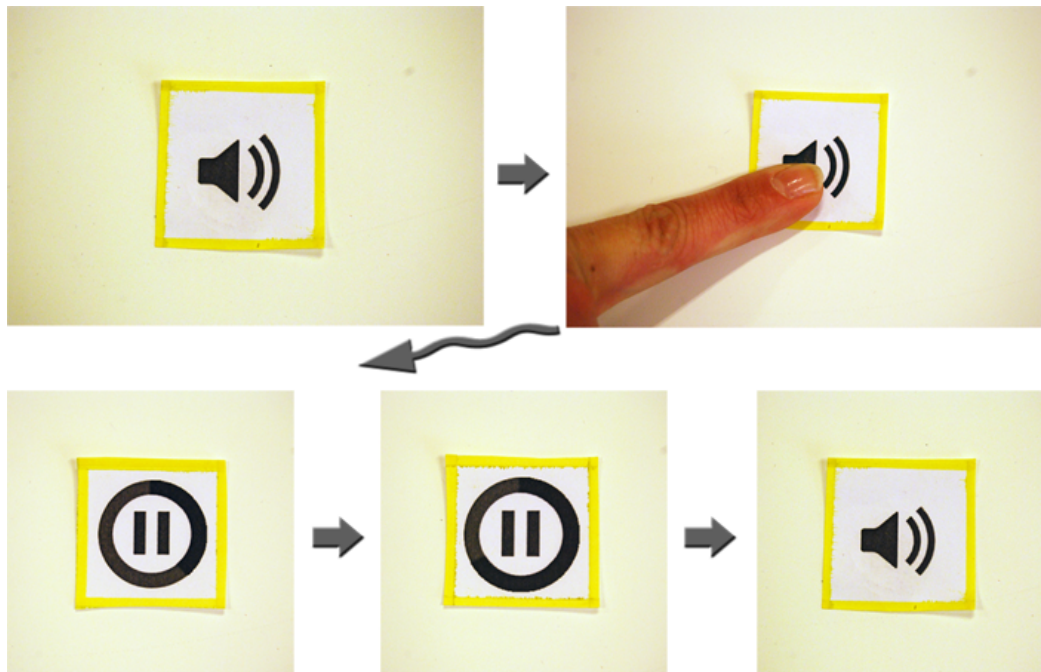
Det siste mediaelementet i MultiViten er lydklipp. Til å representere et lydklipp ble det skrevet ut en papirlapp med et lydikon i størrelsen 3x3 cm. Dette elementet skulle ikke brukerne kunne skalere. Dersom en testbruker trykket på lydikonet ville prototypen spille av lyd, samtidig som lydikonet ble byttet ut med et pauseikon, samt en sirkulær tidslinje rundt ikonet. Nye papirlapper med stadig større del av tidslinjen skyggelagt ville bli byttet ut med den foregående ettersom lydklippet nærmet seg slutten. Ved endt klipp ville papirlappen med lydklippet bli lagt på igjen. Figur 11.19 viser et lydelement som blir spilt av.

For alle museumskortene som ville være en del av scenarioene som testbrukerne skulle gjennom ble det laget forskjellige relaterte mediaelementer. Disse mediaelementene fikk en ramme i samme farge som kategorien til museumsgjenstanden de hørte til. Figur 11.20 viser to puslespillbrikker med mediaelementer på.

Vi erfarte at mediaelementene lett kunne falle av puslespillbrikkene når disse



Figur 11.18: Avspilling av et videoelement.



Figur 11.19: Avspilling av et lydelement.



Figur 11.20: To puslespillbrikker med mediaelementer.

ble flyttet av og på mockupen. Et tynt lag med postit-lim ble derfor påført undersiden av mediaelementene. Limet var nok til at mediaelementene satt fast på puslespillbrikken når den ble flyttet, samtidig som elementene kunne skyves vekk fra brikken av testbrukeren.

11.4.4 Faktaboblene

Når museums kort som ligger på skjermen skyves mot hverandre og de underliggende puslespillbrikkene kobles sammen, vil alle mediaelementene som tilhører disse forsvinne og erstattes av en faktaboble (se Seksjon 11.1.2). Figur 11.21 illustrerer et eksempel på dette. To slike faktabobler ble klippet ut av vanlig, hvitt ark. Faktasetningene ble forfattet på datamaskin, skrevet ut på hvitt ark, klippet ut og limt fast på fremsiden av postit-lapper. På denne måten kunne faktasetningene festes på og fjernes fra faktaboblene under testene.



Figur 11.21: To puslespillbrikker føres sammen og mediaelementene byttes ut med felles fakta om de to.

11.4.5 Bluetooth-ikonene

Ikonene for bluetooth-overføringen ble laget på datamaskin og skrevet ut i farger på vanlig ark. Ikonene som skulle representere mobiltelefoner med påskrudd bluetooth innenfor rekkevidde hadde et telefonsymbol, samt ledig plass under ikonet. Før testen ville testbrukerens navn bli skrevet på en liten lapp og limt fast på denne ledige plassen (eksempelvis “Fredriks mobil”, se Figur 11.22).



Figur 11.22: Ikon for å representere en mobiltelefon innen rekkevidde.

11.4.6 Mobiltelefoner

Det ble laget to mockups av mobiltelefoner som testbrukere ville få utdelt under testen (se Figur 11.23). Bilder av fasaden til to forskjellige smarttelefoner i proporsjonen 1:1 ble skrevet ut i farger og limt på pappstykker i tykkelsen 1 cm. Skjermbilder til disse papptelefonene ble også laget i Adobe Photoshop Elements og skrevet ut i farger. Postit-lim ble smurt på baksiden av disse så de enkelt kunne festes på papptelefonene under testene.

11.4.7 Surface-bordet

For å gjøre testene så realistiske som mulig ønsket vi å utvikle en mockup av et Surface-bord som prototypen skulle bli testet på. Surface har en relativt



Figur 11.23: Mockups av to mobiltelefoner.

enkel design og består hovedsakelig av en skjerm og en kropp.

Skjermen

Planen var i utgangspunktet å bruke en glassplate som Surface-skjerm. Tanken var at glasset ville minne mye om en Surface-skjerm. Det viste seg imidlertid at friksjonen mellom papiret og glassplaten var alt for høy til at GUI-elementene kunne skyves rundt på skjermen. I stedet brukte vi baksiden av kalibreringsplaten som hører til den virkelige Surface (se Figur 11.24). Denne platen er hvit og ca. like stor som en Surface-skjerm. GUI-elementene gled lett på denne platen, samtidig som sandpapiret under puslespillbrikkene fikk godt tak.



Figur 11.24: Surface-mockup.

Kroppen

Til Surface-kroppen brukte vi to pappesker satt inntil hverandre, som sammen ble på størrelse med kroppen til den virkelige Surface. Pappeskene ble tapetsert med en svart avfallspose for å ligne mer på Surface. Skjermplaten ble plassert på toppen av pappeskene.

En oppbevaringsboks ble brukt til å representere oppsamlingsboksen som brukerne kan legge brukte museumskort oppi. Denne ble festet til siden av pappeskene. Over matboksen ble det festet en lapp med teksten “legg brukte museumskort i denne boksen”. Figur 11.25 viser det endelige utseendet til mockupen.



Figur 11.25: Surface-mockup med oppsamlingsboks for brukte kort.

Lyd

Surface har integrerte høyttalere og dermed muligheten til å spille av lyd. MultiViten inneholder lydelementer for noen av utstillingsgjenstandene. Vi ønsket at lydklipp som ble spilt av i MultiViten under testene skulle komme fra selve Surface-mockupen og ikke eksterne lydkilder på utsiden. For å få til dette plasserte vi PC-høyttalere på innsiden av den ene pappesken (se Figur 11.26). Høyttalerne får input fra en 3,5 mm TRS-kabel (“mini-jack”) som ble ført gjennom et hull i esken til utsiden. Lydklipp ble på forhånd spilt

inn på en mobiltelefon (Apple iPhone 4) og denne ble koblet til høyttalerne fra utsiden. En iPhone-app kalt “Tap DJ”² ble benyttet til avspilling av lydklippene. Appen tilrettelegger for avspilling av flere lydklipp samtidig.



Figur 11.26: Høyttalere på innsiden av Surface-mockupen.

Statisk GUI

I MultiViten vil det hele tiden være ei liste langs den ene kortsiden av skjermen der ikoner for bluetooth-enheter innen rekkevidde dukker opp (se Seksjon 11.1.2). For å indikere at denne kanten av skjermen var ei liste ble et langt stykke grått papir teipet fast her, sammen med et bluetooth-symbol. Over denne lista ble et like langt stykke transparent ark festet med teip i den ene siden, slik at det var mulig å løfte opp transparenten fra den andre siden. Når en testbruker aktiverte bluetooth på sin mobiltelefon-mockup løftet wizarden opp transparenten og skjøv telefonikonet for testbrukerens mobiltelefon under den. På denne måten blir testbrukeren hindret til å flytte på eller rotere ikonet, og forhåpentligvis vil han forstå at andre flyttbare mediaelementer må flyttes over telefonikonet i stedet. Figur 11.14 i Seksjon 11.2.2 viser bruken av transparente ark i mockupen.

Meldingen som skal stå på bordet når MultiViten ikke er bruk (se Seksjon 11.1.2) ble skrevet for hånd på et A3-ark (se Figur 11.27). I det samme et museumskort legges ned på skjermen forsvinner dette arket, og settes på igjen når siste kort fjernes.

²<http://itunes.apple.com/us/app/tap-dj-mix-scratch-your-music/id405088414?mt=8>



Figur 11.27: Tekst som vises på skjermen når MultiViten ikke er i bruk.

11.5 Brukbarhetstesten

Parallelt med prototypingsarbeidet av MultiViten ble brukbarhetstesting forberedt. Ved å gjøre disse forberedelsene parallelt kunne vi vite hvilke GUI-elementer som måtte lages og hva som kunne utelates i prototypen basert på testoppgavene. DECIDE-rammeverket ble lagt til grunn ved forberedelsene til brukbarhetstesten (se Seksjon 3.6.2).

11.5.1 Parvis testing

Vi avgjorde i samråd med veileder at det ville være tilstrekkelig å teste *to* testdeltakere samtidig for å kartlegge utfordringene rundt flerbrukertesting. Vi valgte også å dedikere én wizard til hver av testpersonene (se Seksjon 11.3.2). Disse to wizardrollene var det vi selv som utgjorde. På den måten ville vi som forskere selv ta del i forsøkene. Erfaringene vi høstet ville derfor utgjøre en viktig del av resultatene.

11.5.2 Testoppgaver

For å kunne teste prototypens brukbarhet i tilfellene der brukere jobber alene og sammen med andre falt valget på å bruke to parallelle oppgavesett, fremfor et felles (se Seksjon 11.3.1). Til å synkronisere de to testpersonene under testene lot vi oppgavene selv stå for dette (se Seksjon 11.3.3). Oppgavene ble formulert på en slik måte at der vi ønsket å se testpersonene samarbeide var de nødt til å bruke GUI-elementer fra hverandres individuelle oppgaver.

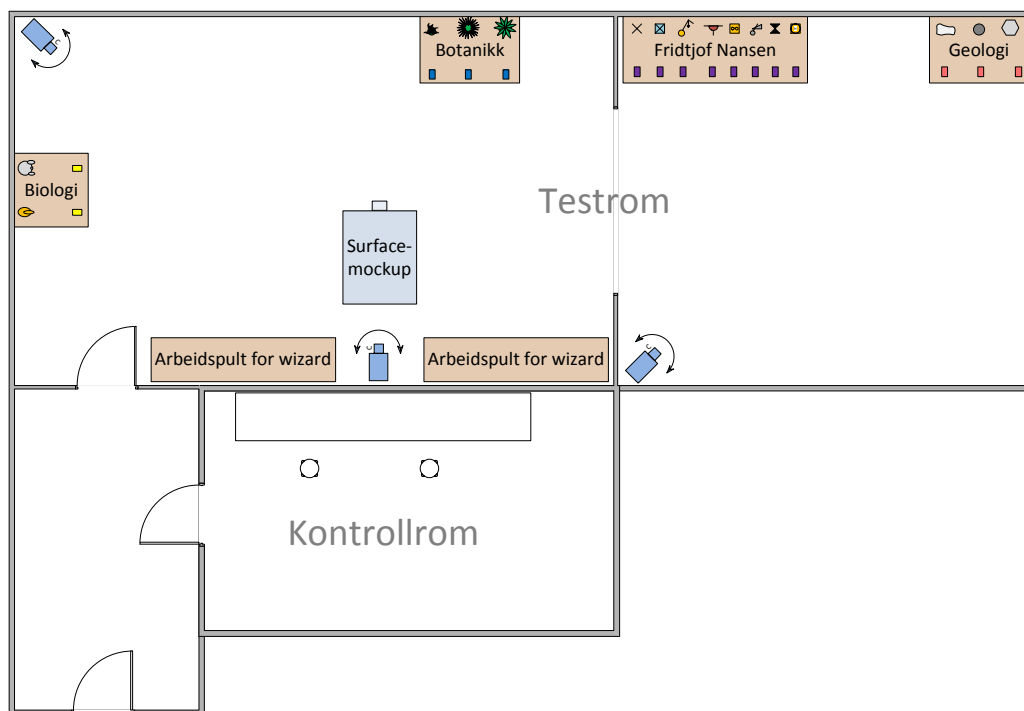
Formulering av testoppgaver til brukbarhetstester er en utfordring. Man ønsker å lede testpersonen til å bruke de delene av prototypen man ønsker å teste uten å gi for mye hjelp. I boka *“Praktisk brukertesting”* [70] skiller Toftøy-Andersen og Wold mellom tre typer oppgaver; konkrete, delvis åpne og helt åpne. Konkrete oppgaver (som er de mest brukte) forklarer hva brukeren skal gjøre, startpunktet for oppgaven og hva som er “riktig svar”. Delvis åpne oppgaver gir brukerne friere tøyler og egner seg godt når det er flere innganger til informasjonen, mens helt åpne oppgaver kan være stilt på måten “vis meg hva du ville gjort når...” og brukes for å lære om brukernes vaner og tankesett. Brukergrensesnittet til MultiViten har et “utforskende” preg og krever at brukerne prøver seg frem. For at testbrukerne skal ha noe konkret å prøve seg frem til må de ha klare mål; oppgaver som kan defineres som utført eller ikke. Samtidig ønsket vi at testbrukerne skulle ha frie tøyler,

slik de ville hatt i en realistisk situasjon på museet. Oppgavene fikk derfor en blanding av konkrete og delvis åpne formuleringer. Det ble lagt særlig vekt på ikke å veilede testpersonene for mye med oppaveformuleringene. Der vi eksempelvis ønsket at testpersonene skulle føre sammen hverandres museumskort formulerte vi oppgaven dit hen at hver testperson skulle se om sitt museumskort hadde noe til felles med andre kort på bordet.

Tabell 11.3 viser oppgavene som ble utformet. Seks oppgaver ble laget for hver testperson. Disse oppgavene skulle dekke all funksjonaliteten til prototypen. Den siste kolonnen viser hva testpersonene må gjøre for å løse oppgavene (hensikten bak oppgaven).

11.5.3 Fullskala test på lab

Til utførelse av brukbarhetstestene hadde vi brukbarhetslaboratoriet på NSEP til rådighet (se Seksjon 10.5). Ettersom vi ønsket å gjennomføre en fullskala test på laboratoriet, innredet vi testrommet om til et fiktivt museum. Figur 11.28 viser en planskisse av innredningen.



Figur 11.28: Planskisse av innredningen av testrommet.

De fire forskjellige utstillingskategoriene vi hadde utredet under utviklingen

Tabell 11.3: Testoppgavene

Oppgave	Testperson 1	Testperson 2	Hva personene må gjøre for å løse oppgaven
1	Besøk utstillingen om Fridtjof Nansen. Finn medaljen som er utstilt der. Finn ut hva det står på baksiden av medaljen.	Besøk isbjørnen som er utstilt på museet. Forsøk å finne ut hvor i verden man finner isbjørner.	Finne museumskortet til gjenstanden, legg kortet ned på Surface-skjermen og forstørre et av bildeelementene.
2	Hvis det er flere personer ved bordet med museumskort: Finn ut om Nansens medalje har noe til felles med andre museumskort på bordet. Hvis det ikke er andre ved bordet, vent og se om det kommer noen.	Hvis det er flere personer ved bordet med museumskort: Finn ut om isbjørnen har noe til felles med andre museumskort på bordet. Hvis det ikke er andre ved bordet, vent og se om det kommer noen.	Personene må føre museumskortene sine sammen så de underliggende puslespillbrikkene kobles. Snakkeboblen som da oppstår må leses.
3	Finn ut årsaken til hvorfor Nansen fikk medaljen sin.	Finn ut hvor lang en isbjørn normalt er.	Museumskortene må skilles fra hverandre igjen og lydelementet på puslespillbrikken må trykkes på.
4	Overfør historien om Nansens medalje til mobiltelefonen din.	Overfør kartet over isbjørnens leveområder til mobiltelefonen din.	Bluetooth må aktiveres på mobiltelefonen. mediaelementet må dras over telefonikonet som dukker opp.
5	Besøk avdelingen for Geologi. Velg to bergarter som du vil undersøke nøyere og finn all informasjonen du kan om disse. Har de noe til felles?	Gå til Botanikkavdelingen og velg to tresorter som du vil undersøke nøyere og finn all informasjonen du kan om disse. Har de noe til felles?	To museumskort må hentes og legges på Surface-skjermen. Kortene må føres sammen slik at personen kan lese snakkeboblen som oppstår.
6	Finn ut om geologi og botanikk har noe til felles.	Finn ut om geologi og botanikk har noe til felles.	Et eller flere kort fra de to kategoriene må føres sammen og snakkeboblen må leses.

av museumskortene (se Seksjon 11.4.1) ble viet hver sin avdeling. Hver avdeling hadde et bord der utstillingsgjenstandene var plassert (se Figur 11.29). Utstillingsgjenstandene var artefakter vi oppdrev som skulle ligne mest mulig på de virkelige gjenstandene de representerte. På ytterkanten av bordet, vendt mot publikum, plasserte vi museumskortene. Over hver avdeling festet vi et papirbanner med navnet på avdelingen, slik at det ville være lettere for testpersonene å skille avdelingene fra hverandre.

Surface-mockupen ble plassert i det største området av testrommet, med god plass rundt bordet slik at både testpersonene og wizardene hadde mulighet til å stille seg der de ønsket (se Figur 11.30). Mockupen var dessuten innen umiddelbar synsvidde for et av kameraene. Dette kameraet ble dermed brukt til å fange opp hva som skjedde på og rundt mockup-skjermen. For at mockupen skulle ha en behagelig arbeidshøyde ble den satt på et bord.

Wizardene hadde hver sin arbeidspult som var organisert med GUI-elementer de sannsynligvis ville få bruk for (se Figur 11.31). Disse bordene var plassert nære mockupen, inntil veggen vendt vekk fra publikumsområdet. På bordene var GUI-elementene sortert i båser som inneholdt ferdig arrangerte puslespillbrikker (brikkene med mediaelementer på) i tillegg til de oppskalerte versjonene av mediaelementene. Øvrige GUI-elementer ble organisert utenfor disse båsene.

I tillegg til kameraet som filmet skjermen, hadde testrommet ytterligere to kameraer. Disse ble vendt slik at de fanget opp mest mulig av de områdene de befant seg i. Det ene kameraet fanget dermed opp det som skjedde rundt mockupen, inkludert wizardenes arbeidsområde. Det andre filmet deler av museumsgjenstandene. Biologi- og botanikkavdelingen befant seg begge i en dødsone for kameraene. Vi vurderte det likevel som viktigst å gjøre videoopptak av det som skjedde på og rundt mockupen, siden dette var essensen i systemet og det var her det aller meste av interaksjonen ville foregå.

11.5.4 Oppsett

Under testene ankommer deltakerne brukbarhetslaboratoriet parvis. De vil bli ønsket velkommen og få en innføring før brukbarhetstesten starter. Etter brukbarhetstesten vil vi sammen med deltakerne diskutere hvordan de opplevde situasjonen. Underveis blir det servert kjeks og kaffe for å skape en trygg og avslappet atmosfære. Til hvert par ble det holdt av én time for hele seansen.



