

Sluttrapport - Telenorprosjekt

Gruppe 2

Av

Siri Dølvik Sandnes (sirids)

Jesper Walberg (jesperbw)

Natalia Pineguina (natalpi)

Universitetet i Oslo

Høst 2011

Innholdsfortegnelse

1.0. Introduksjon	s. 2
2.0. Teori	s. 2-3
3.0. Metoder	s. 3-5
<i>3.1. Low-fidelity</i>	s. 3-4
<i>3.2. High-fidelity</i>	s. 4-5
4.0. Resultat	s. 5-7
<i>4.1. Between-group</i>	s. 5-6
<i>4.2. Two-tailed independent-samples t-test</i>	s. 6-7
5.0. Diskusjon	s. 7-8
6.0. Konklusjon	s. 8
Referanseliste	s. 9

Vedlegg 1	-	Prosjektplan
Vedlegg 2	-	Presentasjon til Telenor
Vedlegg 3	-	Low-fidelity prototype 1 (mockup) + 2 scenarioer
Vedlegg 4	-	Low-fidelity prototype 2 (mockup) + scenario
Vedlegg 5	-	Testresultater
Vedlegg 6	-	Bilder av deltakere

1.0. Introduksjon

Vi valgte oppgaven "Design av visuell selv-betjening på smartphones" og valgte med dette å lage en applikasjon til iPhone for Telenor. Vi idé-myldret over flere alternativer til mulig applikasjon og endte til slutt opp med å ville lage en applikasjon som skulle representere Telenors kundeservice, fra hjemmesiden på web, på en forenklet måte på iPhone og iPad. Vi skrev deretter prosjektplan (vedlegg 1) og forberedte en presentasjon om tankene og ideene våre samt vår visjon som vi skulle ta med i et møte med representanter for Telenor-gruppen (vedlegg 2).

Etter møte med Telenor satt vi igjen med mange gode tilbakemeldinger og forslag til nye løsninger som vi burde fokusere på. Telenor nevnte blant annet at vi burde fokusere på ett område som de tilbyr så ikke oppgaven skulle bli for stor. Med dette mente de videre at en faktura applikasjon kunne være passende og understreket nok en gang at vi burde konkretisere valgene og fokusere på én eller to deler av fakturatjenesten så dette ble overkommelig.

Vi fulgte rådet til Telenor og bestemte oss for applikasjonen "Telenor Faktura" og laget en low-fidelity prototype av denne ideen. Vi viste så prototypen til noen brukere hvor vi fikk tilbakemelding på designet, menyvalg og generelt om applikasjonen. Disse tilbakemeldingene tok vi med oss i den videre utviklingen av high-fidelity prototype. Siden vi ikke fikk laget noen prototype for en faktura applikasjon bestemte vi oss for å lage prototype av informasjon om Telenors tjenester og kundeservice. På denne måten kunne vi teste applikasjonen opp mot Telenors egen hjemmeside og måle tiden deltakere bruker på å finne svar på samme spørsmål ved hjelp av disse to tjenestene (applikasjon og hjemmeside).

2.0. Teori

Storyboarding, scenario og åpent intervju:

Fordeler: Lav kostnad, kort utviklingstid, gir indikasjon på hvordan sluttprodukt *kan* se ut, gir en historie om en mulig bruker og hans/hennes hendelsesforløp, åpner for dialog med bruker ved hjelp av åpent intervju.

Ulemper: Hvis man vil ha mer tilbakemelding på spesifikke ting fra bruker, er konkret med tanke på hendelsen som forekommer slik at bruker kanskje ikke får et helhetsinntrykk.

Experinental design - Between-group:

Fordeler: Tydeligere resultat, deltaker blir bare testet på et vilkår og det tar kortere tid, bedre kontroll av læringseffekten, mindre effekter av utmattelse og frustrasjoner.

Ulemper: Effekt fra individuelle forskjeller, vanskeligere å finne forskjeller mellom vilkår, krever større utvalgsstørrelse (flere deltakere).

3.0. Metoder

3.1. Low-fidelity

Vi valgte å bruke storyboarding og scenarier som low-fidelity prototyper. I motsetning til vanlig storyboarding der man ser produktet bli brukt i en setting valgte vi å lage en historie gjennom skjermbilder som man ser i figur 1. Brukerens miljø er ikke viktig i bruken av denne applikasjonen så vi valgte derfor å ha fokus på det som skjer på skjermen. Gjennom en serie på åtte skjermbilder ser man en oppgave bli utført som f. eks å utsette en faktura.



Figure 1 – Storyboard der skjermbildene forteller en historie

Denne historien av skjermbilder akkompagneres av et scenario som forklarer med ord handlingene som blir utført (vedlegg 4).

