



GÖTEBORGS UNIVERSITET

Användbarhetsutvärdering av gränssnittsprototyper

En jämförelse av heuristisk utvärdering och användarinvolverad utvärdering

Evaluating the usability of user interface prototypes
A comparison of heuristic evaluation and user testing

PONTUS NILSSON TENGNÄS
PATRIK NILSSON

Kandidatuppsats i Informatik

Rapport nr. 2016:018

Abstrakt

När allt mer avancerade applikationer kan köras i webbläsare ställs högre krav på god gränssnittsdesign då detta är en viktig faktor för att system ska bli konkurrenskraftiga. För utvärdering av ett gränssnitts design kan heuristisk utvärdering samt användarinvolverad utvärdering genomföras. De bägge metoderna har i tidigare studier jämförts vid utvärdering av färdigutvecklade gränssnitt. Denna studie syftar istället till att ge kunskap om vad för resultat som kan erhållas när metoderna används under tidiga utvecklingsstadier. En gränssnittsprototyp har utvärderats med bägge metoder och utfallen har analyserats och jämförts. Resultatet visade att bägge metoder lyckats upptäcka användbarhetsproblem men att heuristisk utvärdering i större grad fokuserar på strukturella brister medan den användarinvolverade utvärderingen resulterar i upptäckter gällande specifika gränssnittselement. För att uppnå högsta möjliga grad av användbarhet rekommenderas att bägge utvärderingsmetoder används.

Nyckelord: *Användbarhet, Utvärdering av gränssnitt, Utvärdering vid tidiga utvecklingsstadier, Heuristisk utvärdering, Användarinvolverad utvärdering*

Abstract

As more advanced applications can be used in web browsers there is an increasing demand for good interface design as this is an important factor regarding a system's success or failure. To evaluate the design of a user interface, heuristic evaluation and user testing can be performed. This study aims to compare these two evaluation methods to find what kind of result that is obtained by the respective methods when evaluating user interfaces in early development stages. The reason for focusing on a user interface in an early development stage is that previous research mainly compared the methods by evaluating already finished user interfaces. One user interface prototype has been evaluated with both methods and the results has been analysed and compared. The study showed that both methods has successfully found usability problems and that more problems of a structural sort was found with the heuristic evaluation method. The user testing, on the other hand, found more problems regarding specific interface elements. To achieve a high degree of usability, combining both methods is recommended for use in early development stages.

Keywords: *Usability, Evaluating user interfaces, Evaluation at an early development stage, Heuristic evaluation, User testing*

TACK

Vi vill rikta ett tack till vår handledare Alan Carlsson som bidragit med god och relevant feedback under arbetets gång.

Vi önskar också tacka Vendium AB vilka har gett oss möjligheten att utföra detta arbete samt bidragit med deltagare till vår studie.

Slutligen riktar vi ett stort tack till studiens alla deltagare som gjort denna uppsats möjlig att genomföra.

Innehållsförteckning

1. Introduktion.....	1
1.1. Syfte och frågeställning	2
1.2. Uppsatsens disponering	2
2. Teori.....	2
2.1. Användbarhet.....	3
2.2. Heuristisk utvärdering.....	5
2.3. Användarinvolverad utvärdering	5
2.4. Relaterad forskning.....	7
3. Metod.....	8
3.1. Heuristisk utvärdering.....	9
3.1.1. Deltagare	9
3.1.2. Förberedelse.....	10
3.1.3. Genomförande.....	12
3.2. Användarinvolverad utvärdering	13
3.2.1. Deltagare	13
3.2.2. Förberedelse.....	14
3.2.3. Genomförande.....	15
3.3. Analys	16
3.3.1. Klassificering av brister	17
3.3.2. Tolkning av data.....	19
3.4. Jämförelse	20
4. Resultat	20
4.1. Heuristisk utvärdering.....	20
4.2. Användarinvolverad utvärdering	24
5. Resultatanalys	30
5.1. Heuristisk utvärdering.....	30
5.1.1. Intuitivitet.....	30
5.1.2. Layout	31
5.1.3. Visualisering av information.....	32
5.2. Användarinvolverad utvärdering	32
5.2.1. Intuitivitet.....	33
5.2.2. Layout	33
5.2.3. Visualisering av information.....	34
6. Diskussion.....	34

7. Slutsats	36
7.1. Studiens relevans och överförbarhet	37
7.2. Förslag till fortsatt forskning.....	37
Referenslista.....	39

Bilagor

- Bilaga 1 - Instruktioner för heuristisk utvärdering
- Bilaga 2 - Medgivande om ljud och skärminspelning vid observation samt intervju
- Bilaga 3 - Scenarion vid observation
- Bilaga 4 - Intervjumall
- Bilaga 5 - Skärmdumpar från prototyp

Tabeller och figurer

Figur 1. Skärmdump från formuläret

Tabell 1. Exempeltabell för brister inom kategori intuitivitet som identifierats under heuristisk utvärdering

Tabell 2. Heuristisk utvärdering - Brister inom intuitivitet

Tabell 3. Heuristisk utvärdering - Brister inom layout

Tabell 4. Heuristisk utvärdering - Brister inom visualisering av information

Tabell 5. Heuristisk utvärdering - fördelning av brister utefter allvarlighetsgrad

Tabell 6. Heuristisk utvärdering - fördelning av brister utefter kategori av brist

Tabell 7. Användarinvolverad utvärdering - Brister inom intuitivitet

Tabell 8. Användarinvolverad utvärdering - Brister inom layout

Tabell 9. Användarinvolverad utvärdering - Brister inom visualisering av information

Tabell 10. Användarinvolverad utvärdering - fördelning av brister utefter allvarlighetsgrad

Tabell 11. Användarinvolverad utvärdering - fördelning av brister utefter kategori av brist

Tabell 12. Fördelning av brister

Tabell 13. Antal brister utefter allvarlighetsgrad och utvärdering där de upptäcktes

1. Introduktion

Med Web 2.0 och så kallade *Rich Internet Applications*¹ har webbsidor blivit allt mer kapabla till att ersätta vanliga skrivbordsprogram. Exempelvis är det numera möjligt att utföra relativt avancerade uppgifter så som ordbehandling och fakturering i en webbläsare. För att dessa tjänster ska bli så konkurrenskraftiga som möjligt är det av stor vikt att gränssnittsdesignen är välfungerande då detta är en viktig faktor vad gäller användares benägenhet att använda ett system eller ej (Fernandez, Insfran & Abrahão 2011; Nazrul & Tétard 2014). Vid en bristfällig design kan användare exempelvis uppleva frustration och irritation, något som kan leda till att de överger ett system eller en produkt (Rogers, Sharp & Preece 2011).

Det ställs därmed krav på utvecklare att skapa produkter och tjänster med god användbarhet och design (Delice & Güngör 2009). För att utvärdera huruvida ett gränssnitt är väl designat och användbart används vanligtvis någon form av utvärderingsmetod. Två vanliga metoder för utvärdering är *heuristisk utvärdering* samt *användarinvolverad utvärdering* (Batra & Bishu 2007; Tan, Liu & Bishu 2009). Vilken av dessa metoder som används skulle kunna bero på flertalet faktorer, exempelvis ett projekts tidsram, budget och möjligheten att komma i kontakt med användare (Fernandez, Insfran & Abrahão 2011).

Dessa två utvärderingsmetoder skiljer sig från varandra vad gäller ett antal faktorer men syftet med dem är det samma, att upptäcka brister i ett gränssnitt så att dessa kan åtgärdas (Doubleday, Ryan, Springett & Sutcliffe 1997; Hertzum & Jacobsen 2003). Heuristiska utvärderingar är vanligtvis billigare och mer tidseffektiva medan användarinvolverad utvärdering bidrar med den värdefulla feedback som användare kan ge (Nielsen 1990; Fernandez, Insfran & Abrahão 2011). Då det troligtvis inte alltid är möjligt att genomföra den typ av utvärdering som önskas ligger det i gränssnittsdesigners intresse att veta vad olika typer av utvärdering leder till och vilken typ av designbrister som fångas upp, alternativt riskerar att förbises, med en viss utvärderingsmetod. Därav är det nödvändigt med en kartläggning av dessa två utvärderingsalternativ för att komma fram till vad respektive metod leder till för resultat.

Med kunskap om detta kan gränssnittsdesigners anpassa sitt arbete utefter vilken utvärderingsmetod som kommer att användas och exempelvis vara extra uppmärksamma på faktorer som ofta missas av den aktuella metoden. Kunskaperna ger också designers en bättre möjlighet att bedöma i vilken utsträckning en utvärdering är pålitlig eller ej för den aktuella situationen. Det finns även ett teoretiskt intresse i att få kunskap om huruvida, och på vilket sätt, dessa utvärderingsmetoder kompletterar varandra. Tidigare forskning har i stor utsträckning berört utvärdering av färdigutvecklade gränssnitt (Doubleday et al 1997; Kurečić 2005; Delice & Güngör 2009; Tan, Liu & Bishu 2009) varav vårt fokus lags på webbgränssnitt i ett prototypstadium. Med denna inriktning kan insikter i huruvida tidigare teori kring utvärdering även är applicerbar på prototyper erhållas. För att bidra till en ökad kunskap inom området presenteras

¹ Komplexa och innehållsrika webbapplikationer som delar många egenskaper med traditionella skrivbordsprogram (Maan & Mantha 2014).

i denna uppsats en jämförande studie mellan heuristisk utvärdering och användarinvolverad utvärdering av ett gränssnitt i prototypstadie.

I studien utvärderas en prototyp till ett gränssnitt, utvecklad av författarna, både heuristiskt och av användare. Prototypen har utvecklats på uppdrag av Vendium AB och är ett förslag på ett nytt gränssnitt till deras webbaserade system Struqtur. Struqtur är ett system vars målgrupp är bygg- och anläggningsföretag och innehåller funktioner för bland annat projekt-, faktura- och offerthantering.

1.1. Syfte och frågeställning

Denna uppsats har till syfte att bidra till en ökad insikt i på vilket sätt valet av utvärderingsmetod påverkar den feedback som utvecklarna erhåller. Fokus kommer att riktas mot utvärdering av gränssnitt för webbaserade tjänster i ett prototypstadie. Avsikten med uppsatsen är inte att ställa utvärderingsmetoderna mot varandra i syfte att komma fram till vilken som fungerar mest effektivt, utan snarare att få en ökad kunskap om de respektive metodernas förväntade resultat. Vår förhoppning är att kunna klargöra vad de olika utvärderingsmetoderna upptäcker för typ av designbrister och vad för likheter och skillnader som existerar mellan dem. Med detta önskar vi kunna bidra med rekommendationer och riktlinjer till gränssnittsutvecklare vad gäller utvärderingsmetoder.