Figur 11.29: Avdelingene med museumsgjenstander.



Figur 11.30: Surface-mockupen ble plassert i testrommet med god avstand rundt.



Figur 11.31: Wizardenes arbeidspulter.

Innføringen

Innføringen vil foregå i kontrollrommet og være den samme for alle testparene:

1. Testdeltakerne blir ønsket velkommen.
2. Vi introduserer oss selv og masteroppgaven vår.
3. Vi forklarer hva som skal skje i løpet av sesjonen.
4. Deltakerne blir bedt om å lese gjennom og signere en samtykkeerklæring (se Tillegg C.1).
5. Det blir forklart hva Surface er.
6. En demovideo av Surface blir vist deltakerne. Dette vil bli gjort for å gi deltakerne et tydeligere bilde på hva Surface er og for å gi brukerne et så likt utgangspunkt som mulig dersom noen stiller med mer forkunnskap enn andre. Det var i utgangspunktet meningen å spille inn en video selv, ved å bruke formidlingsdemoene på instituttets Surface-bord. Surface-bordet viste seg for øvrig å være i ustand. Det brukes derfor en kort promoteringsvideo [60] produsert av Microsoft. Videoen viser forskjellige eksempler på bruk av Surface og inkluderer direkte manipulasjon, bruk av fysiske objekter på skjermen, overføring til mobiltelefon og samhandling.
7. Deltakerne får en innføring i brukbarhetstesting. Her bruker vi Bruce Tognazzinis punkter for hvordan man bør gjennomføre brukbarhetstester som grunnlag [71]:
 - (a) Hensikten med testen blir forklart
 - (b) Deltakerne blir fortalt at de kan avbryte når de vil
 - (c) Utstyret i rommet og begrensningene ved prototypen blir beskrevet
 - (d) Deltakerne læres opp til å tenke høyt
 - (e) Deltakerne forklares at det ikke vil bli tilbudt hjelp under testen, men at de kan samarbeide med hverandre
 - (f) Oppgavene blir beskrevet og produktet introdusert
8. Brukerne får utdelt hver sin mobiltelefon-mockup (se Seksjon 11.4.6) og blir forklart at det kan hende at de vil få bruk for den.

9. Deltakerne får vite at vi vil opptre som “datamaskiner” (wizarder) og at hver av deltakerne vil få sin wizard som følger etter ham/henne. De blir forklart at når wizarden sier “frys” betyr det at deltakeren har gjort en handling som gjør at prototypen skal skifte skjermbilde. Deltakeren som har fått “frys” må da vente med å fortsette til wizarden har byttet skjermbilde og sier ordet “start”.
10. Deltakerne får mulighet til å stille spørsmål i forkant av testen.

Testen

Etter innføringen får deltakerne utdelt første oppgave på et A5-ark og blir gitt tid til å lese gjennom denne. De blir så geleidet fra kontrollrommet til testrommet der testen starter. Under testen følger vi som wizarder etter vår respektive deltaker og responderer på handlinger han/hun gjør på prototypen. Når vi anser en oppgave som løst bytter vi ut oppgaven med en ny.

Avslutningen

Etter testen setter vi oss sammen med deltakerne i testrommet. Sammen med deltakerne blir det gjennomført gruppeintervju om hvordan de oppfattet MultiViten med tanke på brukbarhet. De svarer også på et spørreskjema angående brukbarheten (se Tillegg C.2).

Etter å ha diskutert brukbarheten ved MultiViten vil diskusjonen gå over på selve testopplegget. Det vil da fokuseres på hvorvidt prototypen klarte å fremstå som en applikasjon for Surface og hvordan de opplevde å gjennomføre testen med en annen. De vil også her svare på et spørreskjema om testopplegget (se Tillegg C.4).

Avslutningsvis vil deltakerne bli takket for oppmøtet og premiert med kinogavekort og USB-minnepinne.

Kapittel 12

Gjennomføring av Brukbarhetstest

Dette kapitlet tar for seg gjennomføringen av brukbarhetstestene og hvilke brukbarhetsfeil som ble avdekket.

12.1 Testforberedelser

Basert på Nielsens [48] argumentasjon for at det er tilstrekkelig å teste med fem personer (se Seksjon 3.6.1) ønsket vi å gjennomføre fem tester med to deltakere i hver. Dette ville forhåpentligvis gi oss gode nok testresultater både hva gjelder brukbarheten til MultiViten og selve forskningen. Vi rekrutterte derfor ti testbrukere, hvorav to var jenter og åtte gutter, som alle studerte informatikk i første eller andre årskurs. Ingen av deltakerne hadde tidligere erfaringer med et multi-touchbord og få hadde hørt om Surface på forhånd. Totalt fem brukbarhetstester ble kjørt, med to testbrukere på hver test.

Etter at vi hadde satt opp alt utstyret på laboratoriet gjennomførte vi en pilottest med to medstudenter. Dette gjorde vi for å avdekke eventuelle mangler ved prototypen eller testopplegget. Pilotbrukerne ga oss tilbakemelding på at oppgavene var godt formulert og vi så ingen behov for å endre på verken disse eller opplegget.

I forkant av brukbarhetstesten hadde vi laget et detaljert oppsett på hvordan vi skulle gjennomføre brukbarhetstesten (se Seksjon 11.5.4). Dette var et oppsett det var viktig å følge slavisk for samtlige av testene, slik at vi fikk et best mulig sammenligningsgrunnlag. Under gjennomføringen ble dette

oppsettet brukt på alle testparene.

12.2 Testgjennomføring

Til tross for at det ikke forekom noen store forskjeller mellom de ulike testparene var det likevel en del ting som kom mer til syne hos enkelte av de. De følgende underseksjonene beskriver sentrale hendelser under de ulike testparene. Disse beskrivelsene er en kort oppsummering av forløpet og har til hensikt å vise gangen i testingen. Rikere situasjonsbeskrivelser og sitater vil bli gitt i gjennomgangen av resultatene (Kapittel 13).

12.2.1 Testpar 1 - Lasse og Heidi

Lasse og Heidi fant fort ut av at de skulle finne museumskort og legge disse på bordet for å løse den første oppgaven. Siden Lasse kom først frem til bordet med kortet sitt ble Heidi stående og vente på at han gjorde seg ferdig. Mens Heidi holdt på med sin oppgave stilte Lasse seg ved siden av bordet og ventet på at hun skulle fullføre sin første oppgave og forsøkte å gi input til oppgaven hennes.

De hadde en tett dialog under testen og forsøkte å løse oppgavene i fellesskap. Når Heidi skulle løse første oppgave løftet hun opp et mediaelement fra bordet for å kikke nærmere på det. Her fikk hun ingen feedback som tilsa at handlingen ikke lot seg utføre. Begge skjønnte fort at de ulike brikkene kunne føres sammen med andre brikker og løste derfor raskt denne oppgaven.

Også når de skulle spille av lyd ventet de på tur med å spille av sine lydklipp, før de diskuterte sine siste oppgaver i fellesskap. Her var Heidi raskt ute med å forstå at de måtte aktivere bluetooth på mobiltelefonen og Lasse fulgte raskt etter da han så hva Heidi gjorde. De forstod ikke like lett hvordan mediaelementet kunne overføres til mobiltelefonen. Begge forsøkte først å plassere mobiltelefonen sin på skjermen og føre elementene direkte inn på denne. Lasse forstod likevel omsider hvordan det skulle gjøres og førte mediaelementet over bluetoothikonet som hadde dukket opp på skjermen. Heidi så at dette var rette måten å gjøre det på og gjorde deretter det samme.



Figur 12.1: Testpar 1 - Lasse og Heidi.

12.2.2 Testpar 2 - Christian og Torbjørn

Christian og Torbjørn skled fort ut av synkronisering ettersom Christian løste første oppgave lenge før Torbjørn. Dette førte til at Christian ikke kom seg videre før Torbjørn hadde fått fullført sin oppgave 1. Torbjørn skjønnte omsider hvordan første oppgave skulle løses, og fikk samme oppgave som Christian var på. Under oppgave 2 uttrykte Torbjørn at han egentlig ikke så noen link mellom fredsmedaljen og isbjørnen. Christian sa da til Torbjørn: *“Men hvordan kan vi finne noe felles ved å gjøre noe her?”*. *“Det er jo puslespillbrikker da”*, svarte Torbjørn. Christian syntes dette var et godt poeng og foreslo at de kunne prøve å koble de sammen. Dermed løste de oppgave 2 i fellesskap og var synkroniserte igjen.

Da Christian spilte av et lydklipp ble Torbjørn stående og vente for å høre hva som ble sagt. Han så at også han hadde i oppgave å spille av et lydklipp, men sa i stedet *“da venter vi”*, for ikke å avbryte Christians avspilling. Da Torbjørn spilte av sitt lydklipp etterpå fortsatte Christian med sin oppgave, og det ble vanskeligere å høre hva som ble sagt på lydklippet.

I oppgaven der de skulle overføre et element til mobiltelefonen via bluetooth la Torbjørn ned mobiltelefonen-mockupen sin på bordet for så å dra elementet inntil den. *“Det var sånn de gjorde på videoen”*, la han til. Han la ikke merke til at Christian alt hadde klart å overføre elementer via bluetooth, men fant etter hvert ut hvordan dette fungerte.

Da det hadde hopet seg opp flere museumskort med puslespillbrikken på bordet, spurte Christian Torbjørn om det fantes noen måte å fjerne de brikkene man ikke ville ha på bordet lenger. Etter først å ha forsøkt med å trykke på museumskortet, løftet han det og oppdaget at brikken forsvant. Torbjørn påpekte at det fantes en boks med teksten *“vennligst legg kortene her”* og Christian valgte å legge det der da han var ferdig med det.



Figur 12.2: Testpar 2 - Christian og Torbjørn.

12.2.3 Testpar 3 - Helene og Eivind

Med en gang Helene og Eivind ankom museet gikk de hvert til sitt og fant kortene som var beskrevet i oppgaveteksten. Deretter la de kortene ned på bordet omtrent samtidig. De prøvde umiddelbart å trykke på elementene som

dukket opp på brikken, før Helene forsøkte å dra et element ut fra brikken. Da de kom til neste oppgave var Eivind raskt ute med å ta tak i sin egen brikke for så å føre den sammen med Helenes og løste da både sin egen og Helenes oppgave.

Helene valgte å spille av et lydklipp om Fridtjof Nansen. Eivind ble da stående og vente til lydklippet var ferdig spilt, og satte så på lydklippet som var knyttet til sin brikke.

Begge la mobiltelefonene sine ned på bordet og prøvde å føre bildeelementer direkte over på disse. De fant deretter ut at de måtte aktivere bluetooth på mobiltelefonene og merket seg ikonene som dukket opp i listen på siden av skjermen. Eivind skjønnte da at man kunne dra elementet inn på mobiltelefonen via dette ikonet, og Helene så hvordan Eivind gjorde dette.

“Hvis jeg vil finne ut om de har noe til felles, så gjør jeg bare sånn...”, sa Helene mens hun førte sine to brikker sammen. Eivind ble stående og nøle en stund før han gjorde det samme som han så at Helene gjorde.

På siste oppgave dro Helene to puslespillbrikker fra hverandre etter at de hadde vært sammenlignet. Hun fikk dermed “frys” av wizarden sin og ventet på feedback. I mellomtiden førte Eivind to puslespillbrikker mot en av disse som Helene hadde separert. Denne handlingen endret på feedbacken som Helene ventet på.

12.2.4 Testpar 4 - Karl og Øyvind

Øyvind fant fort frem til sitt kort og var i gang på bordet lenge før Karl fant ut hvordan han skulle løse oppgaven sin. Han forsøkte å spørre Øyvind om tips til hvordan han kunne løse oppgaven sin, men Øyvind virket for interessert i sin egen oppgave til å hjelpe Karl. Karl stilte seg etter hvert opp ved siden av bordet og observerte Øyvind. Dette fikk ham inn på tanken å hente isbjørnkortet og legge det på bordet.

Øyvind dirigerte Karl til å slå på bluetoothen til tross for at denne oppgaven ikke kunne løses på den måten. Grunnet til at han foreslo dette var at Øyvind selv hadde oppdaget denne funksjonaliteten og var i gang med å utforske den. Øyvind prøvde også å trykke på ulike bildeelementer fremfor å skalere dem opp. Etter mye forvirring fant de omsider fram til riktig måte å løse sine første oppgaver på.

På oppgave 2 trykket Øyvind på det fysiske museums kortet for å se om noe skjedde før han bestemte seg for å ta tak i de to brikkene som lå på bordet og



Figur 12.3: Testpar 3 - Helene og Eivind.

føre de sammen. Han løste dermed oppgave 2 for både seg selv og Karl, mens Karl sto passivt på siden og kikket på Øyvind som interagererte med bordet.

På oppgave 3 ville Karl overføre et element til mobiltelefonen og forsøkte da å legge elementet oppå den fysiske mobiltelefonen. Øyvind ble inspirert av Karl og la da hele museumskortet oppå den fysiske mobiltelefonen og forventet at noe skulle skje. Begge forsøkte ulike måter som å legge kortet oppå bluetoothikonet, dra elementene direkte inn på den fysiske mobiltelefonen, samt dra kortet inn på mobiltelefonen. Til slutt fikk Øyvind ideen om å dra et element over bluetoothikonet og oppgaven løste seg. “Jeg fant det ut”, sa han til Karl og hjalp ham slik at også han ble ferdig med sin oppgave. Karl valgte så legge isbjørnkortet tilbake der han fant det for å frigjøre plass på bordet.

Godt i gang med nest siste oppgave skulle Øyvind finne en fellesnevner mellom sine to brikker og mente å huske at dette var noe han ville finne ved å legge de to fysiske kortene oppå hverandre. Karl som var fokusert på sin egen oppgave fant ut hva som var felles med sine to kort ved å føre puslespillbrikkene sammen. Øyvind henvendte seg da til Karl: *“hvordan fant man*

ut hva som var felles?”, og fikk svar som førte ham frem til riktig løsning.

I siste oppgave lånte de brikker fra hverandre for å løse sin siste oppgave hver for seg, i stedet for å samarbeide.



Figur 12.4: Testpar 4 - Karl og Øyvind.

12.2.5 Testpar 5 - Jørgen og Frank

Både Jørgen og Frank var raske med å finne museumskort og legge de ned på bordet. Jørgen løste imidlertid første oppgave før Frank. Når han videre skulle løse oppgave 2 tok tak i puslespillbrikken sin og førte den inntil den puslespillbrikken Frank fortsatt jobbet med. Frank løste sin første oppgave rett før Jørgen koblet puslespillbrikken sammen og plutselig var også oppgave 2 løst for begge.

Da Frank spilte av lydklippet om Nansen ble Jørgen stående og vente for å høre på. Jørgen spilte deretter av sitt klipp etter Frank.

På tredje oppgave ville Frank føre lydklippet over til mobiltelefonen sin. Han flyttet lydelementet over telefonikonet som hadde dukket opp i listen

på siden av skjermen, og fikk lydklippet på mobiltelefonen sin. Etter å ha strevet mye med å løse tilsvarende oppgave, merket Jørgen seg dette og fikk løst sin oppgave på samme vis.

Da Frank gikk videre med sin siste oppgave trengte han en botanikkbrikke til å føre sammen med sin geologibrikke. Han spurte da Jørgen som fremdeles var på femte oppgave om han kunne låne en brikke av ham.



Figur 12.5: Testpar 5 - Jørgen og Frank.

12.3 Testresultater

Denne seksjonen oppsummerer brukbarhetsresultatene til applikasjonen MultiViten.

Sett fra et brukbarhetsperspektiv, var hensikten med testene å avdekke problemer relatert til brukbarheten ved MultiViten. Det ble samlet inn både kvantitativ og kvalitativ data under testsesjonene. De kvalitative dataene bestod av video- og lydopptak av oppgaveløsningen, samt intervjuene som ble gjort i

etterkant av testene. For å samle inn kvantitativ data fikk brukerne i oppgave å fylle ut et spørreskjema som tok for seg spørsmål knyttet til brukbarheten.

Ut fra spørreskjemaet, som var strukturert som en fempunkts Likert-skala, ble det regnet ut en score til hver av påstandene. Påstandene med tilhørende totalscore er presentert i Tabell 12.1. Scoren er regnet ut på følgende måte: For påstand 1, 4, 5, 6, 7, 8 og 9 vil scoren være fra 1 (sterkt uenig) til 5 (sterkt enig). For påstand nummer 2, 3, 10, 11 og 12 vil scoren være fra 5 (sterkt uenig) til 1 (sterkt enig). Totalscoren for hver påstand vil være summen av scoren fra de ti brukerne, som betyr at hver påstand maksimalt kan oppnå en score på 50. Det fremkommer av tabellen at påstand 9, 10, 11 og 12 er de som kommer ut med dårligst score. Disse er tilknyttet problemene vi har prioritert høyest med tanke på systemforbedringer.

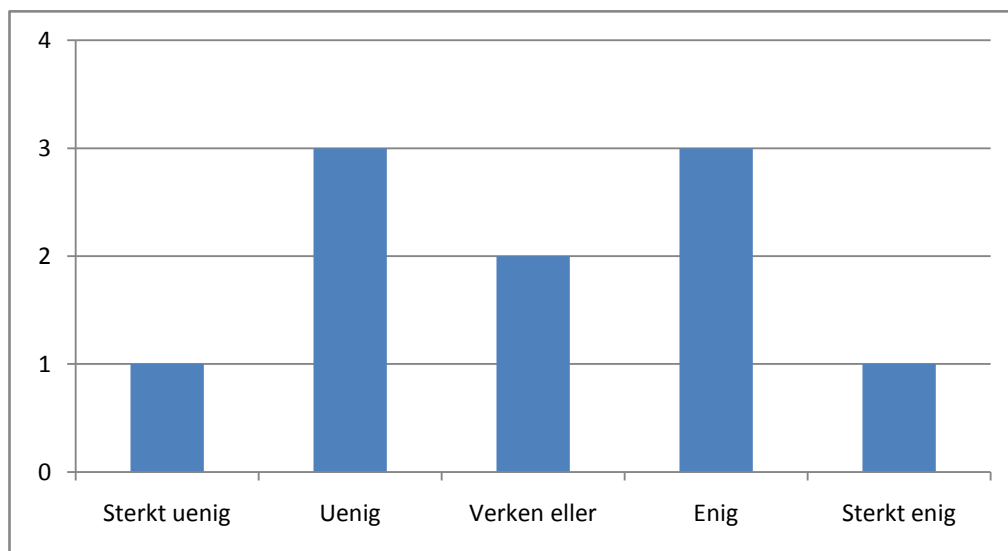
Tabell 12.1: Påstander om brukbarheten ved MultiViten med tilhørende totalscore

Påstand	Totalscore
1. Systemet var lett å bruke.	42
2. Systemet var unødvendig komplisert.	43
3. Jeg tror jeg ville hatt behov for hjelp når jeg brukte systemet.	38
4. Jeg kan forestille meg at de fleste folk ville lært dette systemet raskt.	39
5. Det var lett å forstå at museumskortene kan brukes på Surface-bordet.	42
6. Det var lett å skjønne hvilke av mediaelementene som var bilder.	44
7. Det var lett å skjønne hvilke av mediaelementene som var video.	43
8. Det var lett å skjønne hvilke av mediaelementene som var lydklipp.	48
9. Det var lett å skjønne at man kan flytte fritt rundt på mediaelementene.	39
10. Det var vanskelig å forstå at man ville lære noe mer om museumsgjenstandene dersom man førte museumskort/puslespillbrikkene sammen.	30
11. Det var vanskelig å forstå hvordan man kan overføre et element til mobiltelefonen via bluetooth.	33
12. Puslespillbrikkene tok for mye plass på skjermen.	29

Av spørreskjemaene kommer det tydelig frem at det var bred enighet om at systemet var lett å bruke. Det var ikke unødvendig komplisert og så tilsynelatende lett ut å lære. Det var på forhånd usikkerhet rundt hvorvidt testbrukerne ville forstå bruken av museumskortene. Det viste seg imidlertid at det var kun én av brukerne som hadde problemer med dette. En av grunnene til at testbrukerne forstod dette raskt kan være hjelpeteksten som stod på mockup-skjermen når det ikke var i bruk. Det kan likevel antas at det ville være vanskeligere for en tilfeldig museumsgjenger i en mer realistisk museumsetting å forstå denne bruken av museumskort.