Etter å ha laget et scenario med tilhørende storyboard gjennomførte vi et åpent intervju med potensielle brukere. Først ble de bedt om å lese scenarioet før intervjueren viste dem storyboardet og forklarte handlingen som hører til hvert

skjerm bilde. Intervjueren og den potensielle brukeren hadde så en åpen samtale rundt storyboardet.

På bakgrunn av feedbacken fra de første intervjuene laget vi et nytt storyboard med tilhørende scenario og gjennomførte åpne intervjuer med brukere på samme måte som med den første prototypen.

Begge prototypene er utviklet med Apples guidelines for application design som utgangspunkt, men storyboard nummer én bygger mer på standarder innenfor iOS menysystemer. Begge prototypene bruker hovedmenyer som er like de man finner ellers i iOS og bruker dermed brukerens tidligere erfaringer for å hjelpe dem å manøvrere applikasjonen.

3.2. High Fidelity

Basert på feedback fra low-fidelity prototypene og Apples guidelines for Menu based applications laget vi en high-fidelity prototype for iPad som etterligner de tre hovedkategoriene på Telenor sin hjemmeside som har produkt og kundeinformasjon. Under kategoriene Mobil, Bredbånd og kundeservice la vi inn forskjellige underkategorier som hver hadde en side eller mer med informasjon på Telenor sine websider som man kan se på figur to som viser hovedmenyen og "Mistet PUK"-siden.

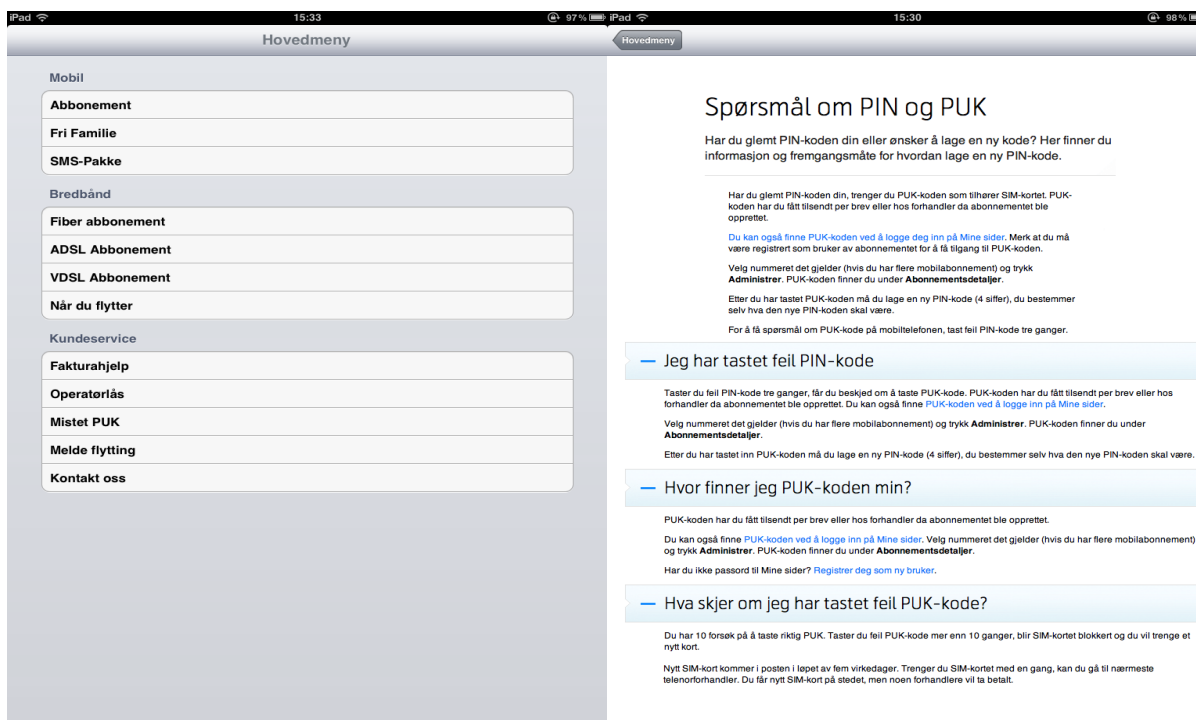


Figure 2 - Hovedmeny og siden med informasjon om PIN & PUK

Med det i baktankene at vi ønsket å teste om informasjonen Telenors kunder ønsket å få tak i var lettere tilgjengelig i appformat enn webformat, fra smarttelefoner og tablets, laget vi sidene med informasjon slik at de lignet mest mulig på Telenors websider. Dermed minimerer vi sjansen for å tilføre bias til resultatene gjennom forskjellig visuell presentasjon av informasjonen.