Mot given bakgrund tar författarna en förståelseinriktad samt normativ ansats och önskar besvara följande frågeställning:

Vilka typer av designbrister upptäcks vid heuristisk respektive användarinvolverad utvärdering av ett webbgränssnitt i prototypstadie?

1.2. Uppsatsens disponering

I kapitel 2 presenteras teori gällande användbarhet, heuristisk utvärdering, användarinvolverad utvärdering samt tidigare jämförelser av utvärderingsmetoderna. Detta för att på ett utförligt vis beskriva uppsatsens mest centrala delar och begrepp samt skapa en förståelse för de utvärderingsmetoder som studeras i denna undersökning. Därefter följer i kapitel 3 en beskrivning av den metod som använts för att besvara uppsatsens frågeställning. I kapitlet beskrivs hur de olika utvärderingarna har genomförts samt hur den data som samlats in i samband med detta analyserats. Resultatet av studien redogörs för i kapitel 4 och följs därefter av en resultatanalys i kapitel 5. Diskussion gällande resultatet och resultatanalysen framförs sedan i kapitel 6. Uppsatsen avslutas med kapitel 7 där slutsatser av studien samt förslag till vidare forskning presenteras.

2. Teori

Detta kapitel ämnar beskriva uppsatsens mest centrala beståndsdelar. Kapitlet är uppdelat i fyra underrubriker varav den första behandlar begreppet användbarhet, vilket är det som utvärderas

med de två utvärderingsmetoder som studeras i denna uppsats. Därefter följer redogörelser för principerna bakom heuristisk utvärdering och användarinvolverad utvärdering samt hur dessa metoder kan appliceras vid en utvärdering. Avslutningsvis presenteras tidigare forskning där jämförelser och granskning av utvärderingsmetoderna genomförts.

2.1. Användbarhet

För att beskriva huruvida ett system har en god eller mindre god design ur ett användarperspektiv kan begreppet *'användbarhet'* (på engelska *'usability'*) användas. Graden av användbarhet brukar inom IT-relaterad forskning syfta till att förklara hur lätt (eller svårt) det är för tänkta användare att nyttja en applikation via dess gränssnitt. Desto lägre svårighetsgrad användare i regel upplever att ett system har avgör hur framgångsrikt systemet kan betraktas vara (Fernandez, Insfran & Abrahão 2011). Andra uttryck som exempelvis *'användarvänlighet'*, *'människa-dator interaktion'*, *'användarcentrerad design'*, *'människa-maskininterface'* och *'user interface design'* har använts av experter genom åren för att på olika sätt beskriva systemdesign ur ett användarperspektiv (Nielsen 1994). ISO 9241-11 (ISO 1998) definierar användbarhet som:

"Den grad i vilken specifika användare kan använda en produkt för att uppnå ett specifikt mål på ett ändamålsenligt, effektivt och för användaren tillfredsställande sätt i ett givet sammanhang." - Översatt av Usability Partners (u.å.)

Denna definition är enligt Bevan (1995) lämplig att utgå från vid bedömning av en applikation eller programvara utifrån användarperspektiv. Förutom att utvärdera det primära användargränssnittet inkluderar detta även andra steg där användaren interagerar med systemet, så som installation och underhåll (Nielsen 1994). ISO-standarderna är väl vedertagna och används för att definiera användbarhet av bland annat Delice och Gungör (2009), Fernandez, Insfran och Abrahão (2011), Liljegren (2006) samt Nazrul och Tétard (2014).

Genom att precisera begreppet användbarhet i mer konkreta och mätbara komponenter går det att systematiskt ta sig an, utvärdera och förbättra system ur ett användarperspektiv (Nielsen 1994). Detta i förhållande till om graden av användbarhet för ett system hade bedömts på ett mer abstrakt och subjektivt plan. Resultatet av en sådan utvärdering hade sannolikt inte varit lika användbar för exempelvis jämförelser över tid i samband med utveckling av ett system, inte heller för att jämföra graden av användbarhet mellan olika system och designförslag (Nielsen 1994).

Nielsen (1994) nämner fem olika aspekter av användbarhet som med fördel kan bedömas individuellt inom ett system med syfte att utröna vilka specifika styrkor och brister som finns. Sammantaget bildar dessa aspekter en grad av användbarhet för systemet, de fem aspekterna listas nedan så som Nielsen (1994, s. 26) beskriver dem. Ursprungstexterna har använts för att undvika missstolkningar och missförstånd samt att lämplig svensk översättning inte hittats:

- *Learnability*
 - The system should be easy to learn so that the user can rapidly start getting some work done with the system.
- *Efficiency*
 - The system should be efficient to use, so that once the user has learned the system, a high level of productivity is possible.
- *Memorability*
 - The system should be easy to remember, so that the casual user is able to return to the system after some period of not having used it, without having to learn everything over again.
- *Errors*
 - The system should have a low error rate, so that users make few errors during the use of the system, and so that if they do make errors they can easily recover from them. Further, catastrophic errors must not occur.
- *Satisfaction*
 - The system should be pleasant to use, so that users are subjectively satisfied when using it; they like it.

- *Definitioner ovan från Nielsen (1994, s. 26).*

För webbaserade gränssnitt menar Fernandez, Insfran och Abrahão (2011) att det uppstått stora utmaningar när det kommer till att designa allt mer användbara applikationer. Detta har lett till att det uppkommit ett stort behov av metoder, tekniker och verktyg som går att använda för att förebygga och åtgärda en låg grad av användbarhet. Vid utvärdering av hemsidor tenderar användarupplevelsen att väga mycket tungt i förhållande till andra aspekter, därför bör en hög grad av användbarhet prioriteras vid design av gränssnitt (Batra & Bishu 2007).

Användbarhet kan mätas på olika sätt med olika metoder. Två av de mest vanliga är heuristisk utvärdering och användarinvolverad utvärdering (Batra & Bishu 2007; Tan, Liu & Bishu 2009). Heuristisk utvärdering anses vara en kostnads- och tidseffektiv metod för att bedöma användbarheten av ett system (Alshamari 2015; Nielsen & Molich 1990; Rogers, Sharp & Preece 2011). Utvärdering utförs av ett antal testare som förslagsvis är experter inom just användbarhet och gränssnittsdesign. Personer utan speciell erfarenhet av dessa områden kan också utvärdera system heuristiskt (Nielsen & Molich 1990), de tenderar dock att hitta färre och andra typer av designbrister än mer erfarna testare (Nielsen 1992). Under utvärdering testas systemet utifrån ett antal regler kallade heuristiker vilket kommer att behandlas mer utförligt i nästa avsnitt. Utvärderingsmetoden har fått kritik då den från vissa håll sägs bedöma användbarhet ur en mer generell kontext och ibland missar syftet med att utvärdera om programvaran är användbar i specifika sammanhang för specifika användare (Alshamari 2015).

Användarinvolverade utvärderingar görs ofta med syftet att bedöma om programvaran är användbar för den specifika användargrupp som systemet är designat för (Rogers, Sharp & Preece 2011). Detta sker genom att låta representativa testanvändare (eller verkliga användare) utöva en rad på förhand bestämda aktiviteter i ett system för utvärdering (Nielsen 1994). Några

av de tekniker som kan användas under utvärderingen är experiment, observationer, intervjuer och enkätinsamling (Rogers, Sharp & Preece 2011). Enligt Nielsen (1994) kan man inför en utvärdering specificera kvantifierbara målvärden som är relevanta för det som ska mätas. Det faktiska resultatet (median- och/eller medelvärden) skulle i sådana fall kunna jämföras med de på förhand bestämda målen för att avgöra om systemet har en tillräckligt hög grad av användbarhet.

2.2. Heuristisk utvärdering

Heuristisk utvärdering är en informell metod som används till att upptäcka problem och bristfällig interaktion i olika typer av gränssnitt. Metoden kan användas inte enbart till att utvärdera gränssnitt för datorprogram utan även andra typer av gränssnitt så som automatiska telefonsvarare (Nielsen 1990). En heuristisk utvärdering utförs genom att ett antal utvärderare, hädanefter benämnda som 'testare', inspekterar och utvärderar ett existerande gränssnitt i förhållande till på förhand bestämda principer eller bedömningspunkter om god design (Nielsen 1990; Nielsen 1992; Rogers, Sharp & Preece 2011). Testarna utför inspektionerna var för sig och noterar de överträdelser, samt graden av allvarlighet hos dessa överträdelser, av principerna som de upptäcker. Därefter sammanställs samtliga överträdelser som har upptäckts av alla testare till en samlad bedömning gällande ett gränssnitts potentiella brister. Det krävs inte att testarna föreslår en lösning på de brister som upptäcks vilket kan ses som en nackdel med metoden (Nielsen 1990). En heuristisk utvärdering bör anpassas till det stadiet i utvecklingsprocessen en produkt befinner sig i. Ett färdigutvecklat system ger testarna möjlighet att utföra olika på förhand bestämda uppgifter, medan en tidigare version kanske inte kan erbjuda detta i samma utsträckning (Rogers, Sharp & Preece 2011). Metoden kräver dock inte att testarna utför verkliga uppgifter och det är därmed möjligt att genomföra heuristiska utvärderingar i ett tidigt utvecklingsstadium (Nielsen 1995b).

Vad som krävs för att en person kan anses vara lämpad att utföra en heuristisk utvärdering tenderar att variera. I en studie av Tan, Liu och Bishu (2009) anses studenter inom människa-dator interaktion som utbildats inom, samt genomfört, minst en heuristisk utvärdering vara experter. Alshamari (2015) definierar i sin studie en expert som en individ med examen i datavetenskap samt tidigare erfarenhet av användbarhetsutvärdering. Heuristiska utvärderingar kan även utföras av personer som saknar erfarenhet av detta (Nielsen 1992). Personer med större erfarenhet av användbarhet tenderar dock att upptäcka fler brister (Nielsen 1992).

Då en heuristisk utvärdering baseras på testares inspektion av gränssnitt krävs inte slutanvändare av det aktuella systemet för att kunna genomföra en utvärdering. Detta innebär att en heuristisk utvärdering är förhållandevis billig och enkel att utföra (Rogers, Sharp & Preece 2011). Metodens tillvägagångssätt innebär även att det är testarnas personliga åsikter och erfarenheter som ligger till grund för de designbrister som upptäcks. Detta innebär att det, enligt Hertzum och Jacobsen (2003), kan förekomma stora skillnader i upptäckter mellan olika testare och Nielsen (1990) rekommenderar att minst tre till fem testare deltar vid en utvärdering. En utvärdering med tre till fem testare bör upptäcka mellan 50-70 procent av de förekommande designbristerna i ett gränssnitt (Nielsen 1990). Enligt Hertzum och Jacobsen (2003) kan det dock vara nödvändigt att använda ytterligare utvärderare, i synnerhet vid kritiska situationer.