12.3.1 Puslespillbrikkene

Selv om testbrukerne hevdet at systemet generelt var enkelt å forstå, var det likevel flere aspekter ved MultiViten som var mer utfordrende enn andre. Selv om flere av brukerne oppfattet bruken av metaforen *puslespillbrikker*, kom det ikke tydelig nok frem at man kunne lære mer ved å føre disse sammen. Selv om Tabell 12.1 viser en lav totalscore på denne påstanden, kan man i diagrammet i Figur 12.6 se at det var stor uenighet om hvorvidt det var forståelig at man ville lære mer om museumsgjenstandene ved å føre brikkene sammen. Noen forstod det altså ikke i det hele tatt, mens andre syntes det var uproblematisk.



Figur 12.6: Påstand 10.

Det var også uenighet blant brukerne når det gjaldt hvor mye puslespill-

brikkene tok av plass. Like mange var enige som uenige i at brikkene tok for mye plass på skjermen. Dette kan bero på at noen av brukerne fikk oppleve at det til tider ble liggende mange brikker samtidig på skjermen under oppgaveløsningen.

12.3.2 Overføring via bluetooth

Noe som kom svært tydelig frem under testene var at brukerne hadde store problemer med å overføre elementer til mobiltelefonen via bluetooth og dette er helt klart noe som burde forbedres. Noe av årsaken til at dette var et markant problem var at brukerne i stor grad ble påvirket av introduksjonsvideoen som ble vist i forkant av testen. Under videoen ble det demonstrert hvordan bilder kunne overføres til mobiltelefonen ved å legge den ned på bordet og dra elementer direkte inn. Dette kan ha ført til misledning. Det kan argumenteres for at det var uheldig å vise akkurat denne videoen. Dette problemet kunne antakeligvis vært avdekket tidligere ved å kjøre flere pilottester enn den éne. Den store andelen av testbrukere som prøvde på nettopp dette tilsier likevel at de følte at akkurat den måten var en intuitiv måte å løse oppgaven på. Dette bør tas til etterretning under en redesign av prototypen.

12.3.3 Trykking på bildeelementene

Under testene observerte vi at flere av testbrukerne prøvde å trykke på bildeelementene fremfor å skalere dem. Det var tydelig at de forventet en feedback på trykkingen. Dette er forståelig siden både video- og lydelementene ga feedback når de ble trykket på.

12.3.4 Flytting av mediaelementer

Det viste seg også at ikke alle forstod at mediaelementene kunne flyttes av puslespillbrikken og beveges rundt på skjermen. Dette var de blant annet nødt til når mediaelementene skulle føres over til mobiltelefonen ved å flytte elementet over telefonikonet sitt.

12.3.5 Forslag til forbedringer

Tabell 12.2 oppsummerer brukbarhetsproblemene som ble avdekket. I et realistisk prosjekt ville disse ha blitt endret på i et redesign og testet ut blant

brukere på nytt. Tabellen viser derfor forslag til redesign av funksjonene som problemene er knyttet til.

Tabell 12.2: Brukbarhetsproblemer og forslag til forbedringer

Brukbarhetsproblemer	Forslag til forbedringer
Det var ikke intuitivt nok å forstå at puslespillbrikkene kunne føres sammen.	Koblingspunktene på puslespillbrikker kan “gløde” når de kommer i nærheten av hverandre. Dette vil forhåpentligvis hinte om at de kan kobles sammen.
Puslespillbrikkene tok til tider for stor plass.	Dynamisk størrelse på puslespillbrikkene. Brukerne skal både kunne skalere denne størrelsen på puslespillbrikkene selv, samtidig som systemet automatisk skalerer de ned når det er mange på bordet samtidig.
Det var vanskelig å forstå hvordan man overfører mediaelementer til mobiltelefonen via bluetooth.	I stedet for et telefonikon som brukeren fører mediaelementer over, kan det dukke opp en firkant som brukeren skal legge den fysiske mobiltelefonen på. Når dette gjøres vil den fysiske mobiltelefonen få en virtuell “sfære” rundt seg som vil følge den fysiske mobiltelefonen hvis den flyttes rundt på skjermen. Brukeren kan så føre mediaelementer inn mot den fysiske mobiltelefonen (inni sfæren) og elementet blir overført via bluetooth.
Det oppleves forvirrende at det ikke skjer noe når man trykker på bildeelementer.	Bildeelementene trenger feedback når de trykkes på. De kan eksempelvis skales opp når de trykkes på i liten størrelse og skales ned fra stor. Evt. kan bildet snu seg rundt og gi informasjon på baksiden.
Det var vanskelig å vite at man kan flytte på medielementer.	Medielementer kan komme “flyvende” inn på puslespillbrikken eller riste litt på seg når puslespillbrikken dukker opp slik at brukerne får følelsen av at de ikke sitter fast på den.

Kapittel 13

Resultater av Brukbarhetstest

Dette kapitlet legger frem resultatene fra brukbarhetstestene som danner grunnlaget for å svare på FS2. Seksjon 13.1 tar for seg funnene som ble gjort. Disse er basert på videoopptakene, gruppeintervjuene, spørreundersøkelsen og våre egne opplevelser av testseansene. Vi vil i Seksjon 13.2 analysere disse funnene før konklusjonen på FS2 presenteres i Seksjon 13.3.

13.1 Funn

Brukbarhetstestene ga klare og interessante funn. Gjennom refleksjon av våre egne erfaringer med testseansene, samt analyse av video- og lydopptak har vi blitt oppmerksomme på utfordringer knyttet til papirprototyping og brukbarhetstesting av Surface-applikasjoner.

13.1.1 Testbrukernes påvirkning av hverandre

Noe vi merket oss under testene var at testpersonene påvirket hverandre. Denne påvirkningen kunne være i form av tydelige handlinger som dialog og støtte, men også som subtil og kanskje underbevisst oppførsel i en sosial setting. Vi har derfor valgt å skille mellom henholdsvis aktiv og passiv påvirkning mellom testbrukerne.

Aktiv påvirkning

Flere av testparene hadde mye dialog seg imellom og påvirket hverandre på den måten svært aktivt. De fungerte som støttespillere for hverandre under testene siden de visste at vi ikke kom til å tilby denne hjelpen fra vår side. Vi merket dette allerede under testingen av det første paret:

Etter tre minutter rådførte Heidi seg med Lasse, som da var kommet til neste oppgave: *“Kan jeg ta det [mediaelementet] ut og forstørre det?”*. Lasse svarte *“Mhm”*, og demonstrerte for henne hvordan han hadde gjort det.

Under de følgende testparene så vi den samme tendensen. Partene hjalp hverandre, både når den ene spurte og også uten at det ble eksplisitt bedt om hjelp. Testpar 4 var særlig aktive til å støtte hverandre:

Karl viste stor forvirring i begynnelsen av testen og skottet flere ganger på hva Øyvind foretok seg. Etter en stund spurte han Øyvind om hjelp. Øyvind var mest interessert i å gjennomføre sin egen oppgave, men når han så hvor mye Karl slet, sa han: *“Ta med deg et sånt museumskort, du”*. Etter dette ble de mer aktive til å støtte hverandre. Dette kom spesielt til syne da begge stod fast på oppgave 4 som innebar at de skulle overføre et mediaelement til telefonen sin. Etter å ha prøvd forskjellige måter å gjøre dette på hver for seg, fant Øyvind plutselig ut hvordan og delte det med Karl: *“Jeg fant det ut!”*. *“Hva gjorde du”*, spurte Karl. *“Du må dra den [mediaelementet] dit”*. Under neste oppgaven returnerte Karl tjenesten da han viste hvordan kortene kunne sammenlignes, da Øyvind hadde glemt dette fra oppgave 2.

Det viste seg at den aktive påvirkningen mellom testpersonene kunne fungere som drivkraft i testene når den ene parten var i ferd med å gi opp. Torbjørn i testpar 2 uttrykte eksempelvis frustrasjon da han ikke skjønnte hvordan oppgave 2 skulle løses:

“Jeg ser egentlig ingen link mellom isbjørn og Nansen”. Christian forsøkte da å få Torbjørn med på sin tankegang: *“Men hvordan kan vi finne noe felles ved å gjøre noe her?”*. *“De er jo puslespillbrikker da”*, svarte Torbjørn. Dermed kom de frem til at puslespillbrikkene måtte kobles sammen.

Den aktive formen for påvirkning var også til tider ødeleggende for den parten som ble påvirket. Dette skjedde særlig under oppgaver der vi ønsket at testpersonene samarbeidet uten at de gjorde det eller i tilfeller der den ene testperson hadde kommet til en slik oppgave før den andre. Et eksempel på dette var da testpar 3 skulle løse siste oppgave:

Helene dro fra hverandre de to museumskortene hun akkurat hadde slått sammen. Denne handlingen skal føre til at faktaboblen som stod der byttes ut med mediaelementene som hører til museumskortenes respektive puslespillbrikker. Helenes wizard ga henne derfor signalet “frys!” og plukket av faktaboblen. Wizarden tok med faktaboblen til arbeidspulten sin der den skulle byttes ut med mediaelementene som skulle tas med tilbake til mockupen. I mellomtiden skjøv Eivind to museumskort (de han nettopp på tilsvarende måte hadde slått sammen) mot de to Helene hadde separert. Dette endret på arbeidet som Helenes wizard var i gang med. I stedet for å bringe mediaelementene til Helenes separerte brikker tilbake, var det nå en faktaboble som skulle plasseres på skjermen. Selv om dette var positivt i den forstand at Eivind løste oppgaven sin og dermed også Helenes, var det ikke bra at handlingen Helene gjorde fikk en annen feedback enn forventet.

En lignende hendelse som den beskrevet over oppstod for testpar 5 da Jørgen hadde gjort seg ferdig med oppgave 1 før Frank:

Jørgen fikk utdelt oppgave 2 mens Frank prøvde å forstå hvilke av mediaelementene på puslespillbrikken hans som var baksiden til Nansens fredsprismedalje. Jørgen forstod raskt at for å løse oppgave 2 måtte han koble sammen puslespillbrikken han hadde brukt i oppgave 1 med brikken Jørgen holdt på med. Uten at Jørgen sa i fra om det førte han sin puslespillbrikke mot puslespillbrikken til Frank slik at disse koblet seg sammen (se Figur 13.1). Frank hadde da heldigvis akkurat funnet ut hvilket element han måtte studere og løste oppgaven sin rett før wizarden til Jørgen fjernet mediaelementene hans for å gi feedback til Jørgens handling.

Hvis Frank ikke hadde vært ferdig med sin oppgave 1 idet Jørgen koblet sammen brikkene ville dette ødelagt betraktelig for testingen av Frank på



Figur 13.1: Jørgen kobler sin puslespillbrikke mot Franks.

denne oppgaven. Denne episoden førte likevel til at Franks oppgave 2 var løst før han i det hele tatt fikk lest oppgaveteksten.

En annen form for aktiv påvirkning brukbarhetstestene avslørte var at testpersonene “lånte” museumskort av hverandre. Dette skjedde som regel i tilfeller der den ene testpersonen hadde kommet lenger i oppgavesettet enn den andre. I testpar 5 hadde eksempelvis Frank kommet til sjette og siste oppgave før Jørgen, som fortsatt jobbet med oppgave 5. Siden oppgave 6 krever bruk av museumskort som begge testpersonene holdt på med i oppgave 5, spurte Frank om han kunne låne museumskortet “Furu” av Jørgen for å sammenligne Geologi og Botanikk.

Passiv påvirkning

Den passive påvirkningen var ofte i den forstand at den ene parten la merke til hva den andre parten gjorde og dermed lærte av dette. Et godt eksempel på det kom til syne under testingen av testpar 3:

Helene og Eivind hadde kommet til oppgave 4 samtidig. Etter gjennomlesing av oppgaven sa Eivind, mens han stirret på telefon-mockupen sin: “Er det no sånn bluetooth...”. Helene skjøt da inn, etter å ha tenkt seg om litt: “Ja, du må trykke der”, mens hun trykket på bluetooth-knappen på telefon-mockupens skjerm bilde. Wizardene sa “frys” og plasserte telefonikonene i lista over bluetooth-enheter innen rekkevidde. Eivind observerte dette og prøvde umiddelbart å føre mediaelementet over telefonikonet sitt idet han fikk “start” fra wizarden sin. I mellomtiden hadde Helene lagt telefonen sin

på skjermen og var i ferd med å føre mediaelementet sitt mot selve telefonen da hun oppdaget hva Eivind hadde gjort. Hun skjøv da mediaelementet i en omvei rundt telefonen og over telefonikonet sitt i stedet.

Eksempelet viser hvordan Helene først aktivt påvirker Eivind med å foreslå hvordan bluetooth aktiveres, før hun deretter blir passivt påvirket av Eivinds handling for å føre over mediaelementet. Den samme passive påvirkningen så vi tydelig i testpar 5 da Jørgen og Frank utførte oppgave 3:

Etter at Frank hadde spilt av lydklippet sitt, tenkte Jørgen høyt:
“Ok, så hvis du kan høre på noe, så kan sikkert jeg også det på samme måte”.

Det var ikke vanskelig å se at brukerne i stor grad var oppmerksomme og tok hensyn til den andre brukeren. Under samtlige testpar ventet den ene parten på at den andre skulle spille ferdig lydklippet sitt før neste lydklipp ble spilt av. Alle antok at lydklippene ville mikses og støye for hverandre hvis begge ble avspilt samtidig, og ønsket derfor ikke å ødelegge for den andre som lyttet til klippet. De uttalte i etterkant at det var naturlig for dem å vente for ikke å skape forstyrrelser. Hensyn ble også tatt i form av plass på skjermen og elementer på denne. Elementene på skjermen ble betraktet som “mine” og “dine” eiendeler og på oppgaver som krevde samarbeid ble det ofte sett på som “låning” av hverandres brikker.

13.1.2 Skillet mellom fysiske og virtuelle elementer

Noe som kom svært godt frem under brukbarhetstestene var at samtlige av deltakerne hadde vanskeligheter med å skille mellom hva som var fysiske og hva som var virtuelle elementer på skjermen. Det var blant annet brukere som forsøkte å trykke på det fysiske museumskortet (se Figur 13.2), dro virtuelle objekter over på den fysiske mobiltelefonen, og flyttet på den virtuelle puslespillbrikken uten å flytte på museumskortet.

Et eksempel som tydelig illustrerer hvordan testpersonene slet med å skille mellom de fysiske og de virtuelle elementene er hvor testpar 4 hadde i oppgave å overføre mediaelementer til mobiltelefonene sine:

“Aha, da tar jeg den her”, sa Øyvind og tok det virtuelle lydelementet og la oppå telefonen. Da han skjønnte at dette ikke ga noen feedback fra wizarden prøvde han med ett å legge selve museumskortet på telefonen i stedet. Febrilsk koblet han mobiltelefonen



(a) Christian i testpar 2

(b) Øyvind i testpar 4

Figur 13.2: To testbrukere forsøker å trykke på fysiske museumskort.**Figur 13.3:** Øyvind legger et kort på mobiltelefon-mockupen.

og museumskortet sammen ved å legge de inntil hverandre. Heller ikke dette ga den ønskede feedbacken og han valgte en ny vri med å legge telefonen oppå kortet (se Figur 13.3). Til slutt prøvde han nok en gang å legge elementet direkte oppå den fysiske mobiltelefonen uten at dette løste oppgaven (se Seksjon 12.2.4).

Vi merket oss to tilfeller hvor testbrukerne trykket på de fysiske museumskortene og ventet at noe skulle skje. Det ene tilfellet var en hendelse fra testpar 2:

Christian var rådvill og ønsket å løse oppgaven med å finne ut hva isbjørnen og Nansen hadde til felles. Etter å ha bommet på

flere ulike interaksjonsalternativer trykket han på museumskortet for å se om det dukket opp noe mer informasjon (se Figur 13.2a). Det var tydelig at han hadde ventet seg at det skulle skje noe når han trykket på museumskortet, og dermed tolket dette som en virtuell “knapp”.

Det andre tilfellet er hentet fra testpar 4 og beskriver Øyvinds forsøk på å løse samme oppgave som Christian:

Umiddelbart etter at Øyvind hadde lest oppgavebeskrivelsen satte han i gang med å trykke på ulike elementer på puslespillbrikken han hadde foran seg (Fridtjof Nansens medalje). Han trykket først på to mediaelementer før han til slutt valgte å trykke på museumskortet. Da dette ikke fungerte prøvde han å sette de to museumskortene oppå hverandre (se Figur 13.4). Dette fungerte heller ikke, i og med at Surface ikke kan vite at de to fysiske kortene er stablet oppå hverandre.



Figur 13.4: Øyvind legger et museumskort oppå et annet.

13.1.3 Superrealisme

Det var sprikende funn når det gjaldt hvorvidt det superrealistiske (se Seksjon 4.4.1) var intuitivt for brukerne eller ikke. Fire av testpersonene brukte

gestures på GUI-elementer for å skalere de opp. Til forskjell fra dette var det to testpersoner som viste tegn på at de ikke skjønnte at bildeelementene kunne skaleres. Den ene av disse var Helene i testpar 3:

Helene skulle på oppgave 1 lese hva det stod på baksiden av Nansens medalje. Bildeelementet med motivet av denne baksiden er nokså lite. Hensikten med oppgaven er at testbrukeren skal måtte skalere opp bildet for å gjøre teksten på medaljen lettere å lese. Det Helene i stedet gjorde var å føre hodet nærmere skjermen og lese teksten på denne måten (se Figur 13.5).



Figur 13.5: Helene fører hodet nærmere bildeelementet i stedet for å skalere det opp.

Øyvind i testpar 4 og Frank i testpar 5 gjorde akkurat det samme som Helene når det gjaldt denne oppgaven. Øyvind forstod imidlertid under en senere oppgave at mediaelementene kunne skaleres.

Av alle de ti testpersonene var det seks personer som aldri skalerte noen av mediaelementene.

13.1.4 Synkronisering

Selv om oppgavene i seg selv var utarbeidet med tanke på å fremme synkronisering var det varierende hvorvidt de ulike parene var synkroniserte under testene.

Hos testpar 1 fungerte synkroniseringen tilsynelatende. Lasse ble stående og vente på at Heidi skulle bli ferdig med sin første oppgave. Når Heidi stod

fast støttet Lasse henne med tips om hvordan hun kunne løse oppgaven sin. Dette førte til at de gikk i gang med oppgavene som involverte samarbeid i fellesskap. Denne dynamikken hadde de gjennom hele testsettet, som gjorde at de til enhver tid var på samme oppgave.

Testpar 2 skled fort ut av synkronisering fordi Torbjørn brukte noe mer tid på å finne museumskortet som var beskrevet i hans oppgave. Dette førte til at Christian ikke kom ordentlig i gang med sin oppgave 2 før Torbjørn var ferdig med oppgave 1. Først når også Torbjørn kom til oppgave 2 kom de frem til den felles løsningen om å føre sammen de to brikkene. Videre ut i testen forble paret forholdsvis synkroniserte og ender opp med å løse den siste oppgaven med å koble sammen geologi- og botanikkbrikkene i fellesskap.

Testpar 3 forholdt seg synkroniserte under hele testen. De påvirket hverandre både aktivt og passivt, slik at de alltid fullførte oppgavene sine nokså samtidig. Helene startet likevel på siste oppgave før Eivind. Hun var forvirret over om hun skulle lete etter et nytt botanikk-kort eller ikke for å slå sammen med geologi-kortene hun hadde brukt. Mens hun funderte over dette hadde Eivind også kommet til siste oppgave og lest denne. Da hun hørte Helenes høyttenkning ble han oppmerksom på at de hadde like oppgaver og sa: "Hvis du har geologi og har henta dem og jeg har botanikk, så kan vi prøve å... ", mens han gestikulerte på skjermen.