Som man kan se er fokuset på hva applikasjonen inneholder forskjellig fra low-fidelity prototypen til high-fidelity prototypen. Vi ønsket å teste funksjoner som var på toppen av listen til kundeservice over hva de får henvendelser om og den klare vinneren var utsette faktura. Dessverre fikk vi ikke tilgang til testbruker på Telenor sin 'Min Side'-funksjon og kunne derfor ikke teste funksjoner som var avhengig av å ha tilgang til Min side.

Siden vi ønsket å gjøre et between-group eksperiment for å sammenligne app med webside ble alternativet å teste de spørsmålene som ikke dreiet seg rundt funksjoner på min side.

4.0. Resultat

H0: Det er ingen forskjell i tiden det tar å finne informasjonen på app og websiden til Telenor.

H1: Det er en forskjell i tiden det tar å finne informasjonen på app og websiden til Telenor.

4.1. Between-group

Med stor fare for læring gjennom utføring av oppgavene valgte vi å gjøre et between-group eksperiment slik at halvparten av deltakerne tester appen og den andre halvparten tester websiden. Ulempene ved at man trenger flere testdeltakere og at forskjeller i individuelle ferdigheter kan påvirke resultatet er verdt det for å forhindre at testerne pådrar seg erfaring fra å utføre oppgavene. Siden kategoriene og innholdet på hver side er laget så like som mulig mener vi

testresultatene vil bli kraftig påvirket om samme tester skal utføre oppgavene på både web og app. Between-group design kan også gjøre det

For å kunne si noe om ferdighetene til deltakerne har vi spurt dem to spørsmål som er ment å måle deres egen oppfattelse av erfaringen de har med smarttelefoner og iPad/iPhone. Spørsmålene ble stilt før testen ble utført.

Hver deltager fikk et tilfeldig utdelt deltagernummer som korresponderte med en av de to testmåtene. Gjennom tilfeldig utdeling av deltagernummer forhindrer man bias gjennom at man bevisst eller ubevisst inndeler deltakere i testgrupper selv.

4.2. Two-tailed Independent-samples T test

Vi gjennomførte en Two sample independent t-test assuming unequal variances. Testen ble brukt fordi vi hadde en uavhengig variabel som forekom under to conditions, app og website.

Grunnen til at vi valgte assuming unequal variances er fordi vi ikke vet noe om variansen mellom de to gruppene og testgruppen er ikke stor nok til at vi kan anta at testvariansen kan erstattes med populasjonsvariansen.

Resultatene er normalfordelt som vi ser av grafene på figur tre.

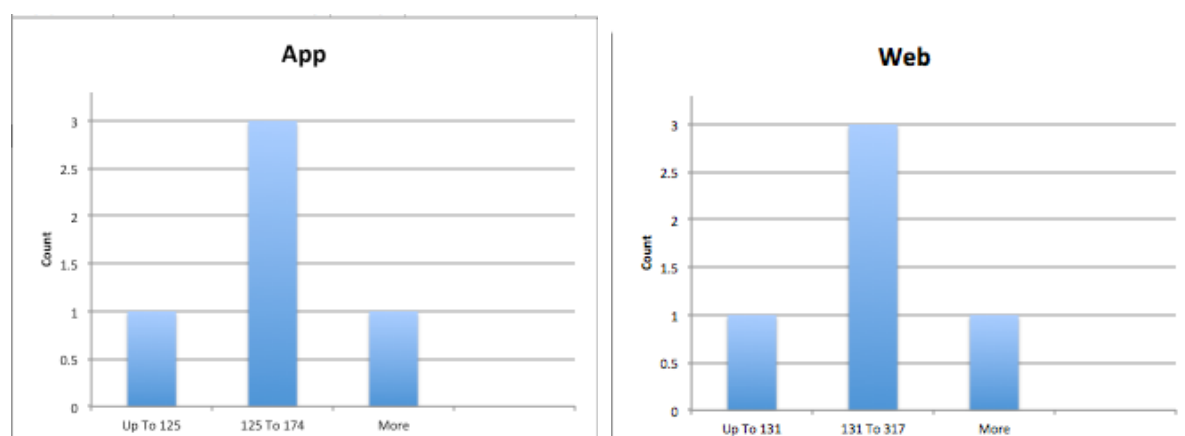


Figure 3 - Histogram viser normalfordeling

Med en t verdi på 1.76866 må den kritiske t verdien være mindre for at vi skal kunne avvise nullhypotesen vår. Den kritiske t verdien for two-tailed

distribution i vår test viste seg å være 2.57058 for 5% kritisk verdi og p-verdien var 0.13718.