För att genomföra en heuristisk utvärdering krävs designprinciper, hädanefter benämnda 'heuristiker', som testarna ska utgå från när de inspekterar gränssnittet. Nielsen (1995a) har definierat tio generella heuristiker som kan användas i samband med heuristiska utvärderingar. Vanligtvis tar dock testare och forskare fram en samling heuristiska principer lämpade för det specifika gränssnitt som skall utvärderas (Rogers, Sharp & Preece 2011). Inför utvärdering kan de heuristiker som bestämts exempelvis bestå av principer från både Nielsen (1995a) och egna definierade principer.

De tio heuristiker Nielsen (1995a) tagit fram har utvecklats och uppdaterats under många år med ambitionen att skapa ett ramverk som går att utgå ifrån vid heuristiska utvärderingar. Dessa heuristiker listas nedan och de som har använts under denna studie redovisas i detalj med tillhörande beskrivning i uppsatsens metodredogörelse, avsnitt 3.1.2.

- *Visibility of system status*
- *Match between system and the real world*
- *User control and freedom*
- *Consistency and standards*
- *Error prevention*
- *Recognition rather than recall*
- *Flexibility and efficiency of use*
- *Aesthetic and minimalist design*
- *Help users recognize, diagnose, and recover from errors*
- *Help and documentation*

Dessa heuristiker behandlar olika aspekter som kan förekomma i gränssnitt och fungerar på så vis som ett stöd för testaren vid utvärdering. Genom att bedöma gränssnittets olika element utifrån dessa principer vägleds testaren i sitt arbete.

2.3. Användarinvolverad utvärdering

Med användarinvolverad utvärdering önskar utvärderare fånga upp faktiska, alternativt framtida, användares åsikter och upplevelser av ett gränssnitt. Detta kan uppnås genom att utföra så kallade 'Think aloud'-sessioner tillsammans med användare. Think aloud är en metod vars syfte är att ta till vara på individers omedelbara tankar och känslor i samband med att de utför en specificerad uppgift. För att kunna registrera dessa tankar och känslor får ett antal användare individuellt utföra olika slag av uppgifter i ett gränssnitt, samtidigt som de säger vad och hur de tänker kring det som utförs (Nazrul & Tétard 2014; Rogers, Sharp & Preece 2011). I samband med detta observeras användaren av en eller flera utvärderare som dokumenterar de problem användaren möter under sessionen. Denna dokumentation kan ske på olika vis, exempelvis genom att testsessionerna spelas in med videokamera (Rogers, Sharp & Preece 2011) eller att utvärderare fyller i ett observationsschema där olika relevanta beteenden prickas av i en tabell i samband med att de inträffar (Patel & Davidson 2011). Vanligtvis låter utvärderarna fler än en användare, fyra

till fem anses vara nödvändigt, utföra uppgifter i systemet vilket leder till att en lista med samtliga upptäckta problem kan sammanställas efter testsessionerna (Hertzum & Jacobsen 2003).

Think aloud-sessioner kan utföras både tidigt och sent i systemutvecklingsprocessen och syftet med att genomföra en utvärdering kan variera. Det är möjligt att genomföra sessioner ute i en verklig miljö (Hertzum & Jacobsen 2003) samtidigt som de också kan utföras i en kontrollerad miljö, exempelvis ett laboratorium (Rogers, Sharp & Preece 2011).

En think aloud-session inleds vanligtvis med att definiera olika typer av uppgifter som användare ska utföra i systemet. Sedan informeras användarna om testets syfte och ombeds att försöka återge så mycket som möjligt av deras tankar i samband med testet. Utvärderaren ger sedan användaren uppgifter som denne ska utföra (Hertzum & Jacobsen 2003). Det kan uppfattas som onaturligt för användare att berätta vad de tänker på (ibid), och att användare är för tysta i samband med sessionerna är en potentiell nackdel med metoden (Rogers, Sharp & Preece 2011). Ytterligare en nackdel med think aloud är att resultaten av en session kan påverkas av de utvärderare som observerar användarna. Hertzum och Jacobsen (2003) menar att det kan förekomma skillnader i vad olika utvärderare uppfattar som ett problem och därmed dokumenterar.

2.4. Relaterad forskning

Vid en jämförelse av heuristisk respektive användarinvolverad utvärdering av webbgränssnitt konstaterar Tan, Liu och Bishu (2009) att det förekommer ett antal skillnader i resultaten som erhålles av dessa metoder. Deras studie genomfördes på fyra olika existerande webbsidor som utvärderades dels heuristiskt av nio testare och dels tillsammans med tolv användare. Resultatet visade bland annat att de heuristiska utvärderingarna upptäckte ett större antal brister i gränssnitten samt att dessa brister i högre utsträckning berörde strukturella problem hos en design. De användarinvolverade utvärderingarna ledde i större utsträckning till mer detaljerade upptäckter gällande specifika element i ett gränssnitt. Av detta drar de slutsatsen att bägge metoder är effektiva och bör användas i kombination vid en utvärdering för att upptäcka så stor andel designbrister som möjligt.

I en studie av Liljegren (2006) har fyra olika utvärderingsmetoder analyserats utefter deras karakteristik samt tillvägagångssätt, och faktiska utvärderingar har inte utförts. Studien är inriktad mot system som används inom sjukvården och Liljegren (2006) menar att användarinvolverad utvärdering är den metod som är bäst lämpad för att upptäcka problem i denna typ av gränssnitt. Detta då användarinvolverad utvärdering anses beakta de mest relevanta aspekterna av användbarhet, i synnerhet vad gäller att upptäcka situationer där det förekommer en hög felfrekvens hos användarna. Därmed skiljer sig det resultat Liljegren (2006) kommit fram till från det Tan, Liu och Bishu (2009) upptäckte i sin studie där utfallet visade att båda metoderna är effektiva och bör kombineras för bästa resultat.

Doubleday et al (1997) har genomfört en studie där de utvärderar ett grafiskt gränssnitt till ett databashanteringssystem heuristiskt samt med think aloud-sessioner. Resultaten av deras studie

visade att de bägge metoderna upptäcker olika typer av problem då 39 procent av de problem som upptäcktes med think aloud-metoden inte upptäcktes av den heuristiska utvärderingen. Samtidigt missade think aloud-utvärderingen 40 procent av de problem som upptäcktes av den heuristiska utvärderingen. Den heuristiska utvärderingen upptäckte mer generella, övergripande, problem medan think aloud-utvärderingen upptäckte mer specifika problem i samband med vissa handlingar i systemet. Samtidigt gav den användarinvolverade utvärderingen inga direkta förslag på lösningar till problemen, vilket experterna i den heuristiska utvärderingen däremot gjorde.

Doubleday et al (1997) resonerar om huruvida dessa skillnader i resultat kan bero på att de heuristiska testarna har svårt att leva sig in i en användares situation och därav inspekterar gränssnittet på ett mer övergripande, strukturellt plan. Användarna å sin sida kan ha svårigheter med att upptäcka och uttrycka denna typ av problem och reagerar i stället på symptomen av dem.

Det existerar en del likheter mellan det resultat som Tan, Liu och Bishu (2009) och Doubleday et al (1997) kommit fram till. Exempelvis vad gäller den användarinvolverade utvärderingens större fokus på specifika detaljer i ett gränssnitt. De menar också att bägge metoderna har unika kvalitéer medan Liljegren (2006) hävdar att användarinvolverad utvärdering är den mest effektiva metoden.

Tan, Liu och Bishu (2009), Doubleday et al (1997) och Liljegren (2006) redovisar för hur resultaten av metoderna skiljer sig vid utvärdering av färdigutvecklade gränssnitt, Tan, Liu och Bishu (2009) och Doubleday et al (1997) visar också på vikten av att kombinera heuristisk utvärdering och användarinvolverad utvärdering då olika typer av brister tenderar att upptäckas. För att bygga vidare på, samt vidga, denna kunskap fokuserar vår studie på gränssnitt i prototypstadie. På så vis kan vi dra paralleller mellan resultatet av denna studie och resultaten av ovan nämnda studier för att se huruvida detta även är applicerbart på gränssnitt i prototypstadie.

3. Metod

För att besvara frågeställningen, "*vilka typer av designbrister upptäcks vid heuristisk respektive användarinvolverad utvärdering av ett webbgränssnitt i prototypstadie?*", har heuristiska utvärderingar och användarinvolverade utvärderingar förberetts och genomförts på ett gränssnitt. Material från respektive utvärdering har samlats in, analyserats och jämförts för att sedan ha fungerat som underlag för denna studies resultat. Då syftet för studien var att jämföra dessa två utvärderingsmetoder har inga justeringar skett på det gränssnitt som testats mellan olika utvärderingstillfällen. Att utvärdering skett vid olika tidpunkter har därför inte påverkat studiens resultat på något vis.

För studien har utvärderingar skett på ett förslag till ett nytt gränssnitt, en digital prototyp, som utvecklats av författarna till denna uppsats. Prototypen har tagits fram på uppdrag av Vendium AB med syfte att utveckla ett designförslag för det användargränssnitt som tillhör företagets webbaserade administrativa verktyg Struqtur. De som använder Struqtur är vanligtvis verksamma inom bygg-, anläggnings- eller hantverksföretag av mindre storlek och systemet innehåller bland annat funktioner i form av projekt-, faktura- och offerthantering. Därmed innehåller systemet

exempelvis information om ett företags kunder, anställda, arbetstider, materiel samt kostnader. Systemet är i huvudsak ett administrativt verktyg och dess användare har varierande grad av datorvana.

Prototypen är uppbyggd som en hemsida bestående av HTML, CSS och Javascript och kan köras i en webbläsare (se bilaga 5 för skärmdumpar från systemet). Det är möjligt att navigera mellan olika platser (HTML-sidor) i gränssnittet, ta del av information som visas samt ange input till systemet.

3.1. Heuristisk utvärdering

För den heuristiska utvärderingen har tre personer med akademiska och praktiska erfarenheter av interaktionsdesign och gränssnittsdesign deltagit. Testarna har självständigt, under en session, utvärderat och bedömt gränssnittsprototypen i enlighet med heuristisk metodik och sedan skickat resultaten till uppsatsens författare. Nedan följer en redogörelse för urvalet av deltagande testare, förberedelser inför utvärdering samt genomförande av den samma.