Testpar 4 hadde under hele testen problemer med synkroniseringen. Det hele startet med at Øyvind var tidlig ute med å utforske bordet, mens Karl brukte lang tid før han forstod at han skulle ta med museumskortet av isbjørnen til bordet. Øyvind virket mer interessert i sin egen oppgave enn å svare på spørsmålene som Karl stilte han. Øyvind var allerede i gang med oppgave 2 før Karl i det hele tatt hadde kommet seg bort til bordet. Etter hvert ble Øyvind mer interessert i å hjelpe Karl slik at de ble synkroniserte. På nest siste oppgave var det derimot Karl som bistod med å hjelpe Øyvind. Dette resulterte i at de var synkroniserte mot slutten og løste den siste oppgaven sammen.

Testpar 5 var svært usynkroniserte. Som nevnt i Seksjon 13.1.1 løste Jørgen oppgave 2 idet Frank akkurat hadde klart oppgave 1. Dette førte til at oppgave 2 var løst for Frank uten at han selv stod for det. Selv om dette førte til at de var synkroniserte, skled de raskt fra hverandre igjen. Under avspillingen av lydklippene ble de stående og vente på hverandre. Ettersom Frank ikke hadde fått med seg alt som ble sagt valgte han å spille av sitt klipp på nytt. I mellomtiden gikk Jørgen videre til oppgave 4. På oppgave 4 skjønnte Frank imidlertid hvordan mediaelementet kunne føres over til telefonen lenge før Jørgen tok etter. Frank lå etter dette én oppgave foran Jørgen

helt til testen var ferdig. Oppgave 6 som skulle være utformet slik at de samarbeidet, unngikk han ved å “låne” en av brikkene til Jørgen.

13.1.5 Spørreundersøkelsen

Avslutningsvis i alle brukbarhetstestene ble det, i tillegg til diskusjonen, gjennomført en spørreundersøkelse (se Seksjon 11.5.4). Tabell 13.1 presenterer resultatene fra denne. I likhet med Tabell 12.1 er det regnet ut en score på hvert spørsmål som indikerer hvor positivt enig testdeltakerne var i påstandene (høy score på negativt formulerte påstander betyr at de er enige i den positive omformuleringen).

13.2 Analyse

Denne seksjonen analyserer funnene som ble gjort.

13.2.1 Testbrukernes påvirkning av hverandre

Grunnen til at det ble gjennomført tester på to testbrukere samtidig var for å skape en realistisk situasjon – nemlig flere brukere som står rundt Surface samtidig og interagerer med skjermen. En slik flerbrukertest vil dermed avdekke hvordan en potensiell Surface-applikasjon fungerer når flere bruker den og avsløre brukbarhetsproblemer som ellers ikke ville vært oppdaget hvis prototypen hadde vært testet på én og én uavhengig testbruker. Brukbarhetstestene vi gjennomførte har likevel avslørt at testing av to personer samtidig er langt mer komplisert enn individuelle tester. Hovedårsaken til dette er at de to testpersonene – variabler vi ikke kan og ikke skal kontrollere – ofte påvirker hverandre i så stor grad at potensielle brukbarhetsfeil i verste fall aldri blir oppdaget under testing. Som nevnt merket vi oss at flere av testparene hadde en tett dialog gjennom testene og var til stor hjelp for hverandre under oppgaveløsingen. Selv om det var ønskelig å gi de slike frie tøyler for realismens del (og de ble oppfordret til dette), førte det til at vi ikke fikk testet systemet fra to helt ulike synspunkt ettersom testparene hadde så stor innflytelse på hverandre.

Funnene viser også at påvirkningen mellom testpersonene hadde konsekvenser for den som ble påvirket. Disse konsekvensene hadde enten positivt eller negativt utfall avhengig av hva påvirkningen innebar. Det kunne også være snakk

Tabell 13.1: Påstander angående prototypen med tilhørende totalscore

Påstand	Totalscore
1. Prototypen klarte å simulere multi-touch.	45
2. Det var vanskelig å forstå hva de ulike elementene på bordet og i rommet skulle representere.	45
3. Det var enkelt å forstå at man kunne rotere, skalere (zoome inn/ut) og flytte på mediaelementene på puslespillbrikkene.	37
4. Det var vanskelig å forstå at det dukket opp en virtuell puslespillbrikke på “Surface-skjermen” når jeg la ned et fysisk museumskort.	49
5. Det var hele tiden lett å skille mellom hva som var fysiske objekter og hva som var virtuelle objekter på Surface-bordet.	22
6. Jeg forstod lett at det egentlig ikke skal være mulig å flytte på puslespillbrikken uten å bevege museumskortet som lå oppå.	19
7. Jeg forstod lett at dersom jeg flyttet et mediaelement vekk fra puslespillbrikken, så ville den stå i ro hvis jeg deretter flyttet museumskortet og puslespillbrikken.	45
8. Det var vanskelig å forstå at puslespillbrikken og alle mediaelementene som var tilknyttet denne ville forsvinne igjen hvis jeg fjernet kortet fra bordet.	42
9. Jeg syntes det var forstyrrende at det var en annen testbruker som brukte prototypen samtidig.	37
10. Jeg følte at vi brukte skjermen “på tur”, dvs. jeg ventet til den andre gjorde seg ferdig og omvendt.	42
11. Det at det gikk veldig treigt under brukertesten ødela mye for opplevelsen av systemet.	40

Tabell 13.2: Modell av de ulike påvirkningsformene og konsekvensene av disse

	Positiv (støttende)	Negativ (ødeleggende)
Aktiv	Testperson 1 hjelper testperson 2 når han står fast.	Handlinger som testperson 1 gjør endrer skjermbildet til testperson 2 slik at enten: <ol style="list-style-type: none"> 1. handlingen testperson 2 venter på returnerer feil feedback, eller 2. testperson ikke får utført oppgaven sin fordi skjermbildet forsvant.
Passiv	Testperson 1 gjør handlinger som testperson 2 blir oppmerksom på og dette hjelper testperson 2.	Testperson 1 utfører feil handling for å løse oppgaven og testperson 2 kopierer denne handlingen.

om enten aktiv eller passiv påvirkning. I Tabell 13.2 presenterer vi en modell av disse konsekvensene. Modellen viser fire forskjellige typer påvirkning og konsekvensene disse har for den som blir påvirket (testperson 2). Den *positive, aktive* påvirkningen var den som forekom oftest i brukbarhetstestene. Den omfatter tilfellene der testpersonene hjelper hverandre enten ved at den ene parten spør eller ved at personen som har funnet løsningen deler dette med den andre. *Positiv, passiv* påvirkning innebærer de omstendighetene der den ene testpersonen legger merke til hva den andre gjør og dette hjelper ham med oppgaven. Den *negative, aktive* påvirkningen handler om situasjoner der den ene testpersonen, uten visshet om at gjerningene hans vil ødelegge for testpartneren, gjør handlinger som direkte påvirker skjermbildet til den andre testpersonen. Dette kan ødelegge for det testperson 2 holdt på med, slik at påvirkningen har en negativ virkning. *Negativ, passiv* påvirkning vil si at den ene testpersonen gjør noe som villeder den andre testpersonen til å tro at det er den "riktige" måten å løse oppgaven på.

Til tross for denne både positive og negative formen for påvirkning mellom testpersonene var det ingen av deltakerne som svarte negativt på spørsmålet om det var forstyrrende å gjennomføre testingen med en annen deltaker sam-

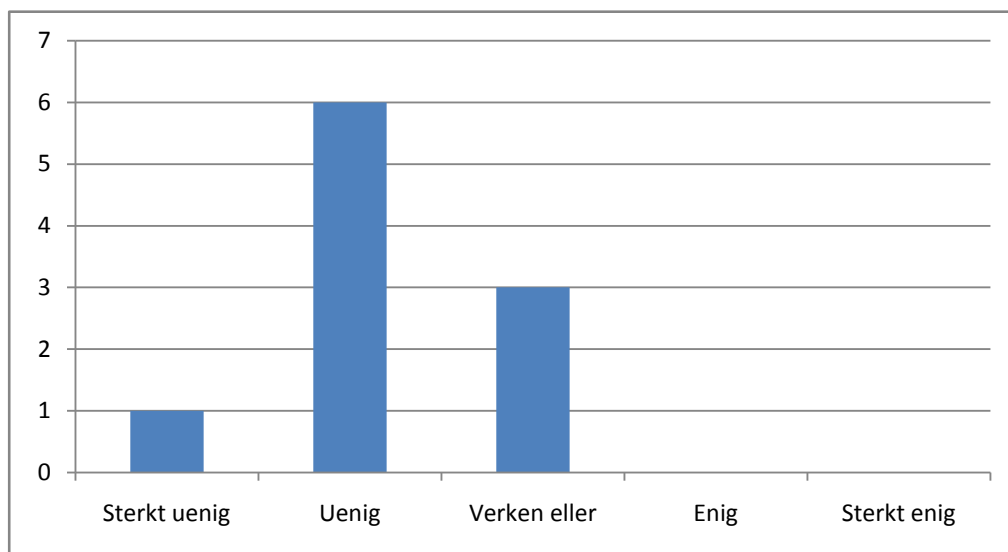
tidig. Dette kan bero på at den positive gevinsten deltakerne følte de fikk ved denne typen testing utlignet eller vektet større enn den negative.

13.2.2 Skillet mellom fysiske og virtuelle elementer

Funnene fremhever konkrete situasjoner som oppstod under testene der brukerne forvekslet de fysiske og virtuelle elementene. Dette viser at det ikke var et tydelig nok skille mellom de to typene GUI-elementer på bordet, til tross for at vi hadde forsøkt å skape et skille mellom de ved å gjøre de fysiske objektene tykkere og mer gripbare enn de virtuelle (se Seksjon 11.2.2).

Ut fra Tabell 13.1 ser vi spesielt to resultater som skiller seg ut. Disse to tilhører påstander som omhandler nettopp skillet mellom fysiske og virtuelle elementer. Den første av disse var påstand 5: *det var hele tiden lett å skille mellom hva som var fysiske objekter og hva som var virtuelle objekter på Surface-bordet*. Et søylediagram for svarene på påstanden vises i Figur 13.6. Dette spørsmålet fikk en svært lav score – 22 – som viser at deltakerne hadde store problemer med skillet. Blant annet uttalte Øyvind: *“Jeg føler plutselig at alt er i skjermen”*, som viser at han så på alt oppå bordet som en del av det virtuelle brukergrensesnittet. Også Torbjørn svarte: *“Etter hvert tenkte jeg at også kortet var et bilde. Men det har kommet jo av at det er papir, ikke sant!”*. Torbjørns svar gir en klar indikasjon på at det var vanskelig å skille mellom de to typene elementer fordi de alle bestod av papir og tynn papp. Måten museumskortene var utformet på var med andre ord ikke tilstrekkelig for å skape tydelig nok skille – de blandet seg i stedet sammen med resten av mengden.

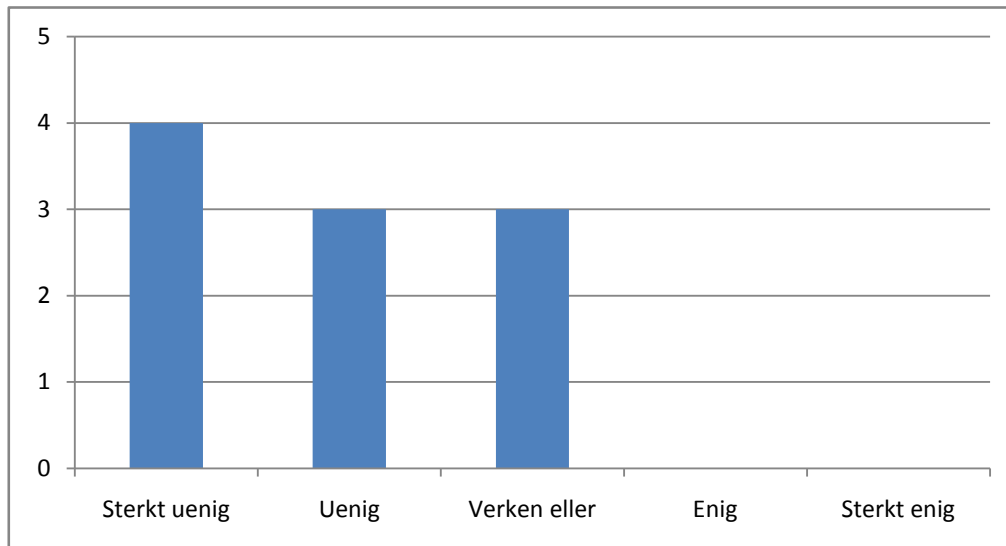
Scoren til påstand 6 var det andre resultatet som skilte seg ut. Påstanden lød som følger: *“jeg forstod lett at det egentlig ikke skal være mulig å flytte på puslespillbrikken uten å bevege museumskortet som lå oppå”*, og søker etter å gi svar på om deltakerne var klar over de begrensningene som en implementert versjon ville hatt og om de tenkte over at papirprototypen hadde de samme begrensningene. Hvis de ikke tenkte over denne begrensningen – et virtuelt GUI-element kan naturligvis ikke flytte på et fysisk objekt som er plassert på skjermen i en implementert Surface-applikasjon – ville dette bety at de ikke klarte å forestille seg dette skillet på en papirprototype. Et søylediagram for svarene på påstanden er presentert i Figur 13.7. Scoren til påstand 6 viste seg å være den laveste – 19 – i spørreundersøkelsen. Deltakerne klarte med andre ord ikke å forestille seg denne begrensningen under testen. Dette er forenlig med videoopptaket som viser at testpersonene til stadighet løftet (se Figur 13.8a) og/eller flyttet puslespillbrikkene uten å bruke museumskortet



Figur 13.6: Søylediagram for svarene på påstand 5: *det var hele tiden lett å skille mellom hva som var fysiske objekter og hva som var virtuelle objekter på Surface-bordet.*

som “flyttebrikke” (se Figur 13.8b). Frank uttalte under diskusjonen: *“hvis det hadde vært et fysisk og et faktisk virtuelt objekt så hadde jeg forstått at du flyttet på den [puslespillbrikken via museumskortet], men nå blir det liksom alt i ett”*. Sitatet understreker at det ikke var vanskelig for deltakerne å forstå puslespillbrikkens avhengighet av museumskortet i en implementert versjon, men at papirprototypen ikke setter nok begrensninger (constraints). Sandpapiret under puslespillbrikkene skapte med andre ord ikke nok lugging til at det hindret flyttingen. Dessuten stod ingenting i veien for at deltakerne kunne løfte hele brikken fra bordet. Siden vi ikke satte nok begrensninger på prototypen gjorde testpersonene det som var lettest for dem, nemlig å ta tak i puslespillbrikken når den skulle flyttes. Eivind poengterte dette under diskusjonen: *“Når det var et ark jeg kunne ta tak i, så var det mye lettere. . . og da var det ikke noe ‘feilpip’. Det gikk, og jeg fikk det resultatet jeg ønsket!”*. Eivind søkte med andre ord etter en slags indikasjon fra wizardene dersom noe slikt ikke skulle være lov. Han gjorde det som virket mest lettvtint for å flytte brikken.

Funnene gir altså klare indikasjoner på at en papirprototype for Surface-applikasjoner har vanskeligheter med å skille mellom hva som er “over og under” skjermflaten. Likevel mener vi, basert på videoopptakene, diskusjonen og spørreundersøkelsen, at testpersonene snarere *glemte* at kortene var



Figur 13.7: Søylediagram for svarene på påstand 6: “*jeg forstod lett at det egentlig ikke skal være mulig å flytte på puslespillbrikken uten å bevege museumskortet som lå oppå*”.

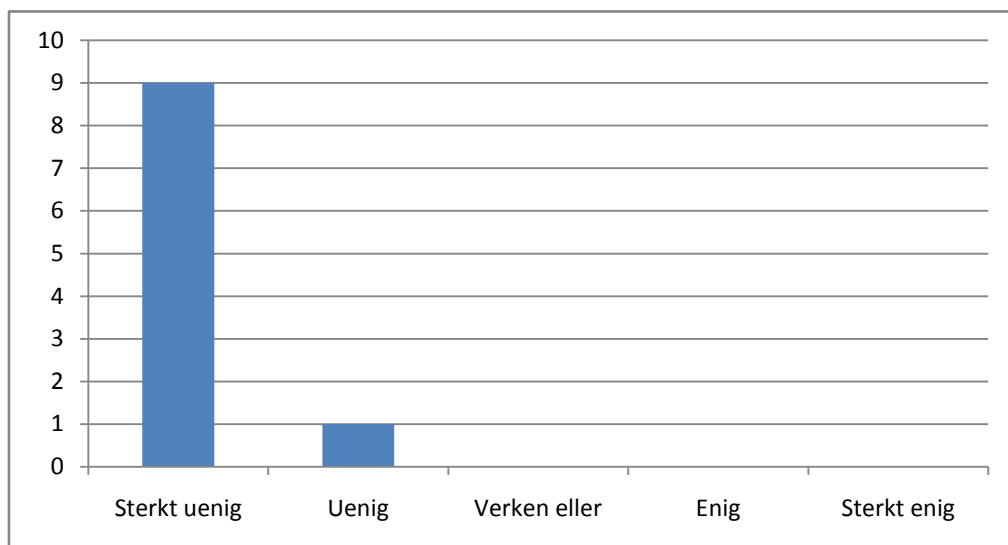


(a) Heidi flytter en puslespillbrikke ved å løfte selve brikken



(b) Eivind og Helene flytter puslespillbrikkene ved å ta på selve brikken.

Figur 13.8: Flytting av puslespillbrikker uten å bruke museumskortet.



Figur 13.9: Søylediagram for svarene på påstand 4: “*det var vanskelig å forstå at det dukket opp en virtuell puslespillbrikke på 'Surface-skjermen' når jeg la ned et fysisk museumskort*”.

fysiske enn at de ikke forstod det. Eivind bekrefter dette under diskusjonen: “*Jeg hadde vel i grunn forstått det, men jeg glemte det bort*”. Påstand 4 i spørreundersøkelsen, “*det var vanskelig å forstå at det dukket opp en virtuell puslespillbrikke på 'Surface-skjermen' når jeg la ned et fysisk museumskort*”, underbygger denne forståelsen Eivind hadde (se søylediagram i Figur 13.9). Påstanden fikk nemlig høyest score av alle påstandene; 49 av 50 mulige. Dette står i sterk kontrast til påstand 5 og 6. Det er derfor ikke urimelig å konkludere med at deltakerne forstod sammenhengen mellom det fysiske kortet og den virtuelle brikken, men at det var noe de ikke var bevisst på underveis i testen. Til å underbygge denne konklusjonen observerte vi følgende: samtlige av testpersonene som ønsket å frigjøre mer plass på skjermen *tok tak i museumskortet og ryddet bort dette uten å røre puslespillbrikken under*. De forventet dermed at den virtuelle puslespillbrikken ville forsvinne. Dette viser at når testpersonene hadde behov for det, så “husket” de at museumskortet var den fysiske linken til den virtuelle puslespillbrikken. Det kan imidlertid argumenteres for at grunnen til at de gjorde dette var at museumskortene var “deres” og puslespillbrikkene og mediaelementene var wizardene sine.

13.2.3 Superrealisme

Det var på forhånd knyttet usikkerhet rundt om papirprototypen ville klare å simulere skalering av mediaelementer, siden slike egenskaper ikke er en intuitiv måte å behandle papirlapper på (se Seksjon 11.2.1). Funnene viser også varierende resultater. Noen forstod godt at mediaelementene kunne skaleres ved å bruke fingergestikuleringen for dette (se Figur 11.17), mens andre viste tegn på det motsatte (se Seksjon 13.1.3). Torbjørn var en av testpersonene som ikke skalerte. Under diskusjonen forsvarte han dette med at han ikke ønsket å skape unødvendig arbeid for wizarden. Utsagnet impliserer at Torbjørn ikke hadde behov for å forstørre noen av mediaelementene for å løse oppgavene. Dette korrelerer med funn der testpersoner fører ansiktet nærmere bordet for å studere elementene fremfor å skalere de opp.