På grunn av en p verdi som er høyere enn vår alphaverdi (5%) og en kritisk t verdi som er høyere enn t verdien kan vi fastslå at testen viser at det ikke er signifikant forskjell mellom gruppene i tiden brukt på å finne informasjonen og dermed kan vi ikke forkaste nullhypotesen ($t(5) = 1.76866, p > 0.05$).

5.0. Diskusjon

Det kan være flere faktorer som har påvirket de endelige testresultatene. *Å vite at man blir testet* er én av disse. Når man vet at man blir testet i noe, er det store sjanser for at atferden endres under testen. I vårt tilfelle, opplevde vi at flere brukere virket stresset av å bli tatt tid på, og dette kan bekreftes i større eller mindre grad av feil i besvarelsene på testoppgavene. Det at det forekom 5 feil svar under testingen av applikasjonen og kun ett feil svar under testingen av websiden kan kanskje forklares ved at brukerne som testet applikasjonen brukte kortere tid på å lete seg frem til riktig side, enn å lese seg frem til svaret. Dermed stresset de med å lese fort når de ble bevisste på hvor lang tid de brukte på selve siden.

En annen faktor som kunne kan ha påvirket testresultatene er *antallet personer i testgruppene med relevant erfaring*. Vi testet totalt 10 brukere, med 5 på hver testgruppe. Brukerne som testet applikasjonen hadde et snitt på 1.8 i erfaring med iPhone/iPad, og et snitt på 3.4 i erfaring for de som testet websiden. Det kan tenkes at selv uten signifikans, med en større gruppe og andre personer med annerledes teknologivaner, ville vi fått andre resultater.

Vi hadde dessverre ikke mulighet til å teste ut dette på grunn av ressursmangel, lite tid og vanskeligheter med å finne flere testbrukere.

En tredje faktor kan være *antallet testoppgaver* som hver testbruker utførte. Vi hadde 3 oppgaver til gjennomføring per deltager, der gjennomsnittlige tidsbruken på applikasjonen ble 150, mens gjennomsnittlige tidsbruken på websiden ble 229,4. Dersom vi hadde gitt flere brukeroppgaver, kunne kanskje

gjennomsnittsverdiene vært annerledes i resultatet, men dette er samtidig ikke sikkert.

Den fjerde faktoren som kunne ha påvirket resultatene er *informasjonsmengde i forhold til informasjonskanal*. Det kan for eksempel tenkes at enkle informasjonsmengder passer best til applikasjoner, mens andre mer innholdsrike ikke gjør det, eller omvendt. Våre oppgaver hadde litt forskjellige informasjonsmengder. Oppgave 1 og 2 var relativt enkle, men begge hadde kortest gjennomsnittlig tidsbruk på testen av websiden. Oppgave 3 var litt mer omfattende, men fikk kortest gjennomsnittlig tidsbruk på applikasjonen. Det er en mulighet for at vi kunne fått andre resultater dersom vi hadde tatt med flere forskjellige brukeroppgaver med ulik informasjonsmengde.

Hvis vi ser på erfaringer som skal begrunne brukernes atferd i tidsbruken, kan det hende at *deltakerne har over- eller undervurdert seg selv*. Hvis det er slik, er dette med på å påvirke det store bildet om hver og enkel testbruker når det gjelder teknologivaner i forhold til utføringen av brukeroppgavene. På en annen side, dersom deltakerne ikke har over- eller undervurdert seg selv, kan det hende at enkelte av disse rett og slett har hatt enten flaks i tidsbruken eller en dårlig dag, som forklarer enten for lite eller for mye tidsbruk på diverse oppgaver.

Det er stor usikkerhet rundt om noen av disse faktorene har påvirket testresultatene, selv om vi kan tenke oss at flere av disse kunne vært mulige.

6.0. Konklusjon

Selv om flere ulike faktorer kunne ha påvirket de endelige testresultatene, er ikke våre resultater statistisk signifikante. Nullhypotesen kan ikke forkastes, og det er bekreftet at det er ingen forskjell i tiden det tar å finne informasjon på applikasjonen i forhold til websiden til Telenor.

Referanser

- Eksperimentell design – Forelesningsslides fra INF2260 den 04.09.11, <http://www.uio.no/studier/emner/matnat/ifi/INF2260/h11/undervisningsmateriale/week3a.pdf> (besøkt 26.11.11).
- Two Sample independent t-Test – Mastee Badii, George Mason University. <http://www.youtube.com/watch?v=SHOSNYLrhus&feature=related> (besøkt 23.11.11).
- Wiley, John & Sons Ltd, 2007, *Interaction design – beyond human-computer interaction*, 2. Utgave, England.
- Lazar, Jonathan mfl., 2010, *Research Methods in human-computer interaction*, John Wiley & Sons Ltd, United Kingdom.

Alle kilder brukt i denne oppgaven er oppgitt.