3.1.1. Deltagare

Enligt Nielsen (1992) blir en heuristisk utvärdering effektiv när deltagande testare har kunskap om hur gränssnitt skall designas ur ett användbarhetsperspektiv. Därför har personer med praktiska samt akademiska erfarenheter av interaktions- och gränssnittsdesign tillfrågats om möjligt deltagande i denna heuristiska utvärdering. Nielsen (1992) har också funnit att resultatet för en utvärdering, d.v.s. hur stor andel av systemets totala antal designbrister som hittas, blir av än högre kvalité om testarna också har erfarenhet av den typ av system som utvärderas. En sådan testare kallas för *'double expert'* och har all den kompetens som kan önskas för deltagande i en heuristisk utvärdering. Förfrågan om deltagande har också ställts till personer som stämmer in på kriterierna för att bli kallad double expert i detta sammanhang.

Tre personer tackade slutligen ja till att heuristiskt utvärdera det gränssnitt som skall bedömas. En av dessa tre anses vara en double expert då personen i fråga har arbetat med att utveckla systemet Struktur vars nya gränssnittsförslag är det som utvärderats under denna studie. De två andra testarna har mer generell kunskap om gränssnitts- och interaktionsdesign, de har också specifik erfarenhet av andra typer av system. Enligt Nielsen (1995b) finner tre testare tillsammans i genomsnitt ca 60 procent av ett systems designbrister under heuristiska utvärderingar. Ett mer optimalt antal testare hade varit fem stycken, som tillsammans brukar finna ca 70 procent av designbristerna (Nielsen 1990; Nielsen 1995b). Vid fler än fem personer som utvärderar minskar den nytta varje enskild testare gör så till vida att många av de brister som upptäcks av den sjätte och senare testaren redan upptäckts av testare 1-5 (Nielsen 1995b).

Premissen för denna specifika utvärdering var att finna testare väl lämpade för att utvärdera gränssnitt och som utan arvode kunde tänka sig att ställa upp. Utifrån dessa förutsättningar har tre personer tackat ja till att medverka som testare i studien. Nielsen (1992) hävdar att det ibland är svårt att få tag på testare till heuristiska utvärderingar och därför har bedömningen gjorts att det

inte varit lönt att utöka processen för att möjligtvis finna fler testare. Något som i sådana fall hade kunnat ske på bekostnad av exempelvis andra förberedelseaktiviteter eller analysarbetet.

De tre deltagande testarna presenteras nedan. Enbart de uppgifter som är relevanta att delge för att ge förståelse kring huruvida varje enskild testare har förmåga att utvärdera gränssnitt har angetts. Då det funnits en ambition att hålla deltagarna anonyma har demografisk information, till exempel rörande deltagarnas kön, ålder samt studie- och arbetsplatser, utelämnats. Samtliga deltagare har samtyckt till att presenteras i nedan stående form samt låtit sitt bidrag till utvärderingen publiceras i denna uppsats.

Testarna har fått var sin unik kod, H1, H2 och H3, som härnäst kommer användas vid hänvisning till den specifika testarens bedömningar, resultat och kommentarer.

Deltagare H1:

Arbetar för närvarande som gränssnittsdesigner och har fyra års arbetslivserfarenhet av detta. Har tidigare studerat Multimedia Design och Fine Art. Tar kandidatexamen i informatik 2016.

Deltagare H2:

Har ett års arbetslivserfarenhet som UX-designer. Studerar för en masterexamen inom interaktionsdesign.

Deltagare H3:

Arbetar som full stack developer på Vendium AB sedan fyra år och har därför både ansvar för Strukturs gränssnitt och bakomliggande system. Har dock ingen medverkan i designen av den prototyp som utvärderas i denna studie utan har endast tagit rollen som testare i samband med utvärderingen. Har sedan tidigare en kandidatexamen inom informatik.

3.1.2. Förberedelse

En av premisserna för heuristisk utvärdering är att testare ska utvärdera ett gränssnitt i förhållande till ett antal på förhand bestämda heuristiker (Nielsen 1990; Nielsen 1992; Rogers, Sharp & Preece 2011). Dessa heuristiker, som även kan ses som kriterier eller regler för god design, har som syfte att ge struktur åt testaren som förhoppningsvis mer effektivt kan leta efter och hitta brister i det gränssnitt som utvärderas. Utan fördefinierade heuristiker tenderar testare att istället utvärdera ett gränssnitt utifrån egen intuition och förnuft (Nielsen 1990).

Inför de heuristiska utvärderingarna har sex heuristiker bestämts, definierats och skickats till testarna i ett dokument. Fem av heuristikerna är direkt hämtade från de generella heuristiker som Nielsen (1995a) tagit fram och presenteras nedan med dess beskrivningar av vad testaren ska leta efter. Den sjätte heuristiken, '*Accessible information*', är framtagen av författarna för att täcka in aspekten av hur enkelt det är att ta till sig information som ges av gränssnittet. Att konstruera egna, för det specifika gränssnittet lämpliga, heuristiker är möjligt och uppmuntras (Nielsen 1995b; Rogers, Sharp & Preece 2011). De sex heuristiker som används i denna studie listas nedan. De som är hämtade från Nielsen (1995a) har angivits i dess ursprungliga definition för att

undvika misstolkningar och missförstånd samt att lämplig svensk översättning inte hittats. För att uppnå ett konsekvent format har även den egenformulerade heuristiken definierats på engelska.

- *Visibility of system status*
 - The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.
- *Match between system and the real world*
 - The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.
- *Consistency and standards*
 - Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.
- *Recognition rather than recall*
 - Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.
- *Aesthetic and minimalist design*
 - Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

-Ovanstående fem heuristiker hämtade från Nielsen (1995a).

- *Accessible information*
 - Information and data should be presented in an accessible and understandable way so that the user easily can comprehend its meaning and act upon it. It should not be difficult to understand the meaning of, and the reason for presenting, different types of information.

Utöver heuristikerna har ett dokument med instruktioner för hur testarna skall genomföra den heuristiska utvärderingen författats. I instruktionerna har testarna kunnat läsa hur de önskas ta sig an utvärderingen, hur de skall ha gått till väga för att starta prototypen samt hur de skall ha dokumenterat de brister som de kunnat identifiera. De har också fått allmän information om studiens syfte, hur en heuristisk utvärdering går till, hur personlig information behandlats i denna uppsats o.s.v. För att läsa instruktionerna i sin helhet, se bilaga 1.

För att ge testarna möjlighet att dokumentera de brister som hittats har ett formulär skapats i Google:s produkt Google Forms. Här har testarna kunnat ange vilken heuristik bristen fallit inom tillsammans med en beskrivning av bristen. De har också fått ange en grad av allvarlighet utifrån hur stor påverkan de bedömt att bristen haft på systemets användbarhet. Figur 1 illustrerar ett stycke i det formulär där angivning av brist skett. Testarna har i formuläret också fått ange sin

deltagar-kod, en beskrivning av vad de haft för erfarenheter inom interaktions- och gränssnittsdesign samt det datum utvärderingen ägde rum.

The screenshot shows a survey form with a purple header and a white body. The title is 'Överträdelser av heuristiker'. Below the title is a purple bar with the text 'Överträdelser'. The main text reads: 'Fyll i de överträdelser du hittar. I formuläret finns utrymme för att fylla i tio stycken. Om du hittar fler än tio överträdelser/designbrister, var god slutför formuläret och påbörja ifyllnad av ett nytt.' Below this is a purple bar with the text 'Överträdelse 1 av 10'. The main question is 'Inom vilken heuristik fann du bristen/överträdelsern?'. Below the question is a dropdown menu with the text 'Choose' and a downward arrow. Below the dropdown is a section titled 'Allvarlighetsgrad' with three radio buttons labeled '1', '2', and '3'. To the left of the radio buttons is the text 'Mycket allvarlig' and to the right is 'Mindre allvarligt'. Below the radio buttons is a text input field with the label 'Beskriv bristen/överträdelsern' and the placeholder text 'Your answer'.

Figur 1. Skärmdump från formuläret

Sista steget under förberedelserna inför utvärderingen har inneburit att samla de resurser som vi ämnat skicka till testarna i en zip-fil. Denna zip-fil har bestått av:

1. Mapp med prototypen som utvärderats
2. Dokument med ovan nämnda heuristiker
3. Dokument med ovan nämnda instruktioner (här finns även en länk till det formulär testarna fyllt i)

Filen har skickats till respektive deltagare via e-mail tillsammans med deras tilldelade deltagar-kod och deadline för utvärderingen.

3.1.3. Genomförande

Var och en av testarna har individuellt utvärderat gränssnittet genom att självständigt granska prototypens olika delar med stöd av angivna heuristiker. Vid uppmärksam överträdelser av heuristiker har brister angetts och beskrivits i formulär. Då heuristisk utvärdering går ut på att få

ärliga och opartiska bedömningar av testarna är det viktigt att se till att testarna inte kan påverkas av andra deltagare inför eller under utvärderingen (Nielsen 1995b). Vi som anordnat studien har dessutom valt att inte närvara under utvärderingstillfällena för att undvika att själva påverka testarna. En av testarna har dock självmant efterfrågat att vi som utvecklade prototypen kunnat närvara på plats under utvärderingstillfället för att lättare kunna få snabba svar på eventuella frågor. Detta har respekterats och vi har därför funnits i testarens närhet, utan att på något sätt ha övervakat eller kommenterat på testarens handlingar under utvärderingstillfället.

Förutom att testarna fått förhålla sig till angivna heuristiker och instruktioner om att grundligt och detaljerat bedöma gränssnittet har de själva, på individbasis, fått välja hur de praktiskt gått tillväga vid utvärderingen. Hur lång tid sessionen fått ta samt huruvida ifyllning av formulär skett under eller efter granskning av gränssnittet är exempel på sådant testarna fått bestämma själva. Enligt Nielsen (1995b) är det möjligt att testare vid utvärdering hittar brister som inte går att para ihop med någon av de på förhand bestämda heuristikerna. Därför bör de uppmanas att ändå ta med bristen eller avvikelser i resultatet. Sådan uppmaning har även givits i detta fall, dock hittades ingen sådan brist av någon av testarna.

Resultatinsamling har skett genom att deltagarna antecknat sina bedömningar av prototypen i det elektroniska formuläret. Varje brist har av testaren tilldelats en heuristik, en allvarlighetsgrad och en beskrivning (se figur 1 från utvärderingsformuläret). När testaren genomfört utvärderingen och valt att skicka in formuläret har det automatiskt sparats i Google Forms. De ifyllda formulärens har samlats säkert på ett sådant vis att enbart vi som anordnat studien haft tillgång till materialet via ett Google-konto. Samtliga deltagande testare har framgångsrikt fyllt i och skickat in var sitt formulär enligt givna instruktioner.