Likevel er det nærliggende å tro at det var flere som ikke forstod at skalering var mulig i det hele tatt. Eivind svarte eksempelvis: *“jeg fikk ikke følelsen av at man kunne skalere. Jeg skjønnte at man kunne flytte og rotere, men ikke skalere”*. Vi mener det er grunn nok til å hevde at papirets mangel på affordance til skalering, kombinert med det faktum at det ikke var høyst nødvendig for å løse oppgavene, gjør at flere testpersoner ikke benyttet seg av skaleringsfunksjonen.

13.2.4 Synkronisering

Testoppgavene var formulert slik at de i seg selv skulle oppfordre parene til å samarbeide på oppgave 2 og 6. Selv om funnene viste at flere av testparene var usynkroniserte under de individuelle oppgavene, samarbeidet 4 av 5 testpar på oppgave 2 og 6.

Hensikten med å få brukerne til å bli synkroniserte under nettopp disse to oppgavene var å se hvorvidt bordet oppfordret til samarbeid. Testpersonene hadde frie tøyler til å gjøre som de ville på oppgavene og ingenting stod i veien for at de kunne “låne” brikkene av hverandre for å løse hver sin oppgave. Likevel var det bare testpar 5 som faktisk gjorde det på denne måten. Med dette kan vi hevde at oppgavene oppfylte noe av målsetningen ved testene: å kartlegge MultiVitens potensial til samarbeid mellom brukere.

Alternativet til å la oppgavene stå for synkroniseringen ville være å holde den ene testpersonen igjen til den andre hadde fullført sin tilsvarende oppgave, og gitt neste oppgave til begge samtidig. Dette ville riktignok synkronisert testpersonene, men det kunne ha ledet til at de ikke ville opplevd å ha frie tøyler, men derimot vært mer “tvunget” til å samarbeide.

13.2.5 Wizard-of-Oz

Bruk av wizard-of-Oz-teknikken på flere brukere samtidig fungerte godt. Dedikerte wizards for hver testbruker viste seg å være nødvendig, siden det til enhver tid var mye interaksjon med prototypen fra begge parter. Vi hadde nok med å følge med på den dedikerte testbrukeren og gi feedback til handlingene som ble gjort og kan ikke se for oss at én wizard ville klart å “behandle” begge testbrukerne samtidig.

Under brukbarhetstestene fikk vi oppleve at det var vanskelig å fokusere på annet enn wizardrollen. Behovet for video- og lydopptak var derfor essensielt for analysen av brukbarhetsfeil som hadde oppstått.

13.2.6 Fungerte brukbarhetstesten til sitt formål?

Til tross for at vi gjennom brukbarhetstestene oppdaget flere problemer tilknyttet flerbrukertesting av en papirprototypet Surface-applikasjon, er det ikke med dette sagt at brukbarhetstestene ikke oppfylte de målene en brukbarhetstest skal. Det primære målet med brukbarhetstesting er å avdekke potensielle brukbarhetsfeil og testene avslørte flere slike.

Brukbarhetstestene avdekket fem sentrale brukbarhetsproblemer (se Tabell 12.2) som det hadde vært nødvendig å utbedre i senere redesign dersom dette hadde vært en reell systemutviklings situasjon. Dette viser at papirprototyping klarer å formidle designet til en multi-touchapplikasjon for Surface godt nok til at testbrukere forstår konseptet, interaksjonsmønsteret og hensikten ved applikasjonen – og at de er kritisk til designet.

13.3 Konklusjon

Basert på analysen av funnene konkluderer vi med at utviklere kan ved hjelp av enkel papirprototyping avdekke potensielle brukbarhetsfeil ved Surface-applikasjoner allerede på et tidlig stadium. Brukbarhetsfeil ble funnet i MultiViten uten å være avhengig av å implementere og teste den på en faktisk Surface-enhet. Det viste seg også at testing av flere brukere samtidig – med dedikerte wizards, parallelle oppgavesett og synkronisering ved hjelp av oppgaveformuleringer – egnert seg godt til å teste prototyper som inneholder mulighet for samarbeid og interaksjon med flere brukere samtidig.

Likevel er det mange utfordringer knyttet til prototyping og flerbrukertesting av applikasjoner for Surface. Når man som utvikler ønsker å gjennomføre brukersentrerte prosesser som dette må man være oppmerksomme på disse utfordringene og ta de til etterretning. For å bidra til anbefalinger rundt prosessen vil vi komme med forslag til hvordan disse utfordringene kan imøtekommes.

13.3.1 Materialer og teknikker for prototyping

Ekspenteringen med materialer og teknikker for prototyping av Surface-applikasjoner ga indikasjoner på hvilke av disse som fungerte bedre enn andre. Først og fremst fant vi ut at fordelene med papirprototyping overgikk bruken av projiseringsteknikken. Dessuten ble vi oppmerksomme på at vi hele tiden spurte oss selv underveis: *“er denne prototypingsteknikken mulig å teste ved bruk av wizard-of-Oz?”*. Vi mener at dette er noe man må ta stilling til når man skal prototype og senere brukbarhetstesten en applikasjon for Surface. Ved regelmessig å teste om prototypen som utvikles lar seg wizard-of-Oz brukbarhetsteste vil man unngå testseanser som ikke lar seg gjennomføre. Gjennom eksperimentering fant vi ut at følgende teknikker og materialer egner seg godt til prototyping og wizard-of-Oz flerbrukertesting av Surface-applikasjoner:

Glasskille

Vi undersøkte om en glassplate kunne skille det som skjedde over skjermen (brukerens verden; bruk av hender/fingre og fysiske objekter) fra det som skjedde under (maskinens verden; endringer av skjermbilder). Glasset bidro til et tydelig skille mellom det fysiske og det virtuelle, men det viste seg imidlertid at teknikken gjorde arbeidet vanskelig for wizarden; GUI-elementene må holdes inntil undersiden av glassplaten, som betyr at prototypen begrenses til det som kan holdes oppe av wizardens to hender. Wizarden må dessuten kunne se hva som skjer på oversiden, samt gjøre ytterligere anstrengelser for å kontrollere lagvise GUI-elementer.

Magneter

Det ble undersøkt om magneter kunne løse problemene forbundet med glasskillet. Teknikken bidro til å lette wizarden for elementer som må holdes inntil

glasset. Magnetene klarer likevel ikke å løse problemet med lagvise GUI-elementer under glasset. Bruk av magneter vil likevel kunne ha verdi for skillet mellom det fysiske og det virtuelle i tilfeller av applikasjoner der det virtuelle ikke har øvre lag som skal kunne flyttes på. Magnethansker ble også testet ut, men introduserte nye problemer. Konklusjonen er at magneter ikke er en egnet som materiale i alle tilfeller.

Postit-lim

Postit-lim fungerte bra til å holde mediaelementene fast på puslespillbrikkene når de ble flyttet av wizardene. Vi måtte for øvrig være forsiktige med mengden slik at det ikke ville være vanskelig å flytte mediaelementene vekk fra brikken og rundt på skjermflaten.

Sandpapir

Bruken av sandpapir fungerte bra mellom museumskortet og puslespillbrikken, men manglet evnen til å “lugge” nok mellom puslespillbrikken og underlaget. Dette viser at sandpapir fungerer bra til å skape affordance, men dårlig når det brukes som constraint.

Transparente ark

Bruk av transparente ark for å skape constrainten som gjorde at testbrukerne ikke kunne flytte på statiske GUI-elementer fungerte veldig bra. Bruken av dette vil derfor anbefales.

13.3.2 Påvirkning

Resultatene avdekket at under flerbrukertesting påvirker testbrukere hverandre i stor grad og på ulike måter. Det gjør at brukbarhetsfeil man ellers ville ha kunnet avslørt hvis den ene testpersonen hadde måttet tenke alene blir forbigått idet han får hjelp av den andre parten. På den andre siden vil man få avdekket problemer ved designet som utelukkende kan oppdages ved at flere opererer på skjermen samtidig, som eksempelvis “mine” og “dine” GUI-elementer. Et annet positivt aspekt ved flerbrukertesting er at testdeltakerne får en trygghetsfølelse. Tanken om at de “spiller på lag” er med på å dempe angsten som kan forekomme dersom de føler at brukbarhetstesten handler om å teste personen og ikke designet.

“Muldvarp”

Under brukbarhetstestene hadde vi ingen kontroll over innflytelsen testpersonene hadde på hverandre. Som funnene viser førte dette til uforutsigbar påvirkning som i flere tilfeller førte til at potensielle brukbarhetsfeil man ellers kunne oppdaget forble skjult. En måte man kan kontrollere testforløpet på, samtidig som opprettholder flerbrukeraspectet, er å bruke en “muldvarp”. Denne muldvarpen vil opptre som en testbruker på lik linje med den andre, men vil samtidig være en del av teststaben. Muldvarpen vil vite hvordan han skal opptre overfor den “egentlige” testbrukeren i forskjellige situasjoner under testen. Han vil oppføre seg som en tilfeldig testdeltaker og samarbeide med den andre parten der det er naturlig, men han vil unngå å ta på seg en ledende, fremdrivende rolle. På denne måten vil sjansen være større for å observere tilfeller der den egentlige testbrukeren ikke forstår designet. Ved bruk av en slik muldvarp vil man også kunne styre den ene variabelen som denne testbrukeren utgjør på andre måter. Eksempelvis vil det være mulig å teste hvordan en bruker interagerer med prototypen når en annen bruker (muldvarpen) opptar “for mye plass” på skjermen.

13.3.3 Skillet mellom fysiske og virtuelle elementer

Problemet som pekte seg tydeligst ut er skillet mellom fysiske og virtuelle elementer. Testdeltakerne viste klare problemer med å holde rede på at museumskortene var fysiske, mens resten av elementene på bordet var virtuelle. Likevel har resultatene vist at deltakerne ikke hadde vansker med å forstå at noen av papirfigurene var fysiske mens andre var virtuelle, de snarere glemte dette skillet underveis i testen siden alle papirbitene så nokså like ut der de lå på bordet. Dette er en av de store fordelene en implementert versjon vil ha overfor en papirprototype.

Tydeliggjøring av skillet

Problemet med at testpersonene hadde vansker med å skille de fysiske fra de virtuelle elementene kan bero på at tykkelsen på pappen som ble brukt til å gjøre museumskortene mer gripbare og forskjellige fra de virtuelle GUI-elementene var for tynn. For å gjøre de fysiske elementene enda mer gripbare og forskjellige fra materialet som ble brukt til de virtuelle elementene kunne det vært brukt tykkere materiale til å representere museumskortene, eksempelvis klosser. Dette ville trolig ha minnet brukerne på at museumskortene var fysiske i motsetning til de andre GUI-elementene under testene.

13.3.4 Superrealisme

Vi har gjennom brukbarhetstestene bekreftet at det er en del egenskaper ved multi-touch som er utfordrende å simulere. Spesielt så vi at den superrealistiske egenskapen skalering ble tatt i mindre bruk enn rotering og flytting. Dette kan bero på måten prototypen simulerte dette på – ved å kreve at brukeren benyttet gesturen “pinch” på papirlappen – ikke var selvforklarende for testbrukerne, men at de ville hatt behov for innføring på forhånd.

Det er verdt å merke seg at testdeltakerne var informatikkstudenter og det er rimelig å anta at de har større teknologisk innsikt enn normalen. Når denne brukergruppen har vansker med å skille mellom fysiske og virtuelle elementer, samt superrealistiske egenskaper kan man med all grunn anta at ikke-teknologiske testbrukere vil slite med det samme.

Innføring i interaksjonsprinsippene

Som nevnt i konklusjonen var skalering en lite utprøvd teknikk under brukbarhetstestene. Årsaken til dette mente vi var at skalering ikke lar seg utføre direkte på en papirlapp og at brukerne burde fått en innføring til interaksjonsprinsippene i forkant av testene. Tanken var at videoen som viste Surface i praksis skulle kunne demonstrere hvordan interaksjonsprinsippene var på den virkelige Surface og at deltakerne ville overføre disse prinsippene til papirprototypen. Imidlertid viser det seg at testbrukere på forhånd bør få en eksplisitt introduksjon til hvilke konkrete gestikuleringer de kan utføre på papirlappene.

13.3.5 Sammenligning med annen forskning

Derboven et al. [18] kom i sitt studie frem til at man må være forsiktig med å generalisere low-fi resultater til hi-fi ved bruk av konvensjonelle verktøy som saks, blyant, teip og lignende (se Seksjon 1.2.2). De etterlyste nye metoder for low-fi prototyping som kunne gjøre resultater fra brukbarhetstester på disse mer forutsigbart for hi-fi resultater. I vårt studie har vi eksperimentert med ulike teknikker og materialer for low-fi prototyping og bekreftet at disse avdekker brukbarhetsfeil – feil som antakeligvis også ville vært til stede i en hi-fi versjon. Vi har på denne måten imøtekommet etterlysningene fra Derboven et al. med et bidrag for videre forskning og praksis.

Sohan [68] peker på viktigheten av flerbrukertesting av multi-touchapplikasjoner som er ment å kunne brukes av flere personer samtidig. Dette fokuset har

vi ivaretatt i dette studiet og har påvist muligheten for å wizard-of-Oz flerbrukerteste en papirprototype for Surface.

Del IV

Forskningsspørsmål 3

Kapittel 14

Anbefalinger

Under arbeidet med FS1 og FS2 har vi utforsket en mulig måte å praktisere brukersentrert utvikling på for applikasjoner til Surface i de tidlige fasene av systemutviklingsløpet. På denne veien har vi avdekket problemer og utfordringer knyttet til metodene som ble brukt, men også fått bekreftet gode løsninger.

På bakgrunn av resultatene fra FS1 og FS2 vil vi gjøre rede for våre anbefalinger. Anbefalingene er ment som en rettleiding for utviklere i prosjekter der det skal utvikles programvare for Surface og brukergruppen er det vi anser som typiske for denne plattformen – personer uten praktisk erfaring med multi-touchbord.

Anbefalingene er gruppert i tre deler:

- Introduksjon av teknologien.
- Low-fi prototyping.
- Wizard-of-Oz flerbrukertesting.

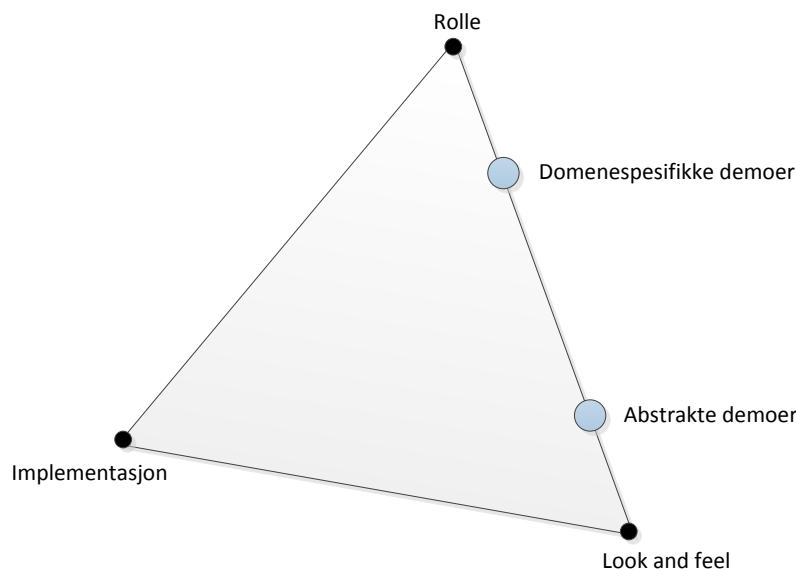
I hver del vil det gis en overordnet begrunnelse for anbefalingene, samt konkrete punkter for praksis.

14.1 Introduksjon av teknologien

FS1 viste at deltakerne ved begge workshopene hadde utbytte av å få en demonstrasjon av egenskapene ved Surface. Etersom teknologien var relativt ukjent for deltakerne var en slik innføring nødvendig for å forstå hvilke

muligheter Surface åpner for ved design av applikasjoner. Resultatene viste ingen stor forskjell mellom antall ideer som ble generert under workshopene. Det var imidlertid større variasjon mellom *typene* av ideer som kom opp. Ideene under den workshopen der deltakerne fikk interagere med domenespesifikke demoer brukte en større del av Surface sitt fulle potensial, mens ideene på motsatt workshop var tilsynelatende mindre påvirket av eksemplene som ble brukt.

Når det gjennomføres designworkshops med et utvalg av brukere i en tidlig utviklingsfase vil vi anbefale å gi deltakerne hands-on erfaring med både abstrakte og domenespesifikke formidlingsdemoer. Disse konseptene bør demonstreres: *direkte manipulasjon, interaksjon med fysiske objekter, samhandling og kombinasjon med andre teknologier*. Deltakerne vil dermed både få en innføring i interaksjonsprinsippene og se hva prinsippene kan brukes til. Dette vil utgjøre to ulike fokus på kontinuumet mellom *rolle* og *look and feel* på Houde og Hills triangel (se Figur 14.1), og sammen gi deltakerne to ulike innfallsvinkler på teknologien; hvordan man interagerer med Surface og hvilken rolle den kan utgjøre. Selv om demonstrasjonene bør implementeres og demonstreres på en ordentlig Surface, vil ikke demoene befinne seg mot *implementasjon* i triangelet. Dette er fordi *hensikten* med demoene ikke er å få svar på hvordan de kan implementeres, men snarere vise hvilke funksjoner teknologien kan utfylle og hvordan den er å se og føle på.



Figur 14.1: Demoenes ulike fokus i Houde og Hills modell av hva prototyper prototyper (se Seksjon 3.5.5).

Anbefalinger:

- Gi brukerutvalget en hands-on demonstrasjon av Surface.
- Demonstrer konseptene direkte manipulasjon, interaksjon med fysiske objekter, samhandling og kombinasjon med andre teknologier ved å kombinere abstrakte og domenespesifikke demonstrasjonseksempler:
 - Gi abstrakte eksempler ved å bruke geometriske figurer og farger for å demonstrere interaksjonsprinsipper.
 - Gi konkrete, domenespesifikke eksempler ved å vise hva prinsippene kan brukes til for å demonstrere teknologiens rolle.

14.2 Low-fi prototyping

Under besvarelsen av FS2 ble det sett på ulike materialer og teknikker som kunne brukes for å low-fi prototype en Surface-applikasjon og hvordan denne kunne brukbarhetstestes. Vi fant ut at en konvensjonell papirprototype klarer å simulere egenskapene ved Surface og samtidig avdekke brukbarhetsfeil ved en applikasjon for denne plattformen. Vi vil derfor anbefale å utvikle papirprototyper for en Surface-applikasjon på et tidlig stadium fremfor å legge mye ressurser i en implementert utgave. Gevinsten vil være stor i form av mange avdekkede feil i en billig og rask prototype.

Anbefalinger:

- Lag en papirprototype for designløsningen av Surface-applikasjonen ved å ta i bruk kreative materialer og teknikker. På bakgrunn av vårt studie har vi følgende anbefalinger til materialer:
 - Postit-lim for å holde lagvise GUI-sammensetninger sammen, men fortsatt tillate at brukere flytter på de ulike elementene i sammensetningen.
 - Sandpapir for å skape friksjon mellom fysiske og tilhørende virtuelle elementer slik at det fysiske objektet “affords” indirekte flytting av det virtuelle elementet.
 - Transparente ark for å hindre flytting av statiske GUI-elementer.
- Sørg for at prototypen lar seg wizard-of-Oz flerbrukerteste ved å gjennomføre regelmessige, interne pilottester under utviklingen.

14.2.1 Vær tydelig på skillet mellom fysiske og virtuelle elementer

Papirprototypen bør ta spesielt hensyn til tydeliggjøring av skillet mellom det som skal være fysiske elementer og det som representerer det virtuelle som vil foregå på skjermen. Resultatene fra FS2 viser at testbrukere lett vil kunne glemme denne forskjellen under testingen. For at dette ikke skal gi unødig forvirring og overskygge problemer relatert til designets brukbarhet, bør materialet som brukes for å representere de fysiske elementene skille seg tydelig fra materialet som representerer det virtuelle.

Anbefalinger:

- Bruk annet materiale for de fysiske objektene, eksempelvis plast eller treverk.
- Gjør de fysiske objektene tredimensjonale og bevar de virtuelle elementene flate.
- Dersom testbrukerne viser problemer med å skille de fysiske elementene fra de virtuelle under brukbarhetstestene bør det gis en forklaring der og da (ta en timeout). Dette er viktig for å ikke spolere testen av designet på grunn av misforståelser rundt prototypen.

14.3 Wizard-of-Oz flerbrukertesting

Papirprototypen bør testes ved bruk av wizard-of-Oz. Før testene bør testbrukerne få en introduksjon til interaksjonsteknikkene på prototypen. Wizardene må også være ekstra oppmerksomme på gestures som testbrukerne gjør på GUI-elementene, siden prototypen skal kunne respondere på slik input i tillegg til ordinære trykk.

Siden Surface-applikasjoner som regel skal kunne brukes av flere personer samtidig vil det være nødvendig å gjennomføre flerbrukertesting. Under gjennomføringen av brukbarhetstestene i dette studiet ble prototypen testet av *to* brukere simultant. Dette fikk avdekket brukbarhetsfeil som ikke hadde vært oppdaget ved testing av én og én person. Vi oppdaget imidlertid at kompleksiteten ved å teste flere brukere på en gang er betraktelig større enn ved enbrukertesting. En slik test krever derfor god planlegging.

En må være spesielt oppmerksom på at under flerbrukertesting vil testbrukerne påvirke hverandre, både positivt og negativt. Dette vil medføre at brukbarhetstestene ikke gir fullt utbytte i form av avdekking av potensielle brukbarhetsproblemer ved designet. Derfor bør en “muldvarp” engasjeres – en med kjennskap til både systemet og testopplegget. Supplerende tester med én testbruker og denne muldvarpen bør så gjennomføres der sistnevnte minsker påvirkningen på testbrukeren.

14.3.1 Parallele oppgavesett

Testbrukerne bør få parallele oppgavesett. Dette viste seg å fungere godt under brukbarhetstestene. Testbrukerne er på den måten fri til å gjøre det som føles naturlig med tanke på samhandling med andre parter. Oppgavene kan til en viss grad kontrollere testforløpet med tanke på å synkronisere testbrukerne, men det må presiseres at det til syvende og sist er testbrukerene selv som bestemmer hvordan testforløpet vil arte seg.

14.3.2 Dedikerte wizards

Flerbrukertesting med wizard-of-Oz krever at testbrukerne har en dedikert wizard hver. Dersom flere brukere skal interagere med prototypen samtidig, vil ikke én wizard klare å fokusere på interaksjonen som testbrukerne gjør uavhengig av hverandre. Brukbarhetstestene vi utførte viste at dedikerte wizards fungerte bra. Verken vi eller deltakerne følte at dette skapte forvirring under testene.

Anbefalinger:

- Gjennomfør brukbarhetstesting av prototypen ved å bruke wizard-of-Oz.
- Forbered testene grundigere enn ved vanlig brukbarhetstesting.
- Gjennomfør flerbrukertesting av prototypen:
 - Test prototypen på fem testpar.
 - Supplér testene med to-tre tester med én testbruker og én muldvarp.

- Gi brukerne parallelle oppgavesett. La oppgavene synkronisere testparene.
- Dedikér én wizard til hver testbruker.

14.3.3 Datainnsamling

Kompleksiteten ved flerbrukertesting krever god datainnsamling for å analysere brukbarheten ved et design. Det vil være vanskeligere å observere interaksjonen underveis i forhold til en tradisjonell enbrukertest. Video- og lydopptak blir dermed viktige kilder for videre analyse. Det anbefales også å gjennomføre testene i et testlaboratorium for å minske ytre påvirkning og støy, og for å sikre god plass rundt prototypen.

Anbefalinger:

- Gjør video- og lydopptak av testene.
 - Sørg for at et kamera filmer prototypen ovenifra på nært hold, mens et annet filmer det som skjer rundt bordet.
 - Sikre godt lydopptak ved å bruke en sensitiv kondensatormikrofon i taket.
- Sikre god plass rundt prototypen og forminsk ytre støy.

Del V

Avslutning

Kapittel 15

Metodediskusjon

Vi har i denne oppgaven tatt for oss hvordan man kan inkludere brukerne under utvikling av applikasjoner for Surface, og kommet frem til anbefalinger for brukerinvolvering i slike prosjekter i de tidlige fasene av utviklingsløpet. Det må likevel tas i betraktning at vi kun har gjort forsøk på en liten gruppe mennesker i en bestemt kontekst og at resultatene ikke nødvendigvis kan generaliseres til alle situasjoner. Våre funn er snarere et produkt av observasjoner og erfaringer fra eksperimentering med brukersentrerte metoder til Surface-plattformen. Noen av funnene er imidlertid svært tydelige og kan antas å være gyldige utover våre konkrete forskningstilfeller, spesielt utfordringene omkring flerbrukertesting og papirprototyping til Surface.

I denne oppgaven har Surface blitt brukt som plattform for å representere multi-touchbord. Surface har egenskaper som andre bord ikke har (eksempelvis tag-gjenkjenning), men mangler også funksjoner som andre igjen innehar (eksempelvis bruk av pekepen, form/omriss- og strekkodegjenkjenning). Likevel har Surface de samme fundamentale multi-touchegenskapene som andre kommersielle multi-touchbord per dags dato. Størrelsesmessig har også andre multi-touchbord samme potensial til å brukes av flere personer samtidig. Vi vil derfor påstå at funnene i denne oppgaven, bortsett fra de som relateres til Surface-spesifikke egenskaper, kan generaliseres til å gjelde for andre kommersielle multi-touchbord.

Forskningsmetodene som er anvendt vil videre drøftes i de neste seksjonene. Det vil argumenteres for forskningens validitet, hovedsaklig basert på Klein og Myers' prinsipper for fortolkende feltforskning (se Seksjon 5.8) og Zimmermans kriterier for Research through design (se Seksjon 5.6). Seksjon 15.1 vil ta for seg validiteten til FS1, mens Seksjon 15.1 argumenterer for validiteten til FS2.

15.1 Validiteten ved FS1

I etterkant av workshopene som ble gjennomført for å svare på FS1 ser vi flere faktorer som kan ha påvirket resultatene. Som nevnt i Seksjon 6.8 vil et utvalg av Klein og Myers' prinsipper om fortolkende feltforskning være relevante for validiteten ved FS1.

15.1.1 Ulik grad av forkunnskaper

Ulikheter mellom gruppenes forkunnskaper, sosiale faktorer innad i gruppene og det lave nivået av utfordring på demoene kan ha bidratt til å skape feilkilder for forskningen. Klein og Myers' syvende prinsipp, *mistanke*, krever oppmerksomhet rundt slike feilkilder. Deltakerne på W2 hadde mer kjennskap til multi-touch generelt og Surface spesielt i forkant av workshopen. Tabell 15.1 viser at tre personer på W2 hadde prøvd Surface tidligere i kontrast til én person på W1. Denne forskjellen kan ha ført til at W2 hadde mindre bruk for formidlingsdemoene enn det W1 hadde for å forstå egenskapene ved bordet. Deltakerne som har prøvd en Surface-enhet tidligere kan også ha blitt påvirket av applikasjonene de har fått prøve da når de skulle komme med egne ideer. Som påpekt i Seksjon 5.2.1 bør variabler som kunnskapsnivå være konstant i slike eksperimenter. Dersom dataene i dette tilfellet skulle vært mer internt valide måtte deltakerne ha innehatt like stor grad av forkunnskaper om Surface og multi-touch generelt. Ideelt sett skulle pre-spørreundersøkelsen vært utført i forkant av workshopene, slik at gruppene kunne vært likt fordelt med grunnlag i forkunnskaper. Likevel kom det fram under gruppeintervjuet at selv om de svarte "ja" på spørsmål om hvorvidt de var kjent med teknologien, var denne kjennskapen langt i fra dyp – de hadde snarere "hørt om Surface et sted". Dette tilsier at ulikheten mellom deltakernes forkunnskaper på de to workshopene ikke var en signifikant variabel.

Tabell 15.1: Sammenligning av forkunnskaper mellom W1 og W2

	Workshop 1	Workshop 2
Hadde hørt om Surface tidligere	1	4
Hadde prøvd Surface tidligere	1	3
Eier enhet med multi-touch	3	3
Eier ikke, men har prøvd multi-touch	2	3

15.1.2 Sosiale faktorer

Det tredje prinsippet til Klein og Myers, *interaksjon mellom forsker og forskningsobjekt*, krever en kritisk refleksjon over de sosiale aspektene som har vært til stede under innsamlingen. Vi var på forhånd klar over at sosiale faktorer kunne ha noe å si for resultatene. Elevene var blant annet satt sammen med det samme kjønn under parjobbingen, mens læreren og den museumsansatte samarbeidet (se Seksjon 6.8). Dette gjorde vi i tro om at dynamikken innad i parene ville være best mulig. Basert på våre observasjoner under parjobbingen viste dette seg å fungere. Under diskusjonen etter parjobbingen og generelt resten av workshopen hadde elevene en god tone både seg imellom og med læreren. Det kan likevel antas at elevene i en viss grad ble påvirket av lærerens nærvær. Lærerne på begge workshopene virket autoritære og elevene hadde tydelig respekt for dem. Dette kan utgjøre både positiv og negativ effekt på arbeidet under workshopene. Det kan både påvirke hvor produktive elevene er og hva ideene inneholder. Det kan tenkes at dette hadde noe av skylden for at flere av ideene var skolerelaterte, men det kan også like gjerne være naturlig for elevene å tenke i slike baner, spesielt når de er klar over at de ble rekruttert gjennom skolen. For å unngå dette kunne vi rekruttert et mer tilfeldig og spredt utvalg av personer til workshopene enn elever og lærere som kjente hverandre fra før.

15.1.3 Formidlingsdemoene

Som nevnt i Seksjon 8.2.3 og Seksjon 8.3.3 hadde ikke deltakerne problemer med å forstå formidlingsdemoene. Noen mente snarere at de var i enkleste laget. Grunnen til dette var at det som ble demonstrert hadde store likhetstrekk med teknologi de hadde prøvd før. At en multi-touchskjerm kan styres via direkte manipulasjon var noe de kjente igjen fra først og fremst mobiltelefoner med multi-touch. Det som var helt nytt for dem var interaksjonen med fysiske objekter. De likte tanken på å kunne feste tags på gjenstander for så å kunne gjøre ting på skjermen med disse gjenstandene. Siden tanken bak workshopene var å formidle teknologi som var ukjent for deltakerne ville det vært hensiktsmessig å fokusere mer på hva som gjør Surface til noe de ikke hadde sett før. Demoene for de grunnleggende prinsippene ved multi-touch, som direkte manipulasjon, måtte fortsatt ha vært med for å vise at bordet har multi-touchegenskaper, men vi kunne ha brukt mer tid og flere demoer på å vise bruken av tags.

15.1.4 Modning

Det var første gang vi hadde planlagt og gjennomført workshops og vi var oppmerksomme på at vi kom til å lære svært mye den første dagen. Derfor bestemte vi oss på forhånd for at vi ville gjennomføre den andre workshopen helt likt som den første, uavhengig om den første kunne vært gjort bedre. Likevel kommer vi ikke utenom det faktum at vi modnet som fasilitatorer ubevisst. Det kan tenkes at vi ble oppfattet som mindre stresset av deltakerne på W2 og at dette har gjort utslag på resultatene. Dersom vi hadde gjennomført flere workshops ville antakelig graden av modning avtatt og stabilisert denne variabelen.

15.1.5 Validiteten

Basert på refleksjonene vil vi påstå at resultatene er internt valide. Til tross for at workshopene hadde ulike forutsetninger og øvrige faktorer varierte, var ikke dette signifikant nok til å motsi den interne validiteten.

Den eksterne validiteten kan det rettes kritikk mot. Det burde ha vært gjennomført flere identiske workshops med andre grupper for å kunne se om resultatene var konsistente. Likevel vil vi hevde at forsøket inkluderte en brukergruppe som var representativ for forskingspørålet og at caset var realistisk. Forsøket var heller ikke avhengig av en spesiell type deltakere – tilsvarende workshops kunne vært gjennomført i et annet domene for en annen brukergruppe. Vi har derfor valgt å presentere funnene som generaliserte resultater i konklusjonen, men vil påpeke at disse må etterprøves for å kunne tilføre ny viten til forskningsmiljøet og praktikere.

15.2 Validiteten ved FS2

For å svare på FS2 ble metoden Research through design brukt som overordnet rammeverk. Basert på Zimmermans kriterier for validitet ved denne typen forskning vil vi argumentere for gyldigheten ved resultatene.

15.2.1 Prosess

Kapittel 11 dokumenterer eksperimenteringsprosessen vi gjennomgikk for å kunne komme frem til hvilke teknikker og materialer som egnet seg best for

prototyping av Surface-applikasjoner. Kapitlet beskriver også i detalj drøftingen vi gjorde rundt gjennomføring av brukbarhetstestene. Gjennomføringen og funnene fra brukbarhetstesting er også grundig skildret med rike situasjonsbeskrivelser og sitater.

Vi vil påstå at denne dokumentasjonen er tilstrekkelig for å kunne gjenskape prosessen eller for å gjennomføre lignende arbeid.

15.2.2 Nyskapning

Prototypingsteknikkene, -materialene og testmetodene vi har eksaminert vil kunne bidra til nyskapende måter å evaluere designløsninger for Surface-applikasjoner på. Anbefalingene vi kommer med er innovative i den grad at vi ikke har lyktes å finne litteratur eller eksempler på praksis som tar for seg samme problematikk.

15.2.3 Relevans

Anbefalingene våre rundt prototyping og brukbarhetstesting av applikasjoner for Surface svarer etterlysningen av low-fi metoder for denne plattformen (se Seksjon 1.2). Med disse grepene vil terskelen være lavere for å gjennomføre brukersentrert utvikling for Surface og videre føre til økt brukskvalitet.

15.2.4 Utvidningsevne

Selv om vi prototypet et konkret applikasjonskonsept for et gitt domene, er anbefalingene abstrahert vekk fra disse konkrete produktene. Vi har dermed gjort det mulig å anvende retningslinjene våre i andre domener og applikasjoner for Surface.

15.2.5 Validiteten

Den interne validiteten for resultatene av FS2 avhenger av eventuelle feilkilder under brukbarhetstestene. Vi sørget for at de eneste variablene under testene var testparene. Vi kunne ha risikert at disse hadde svært ulik grad av forkunnskaper, men under den uformelle samtalen før testforløpet avdekket vi at ingen av deltakerne hadde mer kunnskap enn andre.

Vi har vist at forskningen oppfylte Zimmermans fire kriterier for valid forskning gjennom design og hevder basert på dette at resultatene har ekstern validitet. Testdeltakerne og konteksten var representativt for virkeligheten, samtidig som forskningen ikke var avhengig av dette caset. Funnene kan ikke sies å være statistisk signifikante basert på det lille antallet deltakere og brukbarhetstester. Likevel var funnene så markante at det er all grunn til å tro at de vil forekomme i forskning av større skala.

Kapittel 16

Konklusjon og Videre Forskning

Dette kapitlet besvarer forskningsspørsmål 1 og 2 basert på konklusjonene fra arbeidet som ble gjennomført. Disse resultatene har dannet grunnlaget for å svare på det tredje forskningsspørsmålet, som gir anbefalinger for brukersentrert utvikling for Surface. Seksjon 16.2 avslutter oppgaven med tanker rundt videre forskning på området.

16.1 Konklusjon

FS1: I en brukerdeltakende designworkshop, hvilken merverdi gir domenespesifikke demonstrasjonseksempler fremfor abstrakte ved formidling av mulighetene og begrensningene til Surface?

Til tross for at workshopene resulterte i liten forskjell i antall ideer som deltakerne kom opp med, viste workshopene en merkbar forskjell i ideenes karakter. De domenespesifikke demoene hjalp deltakerne med å se mulighetene ved Surface og ga dem inspirasjon til videre idégenerering. Likevel viste det seg at deltakerne lot seg påvirke av de konkrete eksemplene som ble brukt i de domenespesifikke demoene og bygget ideene rundt disse. Resultatene indikerte også at de abstrakte demoene hadde sin nytteverdi. Selv om disse viste seg å gjøre det vanskelig for deltakerne å se konkrete bruksområder, bidro de til at deltakerne reflekterte mer over hvordan teknologien fungerer.

FS2: Hvilke metoder, verktøy og materialer bør benyttes ved pro-

totyping av applikasjoner for Surface i de tidlige fasene av systemutviklingsløpet, og hvordan kan disse prototypene brukbarhetstestes ved hjelp av wizard-of-Oz?

Resultatene fra brukbarhetstesting viste at konvensjonell papirprototyping av brukergrensesnitt til Surface klarte å avdekke brukbarhetsfeil. Ved å gjennomføre testingen på flere testbrukere samtidig med parallelle oppgavesett vil man også kunne eksaminere hvordan prototypen fungerer i en sosial setting – en setting som Surface ofte er ment for.

Testsesjonene avdekket også utfordringer knyttet til flerbrukertesting og bruk av konvensjonell papirprototyping for Surface. For å gi et svar på forskningsspørsmålet har vi utarbeidet flere forslag til hvordan disse utfordringene kan imøtekommes:

- **Påvirkning:** Testbrukere vil i stor grad påvirke hverandre under testene. Det kan derfor være tjenlig å supplere med noen tester der en “muldvarp” kan kontrollere påvirkningen.
- **Skille mellom fysiske og virtuelle elementer:** Det vil være vanskelig for testbrukere å holde rede på hvilke av elementene på “skjermen” som er fysiske (med tags på) og hvilke som er virtuelle. Det er derfor viktig å gjøre forskjellen mellom disse tydelig, eksempelvis ved å gjøre de fysiske elementene mer tredimensjonale eller bruke andre materialer enn papir og papp.
- **Superrealisme:** Det vil ikke være intuitivt for testbrukere å bruke interaksjonsmønstre på GUI-elementer med superrealistiske egenskaper, eksempelvis skalering. Det vil derfor være nødvendig med en innføring i disse interaksjonsmønstrene på prototypen før testen begynner.

FS3: Hvilke anbefalinger kan vi komme med til prosessen rundt brukersentrert utvikling for Surface i de tidlige fasene av systemutviklingsløpet?

FS1 og FS2 har dannet grunnlaget for å kunne gi råd til hvordan man skal kunne håndtere utfordringene knyttet til brukersentrert utvikling for Surface. Våre viktigste anbefalinger er:

- Demonstrer egenskapene ved Surface for brukerutvalget gjennom hands-on interaksjon med abstrakte og domenespesifikke formidlingsdemoer.

- Utfør konvensjonell papirprototyping, men i tillegg bruk materialene postit-lim, sandpapir og transparente ark.
- Tydeliggjør skillet mellom fysiske og virtuelle elementer i prototypen.
- Wizard-of-Oz-testing av flere brukere simultant anbefales, men krever god planlegging.
- Interaksjon med flere brukere på samme prototype er komplekst og gode video- og lydopptak er desto mer viktig.

Våre anbefalinger bidrar til forskningsmiljøet ved å svare på en etterspørsel etter konkrete metoder for mer aktiv involvering av brukere i utviklingsprosessen for multi-touchbord. Resultatene bygger videre på Klingsheim og Raaes [56] forslag til videre forskning på introdusering av ny teknologi, Derboven et al. [18] sin etterlysning om low-fi prototypingteknikker for multi-touchbord og Sohans [68] problemstillinger rundt flerbrukertesting (se Seksjon 1.2).

16.2 Videre forskning

For å øke anbefalingenes validitet vil det være nødvendig å gjøre videre eksperimentering på et større omfang mennesker i andre kontekster. Det vil være interessant å se om våre anbefalinger bedrer kvaliteten på Surface-applikasjoner i form av økt brukskvalitet og avdekking av brukbarhetsfeil på et tidligere stadium.

Det hadde dessuten vært hensiktsmessig å gjøre isolert forskning på eventuelle fordeler og ulemper ved brukbarhetstesting av en implementert prototype kontra en papirprototype av en Surface-applikasjon. Da ville det være aktuelt å gjennomføre enbrukertesting for å utelukke påvirkningsvariabelen flerbrukertesting medfører.