3.2. Användarinvolverad utvärdering

Den användarinvolverade utvärderingen har bestått av empiriska observationer och intervjuer. Vi som anordnat studien har tagit rollen som observatörer och har också ställt frågorna under de intervjuer som genomförts. Utvärdering har skett vid tre skilda tillfällen med lika många deltagare. Under sessionerna, som skett med var och en av deltagarna individuellt, har de först fått prototypen framför sig på en dator och ombetts utföra ett antal uppgifter genom att interagera med gränssnittet. Deltagaren har samtidigt uttryckt hur denne tänkt utföra efterfrågad aktivitet genom s.k. think aloud-teknik och skeendet har observerats och dokumenterats. Därefter har deltagaren fått svara på hur hon upplevde prototypen under en kvalitativ intervju. Ljud och skärm har spelats in under både observation och intervju.

3.2.1. Deltagare

Nielsen (1994) menar att användbarhetstester vanligtvis genomförs med antingen personer som kan representera användare eller med faktiska slutanvändare. Det viktigaste är att systemet testas korrelerat till dess användare och i det syfte slutprodukten är ämnad att användas för. Vendium AB, som är initiativtagare till samt uppdragsgivare för studien, har haft som ansvar att hitta och

tillförse lämpliga deltagare till den användarinvolverade utvärderingen. De har presenterat ett antal namn samt lämpliga platser och tider där utvärderingar skett.

Även för den användarinvolverade utvärderingen har tre personer deltagit. Dessa deltagare är både användarrepresentanter och slutanvändare. Två av dem är väl insatta i systemet Struqtur, vars nya gränssnittsförslag är det som utvärderas, och en är användare av andra, konkurrerande, administrativa verktyg. En av deltagarna med erfarenhet av systemet Struqtur, som nedan presenteras som deltagare A2, har som yrke att dagligdags utbilda och tala med slutanvändare och har därför väldigt värdefulla kunskaper om vad de faktiska slutanvändarna efterfrågar och har svårigheter med. Deltagaren ansågs därför vara en person som kunnat representera systemets användare. Då deltagare A3 är en faktisk slutanvändare av systemet kan två av de tre deltagarna anses vara användarrepresentanter och en är en faktisk slutanvändare. Enligt Hertzum och Jacobsen (2003) krävs minst fyra till fem deltagare vid think aloud-studier för att kunna identifiera en majoritet av ett systems brister. Då enbart tre personer kunnat delta i utvärderingen har vi därför valt att komplettera utförda think aloud-sessioner med påföljande intervjuer. Själva utförandet kommer mer detaljerat behandlas i det kommande avsnittet "genomförande".

Precis som för den heuristiska utvärderingen har deltagarna fått varsin unik kod, A1, A2 och A3, som hädanefter kommer användas för hänvisning till den specifika utvärderarens intervjuvar, kommentarer och anmärkningar under observationstillfället. Nedan presenteras den användarinvolverade utvärderingens deltagare. Samtliga deltagare har samtyckt till att presenteras i nedan stående form samt låtit upptäckter som gjorts under observation och intervju redovisas i denna uppsats.

Deltagare A1:

Har låg erfarenhet om bygg- och hantverksbranschen men har ett års erfarenhet av att använda ett annat administrativt verktyg där bland annat offerter och fakturor hanteras. Är en representativ nybörjaranvändare av det system som utvärderas.

Deltagare A2:

Har under två år arbetat med användare till det nuvarande systemet Struqtur. Utbildar både nybörjare och mer avancerade användare, hanterar supportärenden samt säljer prenumerationer av systemet. Använder verktyget Teamviewer i arbetet och kan på så sätt övervaka användarnas interaktion under support-sessioner.

Deltagare A3:

Har flerårig arbetslivserfarenhet som projektledare i byggbranschen. Använder Struqtur regelbundet i sitt arbete.

3.2.2. Förberedelse

För att mäta ett systems grad av användbarhet med hjälp av användarrepresentanter är det brukligt att låta dem utföra ett antal på förhand bestämda uppgifter/aktiviteter i systemet (Nielsen 1994). Även Hertzum och Jacobsen (2003) förespråkar detta för studier som utförs med metoden think aloud, vilket görs i denna specifika studie. Föresättningsvis menar Hertzum och Jacobsen

(2003) att förberedelserna inför think aloud-studier består av tre aktiviteter som observatörerna behöver utföra. För det första behöver observatörerna bli bekanta med det system som skall utvärderas, för det andra skall de uppgifter som är ämnade att utföras av deltagarna bestämmas och för det tredje skall det värvas deltagare till studien. Värt att nämna är att det är vi som anordnat studien, och som författat denna uppsats, som tagit rollen som observatörer vid utförda observationer. Då uppdragsgivaren, Vendium AB, hade som ansvar att värva deltagare återstod de två förstnämnda aktiviteterna.

Ett timmeslångt möte sattes upp med en representant på Vendium AB för en genomgång av deras nuvarande system och de viktigaste aktiviteterna som användarna av systemet Struqtur utför. Detta möte hölls redan innan utvecklingen av den prototyp som utvärderats och visade sig vara väldigt viktigt för att få en grundlig inblick i hur systemet används. Genomgången dokumenterades genom inspelning av skärm och ljud då demonstrationen av systemet genomfördes på dator, på så vis har vi vid ett obegränsat antal tillfällen kunnat se genomgången.

Tolv användarscenarion har skrivits som var och ett går ut på att deltagarna skall utföra en relevant uppgift, eller aktivitet, i det gränssnitt som utvärderats. Ett antal olika förutsättningar ligger till grund för vilka scenarion som definierats och varför. Den första förutsättningen är att de uppgifter som vi som observatörer ansett vara relevanta enligt hållen genomgång av systemet Struqtur har definierats. En annan förutsättning är tidsaspekten, deltagarna har gått med på en begränsad tid om totalt 60 minuter som varje utvärderingstillfälle fått ta. Då varje utvärderingstillfälle består av inledande formalia och information, think aloud-observation samt påföljande intervju har antalet scenarion begränsats till tolv stycken. Tidsaspekten har i slutändan inte påverkat studien på så sätt att fler antal scenarion hade behövts för att täcka in samtliga relevanta aktiviteter. De tolv scenarion som togs fram ansågs vara tillräckliga för att utvärdera prototypen i sitt aktuella tillstånd. Den sista förutsättningen är utvecklingsstadiet för den prototyp som har utvärderats. Då prototypen är i ett tidigt utvecklingsstadium begränsar det möjligheten att interagera med gränssnittet på samma sätt som i ett färdigutvecklat system. Detta har medfört att scenarion som hade kunnat testats i ett mer utvecklat system, exempelvis deltagarnas förmåga att tolka feedback-dialoger, inte kunnat användas i denna studie.

Ett kontrakt har upprättats som med den enskilde deltagarens underskrift givit rätten till att spela in samt tolka och publicera det material som tillkommit vid utvärderingstillfället i denna uppsats. Kontraktet kan läsas i sin helhet i bilaga 2.

3.2.3. Genomförande

Den användarinvolverade utvärderingen har skett i sessioner med varje deltagare individuellt. Vid varje utvärderingstillfälle har en eller två observatörer också deltagit. Think aloud-studier börjar med att observatörerna informerar deltagarna om utvärderingen i sig och att de önskas "tänka högt" under sessionen (Hertzum & Jacobsen 2003). Kontraktet om medgivande till inspelning av utvärderingen har också lämnats över till deltagaren för underskrift.

Efter att kontrakt signerats och all formalia behandlats har skärm- och ljudinspelning startats för dokumentation av sessionen. Deltagarna har sedan fått prototypen på en dator framför sig och

har, av observatörerna, blivit tillbedda att granska och interagera med prototypen enligt de scenarion som tidigare förberetts. Vid uppstart av prototypen har deltagaren mötts av en sida som motsvarar Struqtur-systemets startsida i en ny design. Genom de olika scenarion som getts till deltagaren har hon fått navigera runt i systemet för att efter bästa möjliga förmåga utföra det som efterfrågats. När samtliga scenarion genomförts har en majoritet av prototypens element på samtliga av gränssnittets sidor på något sätt hanterats av deltagaren. Det är vanligt att de som observeras glömmar av att tänka högt under think aloud-sessioner (Hertzum & Jacobsen 2003), detta eftersom det inte faller sig naturligt för de flesta att tänka högt. Vid tillfällen då detta skett har observatörerna påtalat detta för deltagaren.

Direkt efter observation har en semistrukturerad intervju genomförts där allmänna frågor om den nyss testade prototypen ställts. Frågorna som använts vid intervjun har varit desamma vid de olika utvärderingstillfällena och deltagarna har fått svara fritt på varje fråga utifrån sina egna erfarenheter och preferenser. Följdfrågor som under intervjun blivit aktuella att ställa för att få ytterligare information från deltagaren angående vissa ämnen har också ställts. Hertzum och Jacobsen (2003) menar att det efter en think aloud-session är brukligt att hålla ett kortare samtal med deltagaren som på så sätt får chansen att uttrycka ytterligare tankar och åsikter om systemet. Som tidigare nämnts har en intervju som är mer utförlig och uttömmande genomförts för att väga upp för deltagargruppen som till antalet är en till två personer för liten. De som ställt frågorna under intervjun är liksom vid observationen vi som författat denna uppsats. Syftet med intervjun har varit att utröna hur användaren upplevde gränssnittet och främst vad hon upplevde var negativt för att på så sätt kunna identifiera eventuella brister som varit svåra att upptäcka under observationen. Deltagarna har under intervjun fått chansen att demonstrera sina förklaringar genom att i prototypen på skärmen framför sig visa vad han eller hon menar. Då skärm- och ljudinspelning fortgått under hela sessionen har även detta dokumenterats på samma sätt som under observationen.

Inför analys av observationer och intervjuer har det inspelade materialet transkriberats. Transkribering har skett genom att filer med det som spelats in från skärmen och ljud som upptagits under de olika utvärderings-sessionerna matats in i ett webbaserat transkriberingsverktyg. Med verktyget är det möjligt att starta, stoppa, spola och förändra hastigheten i video-filen samtidigt som det är möjligt att skriva, vilket gör transkriberingsmomentet så effektivt som möjligt.

3.3. Analys

Efter genomförda utvärderingar har resultaten samlats in, organiserats och analyserats. Brister har identifierats och därefter klassificerats samt förts in i tabeller. Under den heuristiska utvärderingen har testarna själva angett vad de uppfattat som brister, däremot har visst arbete krävts för att identifiera brister från den användarinvolverade utvärderingen. I de transkriberade texterna från observationer och intervjuer har allt som kunnat identifieras som en brist noterats. Dessutom har inspelat video-material från genomförda observationer granskats för att om möjligt identifiera brister utifrån deltagarens interaktion i gränssnittet. En brist har noterats när deltagaren antingen muntligt uttalat att hon haft svårigheter att utföra aktuell uppgift eller om hon inte

effektivt lyckats utföra den uppgift som efterfrågats. Samtliga identifierade brister från de båda utvärderingarna har antecknats och förts över till nästa moment, klassificering.

3.3.1. Klassificering av brister

Klassificering av brister har gjorts för att kunna se vilka olika typer av brister som hittats under utvärderingarna. Syftet med klassificeringen har således varit att standardisera den data som uppkommit av utvärderingarna samt möjliggöra analys och jämförelse av de båda resultaten. Klassificeringsmomentet har varit en del av analysarbetet och har inte använts aktivt av de deltagande testarna i den heuristiska utvärderingen eller under den användarinvolverade utvärderingen. Tre kategorier har definierats som var och en representerar ett område i det gränssnitt som testats däri bristen har sitt ursprung. Kategorierna har inte någon korrelation med de heuristiker som under de heuristiska utvärderingarna fungerat som stöd för testarna. Varje enskild brist har under klassificeringsmomentet tilldelats en av de tre nedan definierade kategorierna:

Intuitivitet

Huruvida användaren omedelbart kan förstå hur interaktion med ett gränssnitt bör ske, hur hon bör handla för att uppnå önskvärt resultat, utan behov av att experimentera, få support eller utbildning.

Layout

Huruvida placering och design av objekt samt element möjliggör för användaren att arbeta effektivt i ett gränssnitt.

Visualisering av information

Huruvida användaren med lätthet kan förstå och tolka information i ett gränssnitt.

Klassificeringsmomentet består även av att bedöma graden av allvarlighet för varje enskild brist. Genom att tilldela varje enskild brist en grad av allvarlighet har resultatet även kunnat visa hur många brister av olika allvarlighetsgrad som upptäckts med respektive utvärderingsmetod. Under denna bedömning har en tregradig skala gällande allvarlighet använts där allvarlighetsgrad 1 ses som mycket allvarlig, allvarlighetsgrad 2 ses som allvarlig och allvarlighetsgrad 3 ses som mindre allvarlig. Då Doubleday et al (1997) och Tan, Liu och Bishu (2009) i sina studier har anammat tregradiga skalor när de bedömt olika bristers allvarlighetsgrad ansågs en tregradig skala lämplig också för denna studie. Att använda en tregradig skala har på så vis också möjliggjort en högre grad av jämförbarhet med dessa tidigare studier.

Hur klassificeringen praktiskt gått till skiljer sig mellan de olika utvärderingsmetoderna. Vid den heuristiska utvärderingen har testarna själva tilldelat varje brist en allvarlighetsgrad som respekterats och kvarstått till klassificeringen. Heuristiken, tillsammans med bristens beskrivning, har legat till grund för vilken av ovanstående tre kategorier bristen tilldelats. Utöver detta steg har de av testarna angivna heuristikerna inte använts i resultatet då syftet snarare varit att använda dem som ett stöd för att underlätta identifiering av brister vilket är utmärkande för metoden. Då det inte använts något motsvarande verktyg under den användarinvolverade utvärderingen har det

därför varit mer relevant att klassificera bristerna inom ovan nämnda kategorier för att möjliggöra jämförelse av metodernas resultat.

De brister som identifierats genom den användarinvolverade utvärderingen har klassificerats utefter hur vi som observerat och intervjuat deltagarna uppfattat de olika bristerna. Om deltagaren exempelvis haft stora svårigheter att utföra ett visst scenario under observationen har en högre allvarlighetsgrad tilldelats. Vilken kategori bristen tilldelats har berott på om deltagaren under observation eller intervju visat på/påtalat svårigheter att förstå hur hon skall gå tillväga (intuitivitet), hämmats i sitt användande p.g.a. bristande design och layout (layout) eller inte kunnat ta till sig information på ett bra sätt (visualisering av information).

För de brister som identifierats under båda utvärderingarna och som varit identiska har det ibland uppstått ett problem när bedömningen av bristens allvarlighetsgrad inte varit den samma. När olika bedömningar gjorts har den grad av allvarlighet som angivits under den heuristiska utvärderingen använts. Denna lösning ansågs vara lämplig då testarna som deltagit i den heuristiska utvärderingen har högre praktisk och akademisk erfarenhet än vad observatörerna själva har.

För att organisera resultaten har bristerna i samband med klassificering förts in i tabeller som skapats utifrån ovan nämnda kategorier. Då resultatanalys av respektive utvärdering till en början skett individuellt har tre tabeller per utförd utvärdering skapats. Således har kategorierna intuitivitet, layout och visualisering av information fått var sin tabell för brister som identifierats under den heuristiska utvärderingen och motsvarande tabeller har skapats för den användarinvolverade utvärderingen. Vilken eller vilka testare och eller användare som upptäckt bristen har angetts i en kolumn, i övriga kolumner har bristens beskrivning, antal förekomster och allvarlighetsgrad angetts (se tabell 1 nedan för exempel). Det är inte alltid testaren eller användarens beskrivning av bristen som angetts i tabellerna då det ibland förekommer tolkningar och konkretiserade beskrivningar som gjorts av oss som anordnat studien. Detta har gjorts då de beskrivningar som angetts av testarna eller användarna ibland varit detaljerade eller bestått av ytterligare information som exempelvis lösningsförslag och därför inte lämpat sig för ett tabellformat där syftet varit att ställa upp resultatet på ett överskådligt vis.

Tabell 1. Exempeltabell för brister inom kategori intuitivitet som identifierats under heuristisk utvärdering

Intuitivitet			
Beskrivning	Allvarlighetsgrad	Förekomster	Upptäckt av ID
En beskrivning av bristen	2	2	H2, H1
Totalt antal brister:	1		
Antal brister av allvarlighetsgrad 1:	0		
Antal brister av allvarlighetsgrad 2:	1		
Antal brister av allvarlighetsgrad 3:	0		
Antal brister upptäckta av alla:	0		
Antal brister upptäckta av 2 personer:	1		

Nederst i tabellerna har sammanfattande statistik skrivits ut. Att kunna beskåda resultatet systematiskt uppställt i tabeller tillsammans med statistik har underlättat analysarbetet där syftet varit att komma fram till vilka typer av brister som uppkommer vid respektive metod. Tabeller som sammanställt hela utvärderingens resultat i mer sammanfattande form har också tagits fram för att underlätta analysarbetet ytterligare.

3.3.2. Tolkning av data

Under analysarbetet har klassificerad och sammanställd data från de båda utvärderingarna initialt granskats och redovisats individuellt, utan att jämföra utfallen mellan de bägge utvärderingsmetoderna. Tabellerna har studerats för att utröna hur resultatet skiljer sig på ett översiktligt plan inom respektive utvärdering. Hur många brister som förekommit inom varje kategori, hur många brister som upptäckts av flera testare inom de olika kategorierna samt andra intressanta utfall och avvikelser har sammanställts och redovisas i resultat-kapitlet av denna uppsats.

En kvalitativ analys av bristerna har också gjorts där de inskickade formulären från den heuristiska utvärderingen och materialet från den användarinvolverade utvärderingen (transkiberade texter, ljud- och video-inspelningar) har studerats. Detta för att möjliggöra för en djupare förståelse om hur deltagaren upplevt sin interaktion med prototypen och vilken information som faktiskt finns i det dokumenterade materialet. Bristerna har i den form de beskrivs av varje enskild deltagare granskats ordagrant för att kartlägga all den information som getts. Hur deltagarna i den användarinvolverade utvärderingen uppträtt har studerats noggrant genom granskning av den inspelade dokumentationen. Detaljrikedomen i beskrivningen av bristen, om det lämnats ett lösningsförslag och om en motivering till varför bristen faktiskt är en brist är exempel på fenomen som studerats. Inom resultatet för respektive utvärdering har också upptäckter jämförts inbördes, har exempelvis ett element i gränssnittet som upplevts som problematiskt av en deltagare upplevts som väldesignat och användarvänligt av en annan deltagare? Resultatet av denna analys redovisas i kapitlet resultatanalys i denna uppsats.

3.4. Jämförelse

Som tidigare nämnts har det insamlade materialet från respektive utvärdering till en början analyserats var för sig, utan att sätta det i jämförelse till den andra utvärderingen. Analys av respektive utvärdering enligt ovan har gjorts med syfte att ge kunskap om vilken typ av designbrister som påpekats, vilka andra typer av upptäckter som gjorts av deltagarna, vilken positiv respektive negativ feedback som framförts samt förbättringsförslag på gränssnittets design. Den sammanställda informationen om de olika metodernas utfall, tillsammans med upptäckter från den kvalitativa analys som gjorts har sedan använts som underlag för att jämföra metoderna och svara på frågeställningen. Förutom att söka ett direkt svar på frågeställningen, *“vilka typer av designbrister som upptäcks vid heuristisk respektive användarinvolverad utvärdering av ett webbgränssnitt i prototypstadiet”*, har också andra betydande upptäckter om de båda metodernas likheter och skillnader noterats under studien och redovisas i denna uppsats.

En redovisning av utvärderingsmetodernas skillnader och likheter har gjorts samt en analys kring vad man som utvecklare kan förvänta sig vid användning av respektive metod till utvärdering av tidiga prototyper för webbgränssnitt. Hur väl resultatet stämmer överens med tidigare teori och forskning om utvärdering har också diskuterats för att ge förslag på hur forskningen kan användas som stöd vid utvärdering av tidiga webbgränssnittsprototyper.

4. Resultat

I detta kapitel redovisas de brister som upptäckts med respektive utvärdering, vilket inkluderar bristernas grad av allvarlighet, kategori samt övriga relevanta upptäckter och sammanställningar.

4.1. Heuristisk utvärdering

De heuristiska utvärderingarna resulterade i att 24 stycken olika brister i gränssnittet upptäcktes. Bristerna redovisas i sin helhet utefter kategori i tabell 2, 3 och 4. I tabellerna presenteras respektive brist med dess beskrivning, grad av allvarlighet, antal testare som upptäckte bristen samt vilka testare som gjorde detta.