For å undersøke verdien av å kjøre flerbrukertester i forhold til enbrukertester ville det også vært verdifullt å isolere slike eksperimenter. Ved å teste en implementert prototype på noen testbrukere individuelt og noen i par vil man kunne avsløre hvorvidt antallet samtidige brukere har innvirkning på mengden og typen brukbarhetsfeil som oppdages. Gjøres dette på en papirprototype vil variabelen som wizarden(e) utgjør kunne påvirke resultatene.

16.2.1 Refleksjoner

Gjennom arbeidet med denne oppgaven har vi møtt utfordringer hvor vi har sett nytten av det å være to. Spesielt i eksperimenteringen av prototypingsteknikkene og gjennomføringen av flerbrukbarhetstester utforsket vi teknikker som ikke ville vært mulig å gjennomføre med bare én person. I eksperimenteringen av prototypingsteknikkene var det dessuten behov for å ha en som kunne ha wizardens rolle og en som kunne ha testpersonens rolle. For å gjennomføre wizard-of-Oz flerbrukertesting var det nødvendig å ha mer enn én wizard tilgjengelig for at testpersonene skulle kunne utføre parallelle handlinger.

For videre forskning kreves det at det involveres mer enn én person for å teste ut våre anbefalinger. Når det gjelder å inkludere en muldvarp samtidig som man praktiserer flerbrukertesting er man nødt til å ha tre eller flere personer i teststaben; én muldvarp og minst to dedikerte wizards for testpersonene.

Del VI

Tillegg

Tillegg A

Dokumenter Brukt i FS1

A.1 Deltakelsesskriv

Brukersentrert utvikling av programmer for Microsoft Surface

Workshop 8. desember 2010

Prosjektet

Prosjektet er et samarbeid mellom NTNU og Vitenskapsmuseet, der det forskes på brukersentrert utvikling til Microsoft Surface. Microsoft Surface er et såkalt multi-touchbord, som vil si en PC der personer kan ta direkte på pc-skjermen uten å bruke datamus eller tastatur.

Ved brukersentrert utvikling involveres brukerne under utviklingen av programmer. En utfordring med Microsoft Surface er at teknologien kan være fremmed for brukerne. Derfor må utviklere formidle mulighetene og begrensningene ved dette bordet på en forståelig måte slik at brukerne kan ta del i utviklingen sammen med utviklerne.

Hensikt

Årsaken til at vi ønsker å gjennomføre en workshop er at den vil gi oss nyttig informasjon om hvordan vi som utviklere best bør presentere teknologien overfor representative brukere. Det er ikke dere som personer vi ønsker å teste, men derimot hvordan vår fremleggelse av teknologien er med på å gi dere inspirasjon til hva Microsoft Surface kan brukes til.

Konfidensialitet

Det vil bli tatt video- og lydopptak av hele workshopen. Dette gjøres for at vi skal kunne vurdere sesjonene i etterkant og for å sikre oss at vi har forstått deres utsagn og handlinger riktig. Vi vil sørge for at materiale vil bli anonymisert slik at det ikke vil være mulig å føre opplysningene tilbake til enkeltpersonene som deltar i prosjektet. Dette innebærer at informasjon som blir formidlet til offentligheten ikke vil kunne settes i sammenheng med den enkelte. Det er kun de involverte i prosjektet som vil kunne se opptakene i ettertid.



Microsoft Surface

Lars Skjelbek
E-post: larsloen@stud.ntnu.no
Tlf: 480 77 580

Kristin Haga
E-post: krihaga@stud.ntnu.no
Tlf: 936 23 763

A.2 Samtykkeerklæring

Brukersentrert utvikling av programmer for Microsoft Surface.

Deltakelse på workshop.

Samtykkeerklæring

Jeg har mottatt skriftlig og muntlig informasjon om studien, og fått anledning til å stille spørsmål. Jeg er klar over at det er frivillig å delta, og at jeg kan trekke meg fra studien når som helst uten å oppgi noen grunn.

Jeg samtykker i å delta i studien.

Trondheim, _____

Underskrift

A.3 Spørreundersøkelse før workshop

Spørreundersøkelse før workshopen

Kjønn: Gutt Jente

Alder: _____

1. Har du hørt om Microsoft Surface tidligere?

Ja Nei

2. Hvis ja, har du prøvd et eksemplar?

Ja Nei

3. Eier du en enhet med multi-touch (f.eks. iPhone, HTC Desire, Samsung Galaxy S, iPod Touch, iPad, Galaxy Tab)?

Ja Nei

4. Hvis nei, har du prøvd en?

Ja Nei

5. Har du besøkt Vitenskapsmuseet tidligere?

Ja Nei

6. Hvis ja, når besøkte du museet sist?

Svar: _____

A.4 Spørreundersøkelse etter workshop

Spørreundersøkelse etter workshopen

Hvor enig er du i påstandene under (sett kryss)?

1. Demoene ga meg et tydelig bilde på hva som er mulig med Surface.

Sterkt uenig					Sterkt enig
	1	2	3	4	5

2. Det var lett å komme opp med ideer til hva Surface kan brukes til på Vitenskapsmuseet.

Sterkt uenig					Sterkt enig
	1	2	3	4	5

3. Omvisningen på museet ga meg god inspirasjon til idédugnaden.

Sterkt uenig					Sterkt enig
	1	2	3	4	5

4. Demoene ga meg god inspirasjon til idédugnaden.

Sterkt uenig					Sterkt enig
	1	2	3	4	5

5. Demoene var lette å forstå**a) Generell multi-touch**

Sterkt uenig Sterkt enig

1	2	3	4	5

b) Interaksjon med fysiske objekter

Sterkt uenig Sterkt enig

1	2	3	4	5

c) Samhandling

Sterkt uenig Sterkt enig

1	2	3	4	5

d) Kombinasjon med andre teknologier

Sterkt uenig Sterkt enig

1	2	3	4	5

6. De ulike demoene ga meg god inspirasjon**a) Generell multi-touch**

Sterkt uenig Sterkt enig

1	2	3	4	5

b) Interaksjon med fysiske objekter

Sterkt uenig Sterkt enig

1	2	3	4	5

c) Samhandling

Sterkt uenig Sterkt enig

1	2	3	4	5

d) Kombinasjon med andre teknologier

Sterkt uenig Sterkt enig

1	2	3	4	5

A.5 Idékort

Gruppe: B

Hvem: _____

Hensikt: _____

Hva går den ut på:

A.6 Temaliste for gruppediskusjon

Temaer for gruppediskusjon, workshop:

1. Hvordan føler dere at workshopen har vært?
2. Føler dere at det var vanskelig å komme på ideer til hva Surface kan brukes til på Vitenskapsmuseet?
3. Hjalp det å få en omvisning på museet for å komme på ideer?
4. Hjalp det å få en demonstrasjon av Surface for å komme på ideer?
5. For abstrakt workshop: tror dere det ville vært lettere å komme på ideer hvis dere hadde fått sett eksempler på hva Surface kan brukes til på museet?
For domenespesifikk: tror dere det ville vært lettere å komme på ideer hvis dere hadde fått sett eksempler som ikke hadde så mye med museet å gjøre?
6. Føler dere at dere skjønnte hva Surface kan brukes til?
7. Var det noe dere likte spesielt godt ved demoene?
8. Var det noe dere likte spesielt dårlig ved demoene?

A.7 Resultater sp.undersøkelse etter workshop

1. Demoene ga meg et tydelig bilde på hva som er mulig med Surface.

	1	2	3	4	5
Abstrakt				5	1
Konkret				4	2

2. Det var lett å komme opp med ideer til hva Surface kan brukes til på Vitenskapsmuseet

	1	2	3	4	5
Abstrakt		4	1		1
Konkret		3	2		1

3. Omvisningen på museet ga meg god inspirasjon til idèdugnaden

	1	2	3	4	5
Abstrakt			1	4	1
Konkret			2	1	3

4. Demoene ga meg god inspirasjon til idèdugnaden

	1	2	3	4	5
Abstrakt		1	3	1	1
Konkret			2	2	2

5. Demoene var lette å forstå

a) Generell multi-touch

	1	2	3	4	5
Abstrakt					6
Konkret					6

b) Interaksjon med fysiske objekter

	1	2	3	4	5
Abstrakt					6
Konkret				1	5

c) Samhandling

	1	2	3	4	5
Abstrakt				1	5
Konkret					6

d) Kombinasjon med andre teknologier

	1	2	3	4	5
Abstrakt				1	5
Konkret				2	4

6. Demoene ga meg god inspirasjon**a) Generell multi-touch**

	1	2	3	4	5
Abstrakt				4	2
Konkret			2	2	2

b) Interaksjon med fysiske objekter

	1	2	3	4	5
Abstrakt			2	4	
Konkret			2	2	2

c) Samhandling

	1	2	3	4	5
Abstrakt		3	1	1	1
Konkret			1	3	2

d) Kombinasjon med andre teknologier

	1	2	3	4	5
Abstrakt			2	2	2
Konkret			1	2	3

Tillegg B

Scenarior for MultiViten

B.1 Scenario 1



NÅR HUN SNUR KORTET OPPDAGER HUN EN SLASS STREKKODE SAMMEN MED TEKSTEN "TA MED DETTE KORTET TIL MULTIMEDIAORDET OG PLOSSER DET PÅ SKJERMEN MED DENNE SIDEN NED."

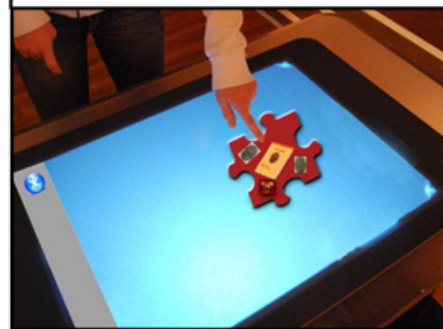


TRINE FINNER FREM TIL MULTIMEDIAORDET...

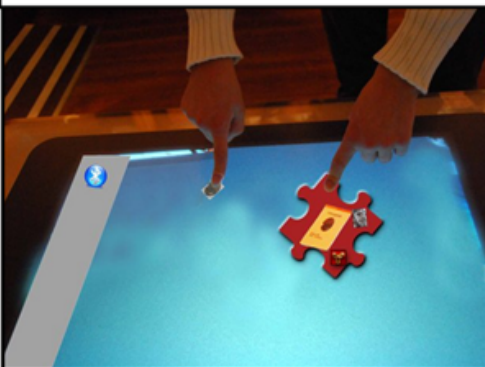
...HVOR HUN FORSØKER Å LEGGE NED KORTET.



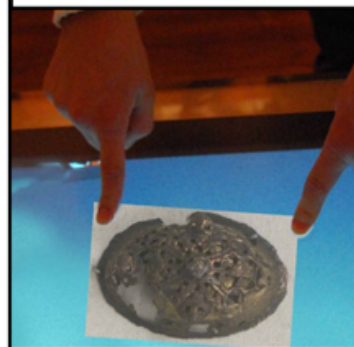
DET DUKKER OPP EN PUSLESPILLBRIKKE MED ULIKE ELEMENTER MED RELASJON TIL VIKINGSKÅLEN.



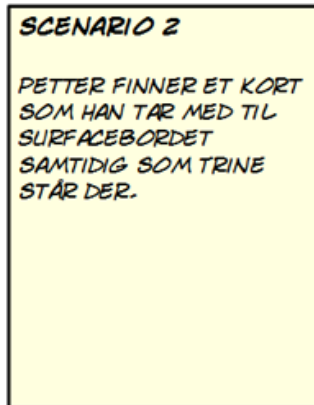
HUN DRAR ET BILDE UT FRA BRIKKEN...



...OG ROTERER OG SKALERER BILDET.



B.2 Scenario 2





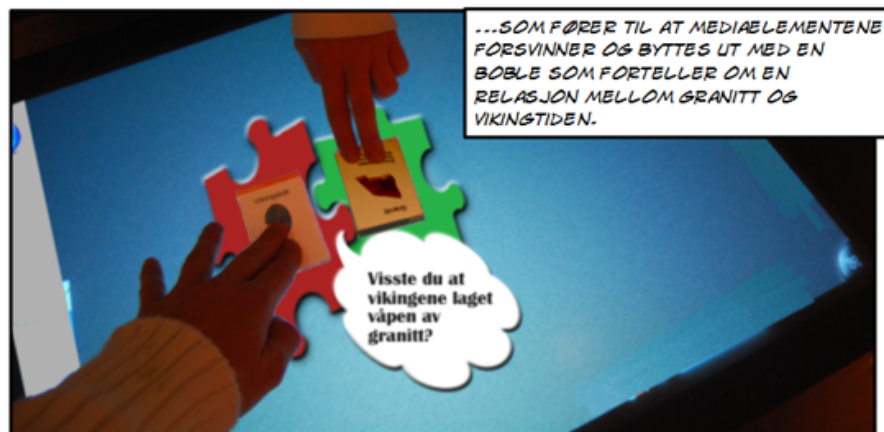
HAN LESER SÅ BAKSIDEN AV KORTET



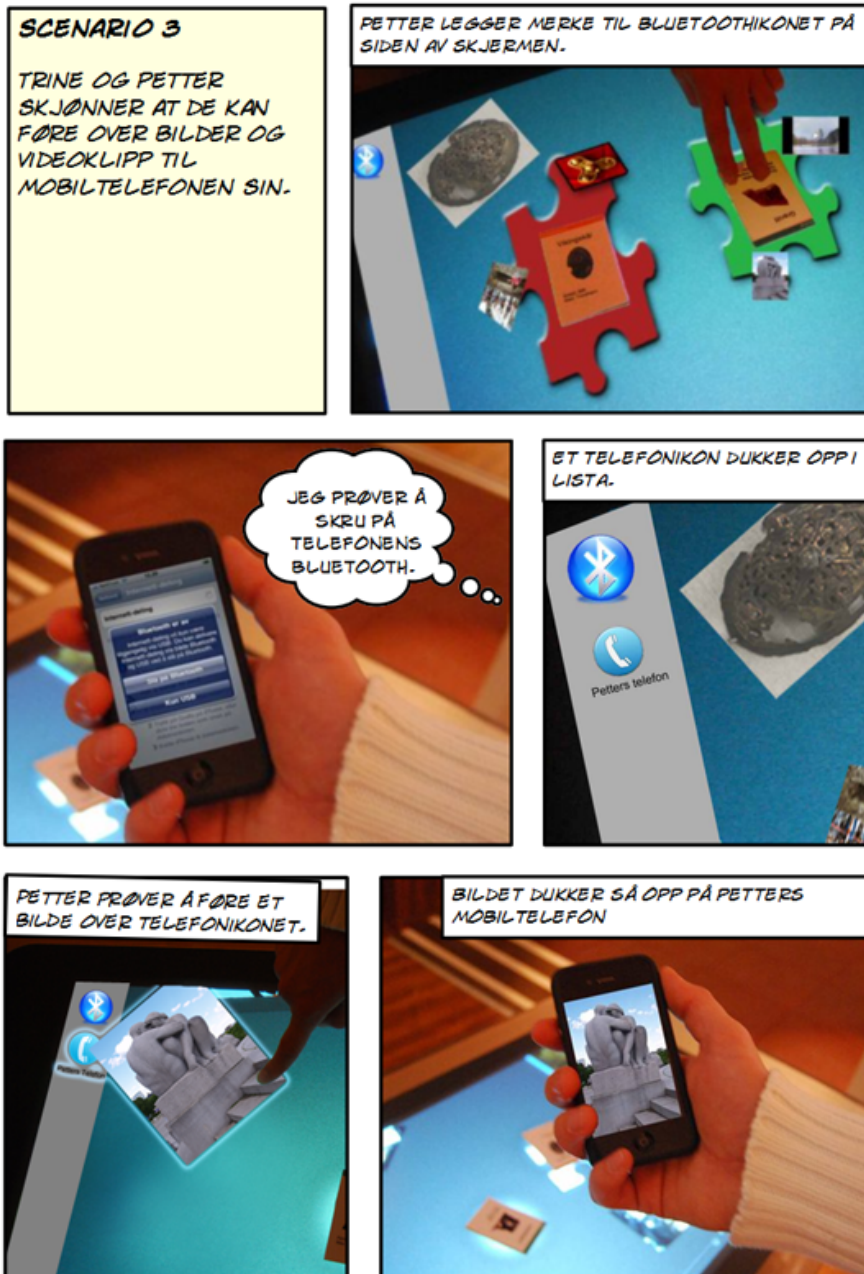
PETTER GÅR NED TIL SURFACE-BORDET SOM HAN TIDLIGERE HAR SETT I UNDERETASJEN.



UNDER KORTET DUKKER DET
OPP EN GRØNN
PUSLESPILLBRIKKE MED
MEDIAELEMENTER SOM
RELATERES TIL GRANITT.



B.3 Scenario 3



Tillegg C

Dokumenter Bruk i FS2

C.1 Samtykkeerklæring for brukbarhetstest

Brukर्सentrert utvikling av programmer for Microsoft Surface.

Deltakelse på brukbarhetstest.

Samtykkeerklæring

Jeg har mottatt informasjon om studien, og fått anledning til å stille spørsmål. Jeg er klar over at det er frivillig å delta, og at jeg kan trekke meg fra studien når som helst uten å oppgi noen grunn.

Det vil bli tatt video- og lydopptak av brukbarhetstesten. Dette gjøres for at vi skal kunne vurdere sesjonene i etterkant og for å sikre oss at vi har forstått deres utsagn og handlinger riktig. Vi vil sørge for at materiale vil bli anonymisert slik at det ikke vil være mulig å føre opplysningene tilbake til enkeltpersonene som deltar i prosjektet. Dette innebærer at informasjon som blir formidlet til offentligheten ikke vil kunne settes i sammenheng med den enkelte. Det er kun de involverte i prosjektet som vil kunne se opptakene i ettertid.

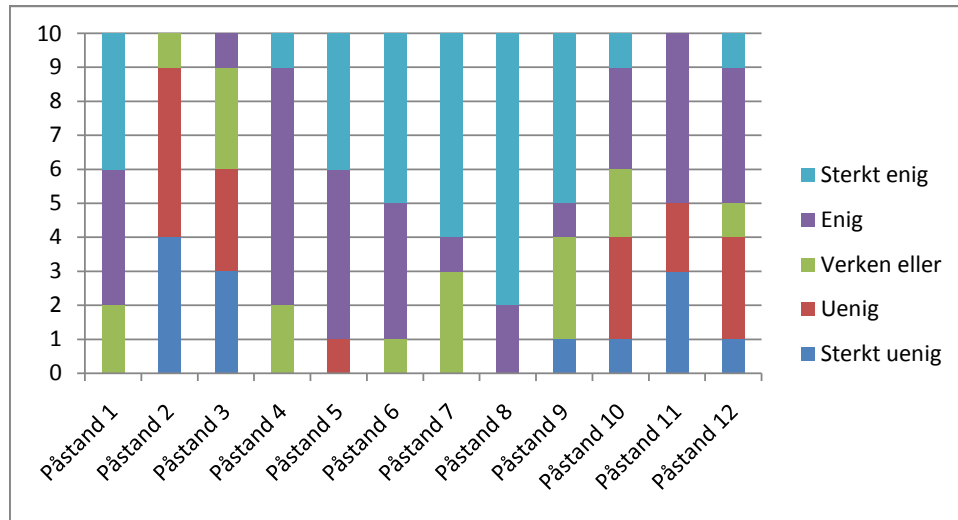
Vi forbeholder oss retten til å bruke bilder fra brukbarhetstesten i oppgaven vår.

Jeg samtykker i å delta i studien.