Tabell 2. Heuristisk utvärdering - Brister inom intuitivitet

Intuitivitet			
Beskrivning	Allvarlighetsgrad	Förekomster	Upptäckt av ID
Svårt att förstå plustecken i menyn, saknas vid tidrapport	2	2	H2, H1
Oklart huruvida det går att interagera med entiteterna i rutorna på startsidan. Skulle kunna vara tydligare	2	1	H2
Oklart huruvida länkar i meny och på startsidan leder till samma destination	3	1	H1
Byt ordet "visa" mot "filtrera" på sidan för tidrapporter då detta mer förklarar vad funktionen gör	3	1	H1
Oklart vilken data som ska fyllas i vid tidrapportering - Bör finnas någon typ av info, exempelvis tooltips eller fler dropdown alternativ	2	2	H1, H2
Menyn känns avskild från innehållet och glöms lätt	3	1	H1
Ingen form av feedbacksystem i prototypen när ett case utförs samt i under-menyn saknas visuell feedback på vart man befinner sig.	1	2	H3, H2
Ikoner på startsidan återanvänds inte när man klickar sig in till den sida som det länkas till	1	1	H3
Man får för sig att det ska gå att klicka på en hel rad i tabellen på sidan för projektöversikt, detta går inte.	3	1	H2
Totalt antal brister:	9		
Antal brister av allvarlighetsgrad 1:	2		
Antal brister av allvarlighetsgrad 2:	3		
Antal brister av allvarlighetsgrad 3:	4		
Antal brister upptäckta av alla:	0		
Antal brister upptäckta av 2 personer:	3		

Nio brister upptäcktes inom denna kategori varav de flesta var av mindre allvarlig art. Ingen brist upptäcktes av samtliga testare. Tre brister upptäcktes av två testare, exempelvis uppmärksammade både testare H1 och H2 att funktionen för tidrapportering är otydlig och efterfrågar mer stöd och hjälp till användaren.

Tabell 3. Heuristisk utvärdering - Brister inom layout

Layout			
Beskrivning	Allvarlighetsgrad	Förekomster	Upptäckt av ID
Olämplig placering av tidrapportsfunktionen på sidorna för specifika projekt. Bör hanteras för sig självt under tidrapports-fliken	1	1	H2
För många kort och mycket information på startsidan. Skulle kunna särskiljas, exempelvis med färger	2	1	H2
Knappen för att enbart visa projekt nära till deadline på projektöversikten är överflödigt vid möjligheten att sortera tabellen (vilket ännu inte är implementerat).	3	1	H2
Ikonen för att hantera projekt är för liten och är svår att lägga märke till, inte självklart vad den gör	1	1	H1
Färger - gul ser ut som en varning	2	1	H1
Hjälpavsnitt brukar i regel finnas uppe till höger i gränssnitt, så bör det göra även i detta. T. ex. Frågetecknet i projektöversikten	2	1	H1
Placera den mest relevanta informationen i fokus	1	1	H1
För liten text generellt i gränssnittet	2	1	H1
För mycket skrollande fram och tillbaka	2	1	H1
Många boxar på projektsidan. Ingen möjlighet att välja vilka man vill se och skala ner. Nu kan man bara se allting hela tiden	2	1	H3
Totalt antal brister:	10		
Antal brister av allvarlighetsgrad 1:	3		
Antal brister av allvarlighetsgrad 2:	6		
Antal brister av allvarlighetsgrad 3:	1		
Antal brister upptäckta av alla:	0		
Antal brister upptäckta av 2 personer:	0		

Bland de tre kategorierna upptäcktes flest antal brister inom kategorin layout. Av dessa brister är samtliga upptäckta av endast en testare. Allvarlighetsgraden är förhållandevis hög då enbart en av bristerna anses vara av mindre allvarlig art.

Tabell 4. Heuristisk utvärdering - Brister inom visualisering av information

Visualisering av information			
Beskrivning	Allvarlighetsgrad	Förekomster	Upptäckt av ID
Otydlig statusindikator för hur långt ett projekt har kvar innan deadline. Framstår som att det menar hur mycket som är kvar att göra i projektet och inte antalet dagar som är kvar. Förklaring av färgerna bör ej vara gömt	2	2	H1, H2
Statusindikatorn som nu består av fyra olika färger skulle kunna ersättas av en färskala för att bättre visualisera projektets åtgång.	3	1	H2
Otydlig eller icke existerande koppling mellan aktiviteter och utfört arbete	2	1	H1
Introduktion / hjälpavsnitt bör finnas för förstagångsbesökare när systemet startas	3	1	H1
För mycket information på projektsida, tar tid att få grepp om informationen	1	1	H1
Totalt antal brister:	5		
Antal brister av allvarlighetsgrad 1:	1		
Antal brister av allvarlighetsgrad 2:	2		
Antal brister av allvarlighetsgrad 3:	2		
Antal brister upptäckta av alla:	0		
Antal brister upptäckta av 2 personer:	1		

Visualisering av information var den kategori där minst antal brister upptäcktes. En av bristerna upptäcktes av två testare och berörde statusindikatorerna för olika projekt i systemet vilka ansågs vara otydliga. En brist är av mycket allvarlig art och angår projektsidan där det anses vara för mycket information som presenteras för användaren vilket gör det svårt att ta den till sig.

Av de brister som hittats under den heuristiska utvärderingen ansågs sex stycken vara av mycket allvarlig art. Fullständig fördelning av bristerna utefter grad av allvarlighet presenteras i tabell 5.

Tabell 5. Heuristisk utvärdering - fördelning av brister utefter allvarlighetsgrad

Totalt antal	24
Mycket allvarliga	6
Allvarliga	11
Mindre allvarliga	7

Variationen av upptäckter mellan de enskilda testarna var hög, exempelvis upptäcktes ingen av bristerna av samtliga testare och enbart fyra av bristerna upptäcktes av mer än en testare. Antalet upptäckter per testare varierade från tre till femton.

Brister som tilldelats allvarlighetsgrad mycket allvarlig har i störst utsträckning berört det strukturella upplägget av prototypens design. Dessa brister gäller exempelvis olämplig placering av viktig information, bristfällig visuell feedback gällande vart användaren befinner sig i systemet samt att ikoner inte används konsekvent genom hela prototypen. Brister av lägre allvarlighetsgrad är även dessa till största del av en övergripande karaktär i form av att gränssnittet kräver för mycket "skrollande" och att menyn upplevs som avskild från det faktiska innehållet på sidan. Det förekommer också ett antal brister gällande enskilda detaljer i gränssnittet och det är även bland dessa brister det råder störst samstämmighet bland testarna. Exempelvis upplever både testare H1 och testare H2 att plustecknet vid vissa menyalternativ är otydligt och ej självförklarande. De noterade även att statusindikatorn på sidan för projektöversikt är otydlig och borde designas annorlunda alternativt förklaras bättre.

Klassificering av funna brister utefter kategorierna intuitivitet, layout och visualisering av information visade att flest brister, tio stycken, var av typen layout. Fullständig fördelning av bristerna utefter kategori av brist presenteras i tabell 6.

Tabell 6. Heuristisk utvärdering - fördelning av brister utefter kategori av brist

Intuitivitet	9
Layout	10
Visualisering av information	5

Inga av de upptäckta bristerna står i motsättning till varandra och vid sju av bristerna gavs även ett förslag på hur problemet kan åtgärdas till senare versioner av gränssnittet.

4.2. Användarinvolverad utvärdering

Think aloud-sessionerna tillsammans med efterföljande intervjuer resulterade i att 30 olika brister upptäcktes. Bristerna redovisas i sin helhet utefter kategori i tabell 7, 8 och 9. I tabellerna presenteras respektive brist med dess beskrivning, grad av allvarlighet, antal användare som upptäckte bristen samt vilka användare som gjorde detta.

Tabell 7. Användarinvolverad utvärdering - Brister inom intuitivitet

Intuitivitet			
Beskrivning	Allvarlighetsgrad	Förekomster	Upptäckt av ID
Relevant information som man behöver scrolla till blir uttömmande för besökaren som ibland fastnar på den första informationen. Gränssnittet visar inte vart man skall.	1	1	A1
I tabellen på projektöversikt skiftar raderna färg när musen hålles över dessa. Detta antyder att man skall kunna klicka på raden, något som inte går.	3	1	A1
Det framgår inte att det finns en kalenderfunktion i fältet för att välja datum.	2	1	A1
Försöker jobba enligt gammal vana när det gäller tidrapporterna. Försöker gå in på rapportgeneratoren i stället för att använda den nya översiktsrutan.	3	1	A3
Huvudmenyn (den orange) har en något ologisk ordning. Exempelvis bör Offert, Projekt och Faktura ligga bredvid varandra.	3	1	A3
Datumväljare - Utgår inte från dagens datum	3	1	A2
Datumväljare - Interaktions- och navigeringssvårigheter	3	1	A2
Tidrapportering - Projekt förvalt, borde stå "välj projekt" exempelvis	2	1	A2
Tidrapportering - Otydligt om antal dagar kan genereras utifrån inrapporterade timmar	2	1	A2
Man bör kunna namnge/märka projekt där namnet också ses som en rubrik på projektsidan	1	1	A2
Tidrapportering bör vara tydligare kopplat med aktiviteter på enskild projektsida	1	1	A2
Projekt - Aktiviteter. Går inte att hantera vilken anställd som utför aktiviteter	2	1	A2
Totalt antal brister:	12		
Antal brister av allvarlighetsgrad 1:	3		
Antal brister av allvarlighetsgrad 2:	4		
Antal brister av allvarlighetsgrad 3:	5		
Antal brister upptäckta av alla:	0		
Antal brister upptäckta av 2 personer:	0		

Av de tolv brister som upptäcktes inom kategori intuitivitet upptäcktes ingen av mer än en användare. Graden av allvarighet var relativt jämnt fördelad och bristerna berör till stor del specifika element, exempelvis datumväljaren, i gränssnittet.

Tabell 8. Användarinvolverad utvärdering - Brister inom layout

Layout			
Beskrivning	Allvarlighetsgrad	Förekomster	Upptäckt av ID
Liten knapp på projektöversikten för att hantera projekt som dessutom ligger för sent (för långt till höger) i tabellen.	1	3	A1,A2,A3
Checkboxen för arbetsmoment i rutan för tidrapportering var svår att se	3	1	A1
Planerade/Pågående aktiviteter är en viktig del och således för långt ned på sidan.	1	1	A1
På startsidan tar i synnerhet mail och kommande händelser stor plats, i allmänhet gör även de andra rutorna på denna sida det	3	1	A3
Det finns en begränsning i att man enbart kan hantera ett projekt i taget	1	1	A3
Går inte att välja flera anställda i rullistan för rapporterade tider på sidan för tidrapport.	1	1	A3
Startsidan är något rörig och känns inte tillräckligt välkomnande eller informativ.	2	1	A3
Knappar, Text, Färg - För liten storlek på knappar och text på vissa ställen. Blir svårt att läsa, speciellt på en surfplatta utomhus. Även färgvalen gör det svårare att läsa ibland.	1	2	A1,A2
Tidrapportering - Saknar funktion för att gå in på en specifik dag för att kunna se samt redigera uppgifter på den dagen. Det finns men framgick ej.	2	1	A2
Projekt - Funktionaliteten för att tidrapportera upplevs som malplacerad. Användaren får en känsla av att vara in på tidrapportering när denne önskar att se utfört arbete och lägga till materialkostnader.	1	1	A2
Projekt - Utfört arbete, material och övriga kostnader bör vara det första användaren ser då det är detta som önskas arbetas med på denna sida.	2	1	A2

Projekt - Rapporterade dagar är inte sida vid sida med materialkostnader vilket gör att användaren måste scrola fram och tillbaka för att få en överblick.	2	2	A2,A3
Projekt - Bristfällig placering av olika element. Önskar att det mer dagliga sysslorna placeras överst medans sådant som administreras i efterhand placeras nederst.	1	2	A2,A3
Totalt antal brister:	13		
Antal brister av allvarlighetsgrad 1:	5		
Antal brister av allvarlighetsgrad 2:	6		
Antal brister av allvarlighetsgrad 3:	2		
Antal brister upptäckta av alla:	1		
Antal brister upptäckta av 2 personer:	3		

Samtliga användare uppmärksammade att knappen för att hantera projekt på projektöversikten är för liten och olämpligt placerad. Detta är den enda brist som upptäckts av samtliga användare. De två rödmarkerade raderna i tabell 8 indikerar att olika användare identifierat brister som rör samma gränssnittselement. Eftersom användarna upplever att de hämmas i sitt användande på olika sätt står bristerna i konflikt med varandra; en lösning till den ena bristen skulle innebära att den andra kvarstår och vice versa.