Trondheim, _____

Underskrift

C.3 Resultater spørreundersøkelse om brukbarhet

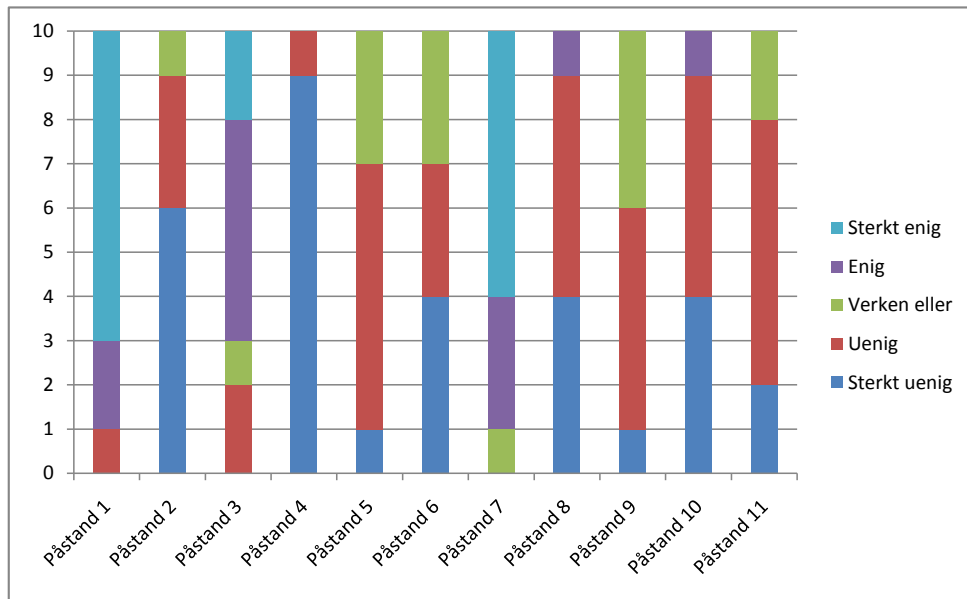


C.4 Spørreundersøkelse om testopplegget

Spørsmål angående brukertesten

- | | Sterkt uenig | Sterkt enig |
|---|--|--|
| 1. Prototypen klarte å simulere multi-touch. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| 2. Det var vanskelig å forstå hva de ulike elementene på bordet og i rommet skulle representere. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| 3. Det var enkelt å forstå at man kunne rotere, skalere (zooome inn/ut) og flytte på mediaelementene på puslespillbrikkene. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| 4. Det var vanskelig å forstå at det dukket opp en virtuell puslespillbrikke på "Surface-skjermen" når jeg la ned et fysisk museumskort. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| 5. Det var hele tiden lett å skille mellom hva som var fysiske objekter og hva som var virtuelle objekter på Surface-bordet. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| 6. Jeg forstod lett at det egentlig ikke skal være mulig å flytte på puslespillbrikken uten å bevege museumskortet som lå oppå. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| 7. Jeg forstod lett at dersom jeg flyttet et mediaelement vekk fra puslespillbrikken, så ville den stå i ro hvis jeg deretter flyttet museumskortet og puslespillbrikken. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| 8. Det var vanskelig å forstå at puslespillbrikken og alle mediaelementene som var tilknyttet denne ville forsvinne igjen hvis jeg fjernet kortet fra bordet. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| 9. Jeg syntes det var forstyrrende at det var en annen testbruker som brukte prototypen samtidig. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| 10. Jeg følte at vi brukte skjermen "på tur", dvs. jeg ventet til den andre gjorde seg ferdig og omvendt. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| 11. Det at det gikk veldig treigt under brukertesten ødela mye for opplevelsen av systemet. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |

C.5 Resultater spørreundersøkelse om testopplegget



Tillegg D

Øvrige Tillegg

D.1 Samtale med Einar Ingebretsen

Notater fra samtale med Einar Ingebretsen, 16.11.10

utvikler av applikasjon for "De ukjente" på NRK.

Einar har jobbet tett med Microsoft Norge, og det var gjennom dem han fikk forespørselen om å utvikle applikasjonen fra NRK.

NRK var opprinnelig i kontakt med Microsoft om noe helt annet. Planen til NRK var å bruke en mer "oldschool" løsning i programmet "De ukjente".

Et problem de støtته på var tilgangen på bord. Det var vanskelig å få tak i to surface bord, spesielt ettersom det bare var en måned igjen før filmingen av programmet skulle starte. I starten fokuserte de dermed på å lage en applikasjon til "HP-all in one", en løsning som baserte seg på en kombinasjon av PC, monitor og multi-touch.

Utviklingsprosessen ble noe todelt. Ettersom det var dårlig med tid fikk Einar mye frihet når det gjaldt utviklingen. Det forelå noen nokså enkle basiskrav som å skalere og flytte på bilder. Når det gjaldt applikasjonens utseende var det to grafikere fra NRK som hadde ansvaret for dette.

En av de største utfordringene med å få testet dette på brukere var tilgangen på bord, siden det kun var to tilgjengelige bord totalt og Einar måtte ha det ene for å utvikle på. For å fremstille en skisse av konseptet ble det brukt noe så enkelt som Excel.

Einar laget en filmsnutt av applikasjonen sin og sendte denne inn til NRK for å få feedback. NRK er veldig flinke til å samarbeide som et team, så her var mange av de som er involvert i programmet med (også programleder).

Applikasjonen som vises på tv ser ut til å være ganske enkel, men i bakgrunnen skjuler det seg også to andre applikasjoner som skulle utvikles. Blant annet en administrasjons-app - for å legge opp løsningen, score, navn på laget, bilder og hint (sistnevnte ble aldri brukt). I tillegg en monitor til programlederen som skulle fungere som en hjelpeapplikasjon.

NRK var veldig fornøyde med produktet, og ønsker mer. Ettersom oppetiden på et slikt system er særdeles viktig ble dette utviklet testdrevet. Dette førte med seg at NRK ikke hadde noe problem med applikasjonen i noen av programmene.

I senere tid har det vært lite oppfølging. Det har vært et par småting, og noen bugs som har blitt fikset. Ellers ytret NRK ønske om noen justeringer som de ønsket å gjøre, men som Einar var nødt til å fraråde de, da det kunne by på problemer å endre på kodebasen. Han bekrefter at han ikke har skrevet en kodelinje siden pilottesten.

Det ble også nevnt at NRK i Trondheim har benyttet seg av Microsoft Surface i en applikasjon for spillreviews.

Bibliografi

- [1] ISO 9241-210. Ergonomics of human-system interaction — part 210: Human-centred design for interactive systems, 2010.
- [2] Chadia Abras, Diane Maloney-Krichmar, and Jenny Price. User-centered design. *Encyclopedia of Human-Computer Interaction*, 2004.
- [3] Scientific American. How it works: Multi-touch surfaces explained. Hentet 11. mai 2011 fra <http://www.scientificamerican.com/article.cfm?id=how-it-works-touch-surfaces-explained>, 2008.
- [4] Daryl J. Bem. Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, 74(3):183–200, 1967.
- [5] Jeanette Blomberg and Finn Kensing. The integration of computing and routine work. *Computer Supported Cooperative Work*, 4(3):167–185, 1998.
- [6] John Brooke. Sus: A quick and dirty usability scale. In P. W. Jordan, B. Weerdmeester, A. Thomas, and I. L. Mclelland, editors, *Usability evaluation in industry*. Taylor and Francis, London, 1996.
- [7] Bill Buxton. *Sketching User Experiences: Getting the Design Right and the Right Design*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2007.
- [8] Bill Buxton. Multi-touch systems that i have known and loved. Hentet 24. januar 2011 fra <http://www.billbuxton.com/multitouchOverview.html>, 2010.
- [9] Youtube camillesg. Baby masters the iphone interface. Hentet 26. januar 2011 fra http://youtu.be/4-1JP_4VoeY, 2008.

- [10] Erran Carmel, Randall D. Whitaker, and Joey F. George. Pd and joint application design: a transatlantic comparison. *Communications of the ACM*, 36(4):40–48, 1993.
- [11] Youtube ChadTheTube. iphone used by 1 year old baby. Hentet 26. januar 2011 fra <http://youtu.be/oZwKPDvYA2M>, 2007.
- [12] Jennifer Colegrove. The state of the touch-screen market in 2010. Hentet 24. januar 2011 fra <http://bit.ly/iwmqnR>, 2010.
- [13] Hewlett-Packard Development Company. Hp touchpad. Hentet 11. mai 2011 fra <http://www.palm.com/us/products/pads/touchpad/index.html>, 2011.
- [14] Microsoft Corporation. Microsoft surface user experience guidelines, 2009.
- [15] Microsoft Corporation. Windows speech recognition. Hentet 26. januar 2011 fra <http://www.microsoft.com/enable/products/windowsvista/speech.aspx>, 2011.
- [16] Mary Jo Davidson, Laura Dove, and Julie Weltz. Mental models and usability. Hentet 29. oktober 2010 fra <http://www.lauradove.info/reports/mental%20models.htm>, 1999.
- [17] Stewe Dawson. Lonely planet on microsoft surface. Hentet 26. november 2010 fra <http://emergingexperiences.com/2009/09/lonely-plant-on-microsoft-surface/>, 2009.
- [18] Jan Derboven, Dries De Roeck, Mathijs Verstraete, David Geerts, Jan Schneider-Barnes, and Kris Luyten. Comparing user interaction with low and high fidelity prototypes of tabletop surfaces. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, NordiCHI '10, pages 148–157, New York, NY, USA, 2010. ACM.
- [19] Information & Design. Participatory design. Hentet 26. januar 2011 fra <http://www.infodesign.com.au/usabilityresources/participatorydesign>, 2011.
- [20] Thomas T. Hewett et. al. *ACM SIGCHI curricula for human-computer interaction*. Association for Computing Machinery, Inc., Hentet 21. oktober 2010 fra <http://old.sigchi.org/cdg/index.html>, 1996.

- [21] Evoluce. Evoluce one multi-touch lcd screen. Hentet 24. januar 2011 fra http://www.evoluce.com/en/products/multi-touch_lcd_screen.php, 2011.
- [22] Kraig Finstad. The system usability scale and non-native english speakers. *Journal of Usability Studies*, 1(4):185–188, 2006.
- [23] Grete Fischer and Nils Sortland. *Innføring av sosialpsykologi*. Universitetsforlaget, 2001.
- [24] George W. Fitzmaurice, Hiroshi Ishii, and William A. S. Buxton. Bricks: laying the foundations for graspable user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '95, pages 442–449, New York, NY, USA, 1995. ACM Press/Addison-Wesley Publishing Co.
- [25] Therese Fjellgård. Hvordan ser året ut i hodet ditt? Bloggpost hentet 29. oktober 2010 fra http://sordid.blogg.no/1231322738_hvordan_ser_ret_ut_i_.html, 2009.
- [26] Suzanne Ginsburg. *Designing the iPhone User Experience - A User-Centered Approach to Sketching and Prototyping iPhone Apps*. Addison-Wesley, 2010.
- [27] Sigmund Grønmo. *Samfunnsvitenskapelige metoder*. Bergen : Fagbokforlaget, 2004.
- [28] NUI Group. Nui group community faqs. Hentet 26. januar 2011 fra <http://nuigroup.com/faq/>, 2009.
- [29] Eric Havir. Swedish medical center baby photo contest. Hentet 26. november 2010 fra <http://blogs.msdn.com/b/surface/archive/2010/05/07/swedish-medical-center-baby-photo-contest.aspx>, 2010.
- [30] Rudy A. Hirschheim. User participation in practice: experiences whit participative systems design. *Participation in systems development*, pages 194 –204, 1989.
- [31] Steve Hodges, Shahram Izadi, Alex Butler, Alban Rrustemi, and Bill Buxton. Thinsight: Versatile multi-touch sensing for thin form-factor displays. *Proceedings of the 20th annual ACM symposium on User interface software and technology*, 2007.

- [32] Stephanie Houde and Charles Hill. What do prototypes prototype? In *Handbook of Human-Computer Interaction (2nd Ed.)*. Elsevier Science B.V.: Amsterdam, 1997.
- [33] Ideum. Mt-50 multitouch table. Hentet 24. januar 2011 fra <http://www.ideum.com/products/multitouch/>, 2011.
- [34] Apple Inc. Apple reinvents the phone with iphone. Hentet 24. januar 2011 fra <http://www.apple.com/pr/library/2007/01/09iphone.html>, 2007.
- [35] Apple Inc. Apple launches ipad. Hentet 25. januar 2011 fra <http://www.apple.com/pr/library/2010/01/27ipad.html>, 2010.
- [36] Apple Inc. Apple ipad. Hentet 11. mai 2011 fra <http://www.apple.com/ipad/>, 2011.
- [37] Apple Inc. Apple reports first quarter results. Hentet 24. januar 2011 fra <http://www.apple.com/pr/library/2011/01/18results.html>, 2011.
- [38] Apple Inc. Apps for iphone. Hentet 24. januar 2011 fra <http://www.apple.com/no/iphone/apps-for-iphone/>, 2011.
- [39] Motorola Mobility Inc. Motorola xoom. Hentet 11. mai 2011 fra <http://www.motorola.com/staticfiles/Consumers/xoom-android-tablet/us-en/overview.html>, 2011.
- [40] Heinz K. Klein and Michael D. Myers. A set of principles for conducting and evaluating interpretive field studies in information systems, 1999.
- [41] Microsoft Developer Network (MSDN). Microsoft surface bluetooth connect code sample. Hentet 26. november 2010 fra <http://code.msdn.microsoft.com/surfacebluetooth>, 2010.
- [42] Microsoft Developer Network (MSDN). Byte tags. Hentet 11. mai 2011 fra [http://msdn.microsoft.com/en-us/library/ee804885\(v=surface.10\).aspx](http://msdn.microsoft.com/en-us/library/ee804885(v=surface.10).aspx), 2011.
- [43] Enid Mumford. Participation – from aristotle to today. *Beyond productivity: information system development for organizational effectiveness*, pages 95–104, 1984.
- [44] Michael Myers. Investigating information systems with ethnographic research. *Commun. AIS*, 2, December 1999.

- [45] Michael D. Myers. Qualitative research in information systems. *MIS Q.*, 21:241–242, June 1997.
- [46] Jakob Nielsen. The usability engineering life cycle. *Computer*, 25(3):12–22, mar. 1992.
- [47] Jakob Nielsen. *Usability Engineering*. Academic Press, Inc, 1993.
- [48] Jakob Nielsen. Why you only need to test with 5 users. Hentet 5. november 2010 fra <http://www.useit.com/alertbox/20000319.html>, 2000.
- [49] Donald A. Norman. *The psychology of everyday things*. New York : Basic Books, 1988.
- [50] Donald A. Norman. *The design of everyday things*. New York : Doubleday/Currency, 1990.
- [51] Espen Solberg Nygård. Multi-touch interaction with gesture recognition. Master's thesis, NTNU, Norges teknisk-naturvitenskapelige universitet, 2010.
- [52] Briony J. Oates. *Researching Information Systems and Computing*. SAGE Publications Ltd, 2006.
- [53] University of Minnesota. Human-computer interaction. Hentet 23. mai 2011 fra <https://wiki.umn.edu/InterfaceDesign/Human-computerInteraction>, 2007.
- [54] Anne-Marie Oostveen and Peter van den Besselaar. From small scale to large scale user participation: a case study of participatory design in e-government systems. pages 173–182, New York, NY, USA, 2004. ACM.
- [55] Robert A. Peterson. *Construction effective questionnaires*. SAGE Publications, Inc, 2000.
- [56] Benedicte Raae and Tuva Foldøy Klingsheim. Introducing new technologies to users in user-centered design projects: An experimental study. Master's thesis, NTNU, Norges teknisk-naturvitenskapelige universitet, 2009.
- [57] Amnesia Razorfish. Staffing directory on microsoft surface. Hentet 26. november 2010 fra <http://www.microsoft.com/showcase/en/us/details/a95bdae6-be36-490b-8797-c5c2014523fe>, 2009.

- [58] Amnesia Razorfish. Table toss, a microsoft surface game. Hentet 26. november 2010 fra <http://www.microsoft.com/showcase/en/us/details/1a8de479-0192-4f59-a6d0-8c2ef7433c47>, 2009.
- [59] Jeremy Reimer. A history of the gui. Hentet 26. januar 2011 fra <http://arstechnica.com/old/content/2005/05/gui.ars/>, 2005.
- [60] Youtube rluecke3. Microsoft “surface” - the possibilities. Hentet 1. mars 2011 fra <http://youtu.be/6VfpVYYQzHs>, 2007.
- [61] Colin Robson. *Real world research: A resource for social scientists and practionerresearch*. Mass: Addison-Wesley, 2002.
- [62] Y. Rogers, H. Sharp, and J. Preece. *Interaction Design: Beyond Human-Computer Interaction*. John Wiley and Sons Ltd, 2002.
- [63] Bjørn H. Samset. Mitt år er en sirkel mot klokka, hvordan er ditt? Bloggpost hentet 29. oktober 2010 fra <http://www.forskning.no/blog/b.h.samset@fys.uio.no/240124>, 2010.
- [64] Ulrike Schultze. Reflexive ethnography in information systems research. Fra boken *Qualitative Research in IS: Issues and Trends* av Eileen M. Trauth (Pennsylvania State University, USA), 2001.
- [65] G. C. Smith. The marble answering machine. In *The Hand That Rocks the Cradle*, pages 60–65, 1995.
- [66] Serengul Smith-Atakan. *Human-Computer Interaction*. Thomson Learning, London, 2006.
- [67] Carolyn Snyder. *Paper Prototyping: The Fast and Easy Way to Design and Refine User Interfaces*. Morgan Kaufmann, 2004.
- [68] S. M. Sohan. *Problems with Testing Multi-User Multi-Touch Table-top Applications*. Hentet 22. februar fra <http://www.smsohan.com/publications>, 2011.
- [69] Norges teknisk-naturvitenskapelige universitet. Ntnu vitenskapsmuseet. Hentet 15. desember 2010 fra <http://www.ntnu.no/vitenskapsmuseet/>, 2010.
- [70] Eli Toftøy-Andersen and Jon Gunnar Wold. *Praktisk brukertesting*. Cap-pelen Damm akademisk, 2011.
- [71] Bruce Tognazzini. *Tog on Interface*. Mass: Addison-Wesley, 1991.

- [72] TUIO.org. Tuio. Hentet 18. mai 2011 fra <http://www.tuio.org/>, 2011.
- [73] Stanford University. Stanford encyclopedia of philosophy. Hentet 26. november 2010 fra <http://plato.stanford.edu/entries/abstract-objects/>, 2010.
- [74] Craig Villamor, Dan Willis, and Luke Wroblewski. Touch gesture reference guide. Hentet 26. januar 2011 fra <http://www.lukew.com/ff/entry.asp?1071>, 2010.
- [75] Pierre Wellner. The digitaldesk calculator: Tactile manipulation on a desk top display. In *In Proceedings of UIST'92, the ACM Symposium on User Interface Software and Technology*. (Nov, pages 27–33, 1991.
- [76] Wikipedia. Pong. Hentet 10. desember 2010 fra <http://en.wikipedia.org/wiki/Pong>, 2010.
- [77] Wikipedia. Command-line interface. Hentet 26. januar 2011 fra http://en.wikipedia.org/wiki/Command-line_interface, 2011.
- [78] Wikipedia. Likert scale. Hentet 16. mai 2011 fra http://en.wikipedia.org/wiki/Likert_scale, 2011.
- [79] Wikipedia. Microsoft surface. Hentet 11. mai 2011 fra http://en.wikipedia.org/wiki/Microsoft_Surface, 2011.
- [80] Wikipedia. Multi-touch. Hentet 10. mai 2011 fra <http://en.wikipedia.org/wiki/Multi-touch>, 2011.
- [81] Wikipedia. Tangible user interface. Hentet 18. mai 2011 fra http://en.wikipedia.org/wiki/Tangible_User_Interface, 2011.
- [82] Wikipedia. Ubiquitous computing. Hentet 11. mai 2011 fra http://en.wikipedia.org/wiki/Ubiquitous_computing, 2011.
- [83] Xbox. Introducing kinect for xbox 360. Hentet 26. januar 2011 fra <http://www.xbox.com/Kinect/>, 2011.
- [84] Per Øystein Saksvik and Kjell Nytrø. *Ny personalpsykologi for et arbeidsliv i endring*. Cappelen Akademiske Forlag, 2006.
- [85] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. Research through design as a method for interaction design research in hci. In *Proceedings of the SIGCHI conference on Human factors in computing systems, CHI '07*, pages 493–502. ACM, 2007.