Tabell 9. Användarinvolverad utvärdering - Brister inom visualisering av information

Visualisering av information			
Beskrivning	Allvarlighetsgrad	Förekomster	Upptäckt av ID
Mycket och omfattande information gör det svårt att hitta specifik information direkt på sidan för specifikt projekt.	1	1	A1
Tar tid för deltagaren att hitta hur många dagar det är kvar till deadline på projektöversikten.	2	1	A1
Den blåa färgen under status i tabellen på projektöversikt förvirrar något. Bättre eller sämre än grön?	3	1	A1
Projektaktiviteter saknar bättre visualisering och hantering.	2	1	A2
Projekt - Aktiviteter. Visa som kalender eller "schema" för aktiviteterna för bättre överblick.	2	1	A2
Totalt antal brister:	5		
Antal brister av allvarlighetsgrad 1:	1		
Antal brister av allvarlighetsgrad 2:	3		
Antal brister av allvarlighetsgrad 3:	1		
Antal brister upptäckta av alla:	0		
Antal brister upptäckta av 2 personer:	0		

I kategori visualisering av information var bristerna till största del av en mer specifik art och berör exempelvis statusindikatorerna på projektöversikten vilka användare A1 upplevde som förvirrande och svåra att avläsa.

Av de brister som identifierats av den användarinvolverade utvärderingen ansågs sex stycken vara av mycket allvarlig art. Fullständig fördelning av bristerna utefter grad av allvarlighet presenteras i tabell 10.

Tabell 10. Användarinvolverad utvärdering - fördelning av brister utefter allvarlighetsgrad

Totalt antal	30
Mycket allvarliga	9
Allvarliga	13
Mindre allvarliga	8

Det förekommer stor variation i upptäckter mellan de olika användarna. Av 30 olika brister upptäcktes tre stycken av två användare medan en brist upptäcktes av samtliga. Den brist som upptäcktes av samtliga användare var av allvarlig karaktär och berörde knappen för att hantera projekt på projektöversikten vilken användarna upplevde som för liten till storlek.

Bristerna är till största del av en relativt detaljerad och specifik karaktär. Exempelvis berör bristerna områden så som att datumväljaren inte utgår ifrån dagens datum, att mail och

kommande händelser tar upp för mycket plats på startsidan samt att vissa färgval gör text svårsläst. Vad gäller datumväljaren upplevde två av användarna denna som bristfällig och noterade inte omedelbart att det går att klicka fram en kalender där man kan välja datum. En av användarna uttryckte sig på följande vis på frågan om huruvida det framgick för deltagaren att denna funktion existerade:

“Nej det framgår inte. Det måste ni försöka visa bättre med någon ikon eller fyrkant så man förstår.” (intervju, användare A1)

Att detta inte framgick ledde till att de båda användarna upplevde interaktionen vid inmatning av datum som problematisk då de inledningsvis fyllde i datum manuellt i textfältet i stället för att använda datumväljaren.

Storlek på knappar och text, samt vissa färgval ansågs skapa problem enligt två användare och kommenteras bland annat på följande vis:

“Det som är ett problem, men den vita texten på den gula knappen ser man inte utomhus när man arbetar på en padda. Den är klurig. Man jobbar ofta i systemet utomhus.”
(intervju, användare A2)

“Vidare på projektöversikten hade jag velat ha knappen för hantera projekt mycket större.” (intervju, användare A1)

Genomgående under både think aloud-observationerna och intervjuerna refererade användarna ett flertal gånger till hur det nuvarande systemet (Struktur) ser ut och poängterade sådant som de upplevde var både sämre och bättre i förhållande till prototypen. Användarna utgick till viss del ifrån att prototypen skulle fungera på samma sätt som det nuvarande systemet vilket ibland ledde till felaktiga navigeringar. Följande citat illustrerar en av användarnas avsaknad av att kunna arbeta med flera projekt samtidigt, samt möjligheten att se de senast uppladdade bilderna på startsidan (funktioner som existerade i det gamla systemet men inte i prototypen):

“Ja det är i min värld en klar nackdel. Att jag inte kan jobba i två projekt samtidigt.”
(intervju, användare A3)

“[...] för när man loggar in så brukar jag kolla för att se om dom lägger upp bilder för att se vad som händer och så. Jag är inte aktivt delaktig i alla projekt men det är kul att se lite granna vad som händer i alla projekt. Det är bara kul att liksom scrolla igenom dom.” (intervju, användare A3)

Författarnas kategorisering av bristerna utefter typ av brist, intuitivitet, layout eller visualisering av information, presenteras i tabell 11. Merparten av bristerna kategoriserades som brister inom layout och intuitivitet.

Tabell 11. Användarinvolverad utvärdering - fördelning av brister utefter kategori av brist

Intuitivitet	12
Layout	13
Visualisering av information	5

I samband med uppmärksammandet av en brist föreslog användarna vid nio tillfällen en lösning på problemet.

5. Resultatanalys

I följande kapitel redogörs för den analys av resultatet som genomförts. Analysen presenteras för respektive utvärdering var för sig, dock kan vissa kopplingar och jämförelser mellan utvärderingarna förekomma där detta upplevts relevant för att beskriva sammanhang och samband.

5.1. Heuristisk utvärdering

Den heuristiska utvärderingen resulterade i att 24 olika brister upptäcktes och ingen av dessa brister upptäcktes av samtliga testare. Samtidigt var det enbart fyra brister som upptäcktes av två testare. Denna variation kan bero på att testarna har haft stora friheter gällande hur de genomfört utvärderingen. Då författarna inte givit testarna några instruktioner gällande exempelvis specifika element eller uppgifter att utföra i systemet kan detta ha inneburit att testarna koncentrerat sig på olika områden vid sin utvärdering vilket lett till dessa resultat. Variationen kan även bero på att testarnas personliga åsikter och erfarenheter kan påverka vad de upptäcker och vad de väljer att rapportera, vilket enligt Hertzum och Jacobsen (2003) bör tas i beaktande vid användning av metoden. Resultatet innebär att varje enskild testare bidragit med upptäckter som de andra inte har gjort.

5.1.1. Intuitivitet

Inom kategori intuitivitet upptäcktes två brister av högsta allvarlighetsgrad som ej uppmärksammades av den användarinvolverade utvärderingen. Bägge dessa brister berör generella problem med prototypens utförande och beskrivs på följande vis:

Ingen form av feedbacksystem i prototypen när ett case utförs samt i under-menyn saknas visuell feedback på vart man befinner sig. (testare H3, H2)

Vilket syftar till att det inte förekommer någon form av feedback gällande huruvida en handling i systemet faktiskt utförs vid knapptryck samt att det i menyn inte alltid markeras vad användaren för tillfället har valt för alternativ. Samt:

Ikoner på startsidan återanvänds inte när man klickar sig in till den sida som det länkas till. (testare H3)

Denna brist påpekar att de ikoner som används för bland annat projekt och tidrapportering endast förekommer på prototypens startsida och aldrig återanvänds på någon annan sida.

Bristerna beskriver därmed övergripande generella problem vilka kan leda till förvirring hos en användare då denne inte alltid kan tyda vilket menyalternativ som är valt samt saknar möjlighet till igenkänning av ikoner. Då den användarinvolverade utvärderingen inte upptäckte dessa brister stämmer resultaten överrens med vad Doubleday et al (1997) och Tan, Liu och Bishu (2009) kom fram till i sina studier gällande att heuristisk utvärdering i större utsträckning upptäcker övergripande och strukturella problem. Gällande brister av högsta allvarlighetsgrad inom de övriga kategorierna skiljer sig dock denna studies resultat då samtliga av dessa brister även upptäcktes vid den användarinvolverade utvärderingen.

5.1.2. Layout

Noterbart gällande denna kategori är att testarna nämner en del brister som beror på avsaknad av en viss funktion eller möjlighet. De ger även förslag på vad som kan läggas till eller förändras för att åtgärda problemet. Testare H3 föreslår exempelvis att användaren bör ha möjligheten att dölja de element i gränssnittet som denne inte önskar arbeta med:

Många boxar på projektsidan. Ingen möjlighet att välja vilka man vill se och skala ner. Nu kan man bara se allting hela tiden. (testare H3)

Testare H1 föreslår att gränssnittets hjälpavsnitt placeras uppe till höger för att på så vis följa en mer etablerad standard:

Hjälpavsnitt brukar i regel finnas uppe till höger i gränssnitt, så bör det göra även i detta. T. ex. Frågetecknet i projektöversikten (testare H1)

Doubleday et al (1997) menar att heuristiska utvärderingar i större utsträckning upptäcker de strukturella problemen och att en användarinvolverad utvärdering uppmärksammar symptomen av dessa strukturella problem. Detta kan noteras i denna studies resultat då testare H1 menar att det som är mest relevant för användaren bör placeras på ett tydligt vis:

Placera den mest relevanta informationen i fokus (testare H1)

Samtidigt uppmärksammade samtliga användare att vissa objekt inte var placerade där de hade önskat och att detta tvingade dem till omständlig navigering (scrollning) upp och ned på sidorna. Användare A3 illustrerade detta på följande vis:

Och när jag klickar upp projektet så får jag dagarna och materialet och det är generellt sett det jag håller på med i den projektvyn och det kommer jag väldigt tydligt till. Nu kommer dem (klickar på ett projekt och visar "arbetade tider" och "material") såhär och jag kan inte arbeta i dem utan att scrolla mellan dem upp och ner. (intervju, användare A3)

Att samtliga användare reagerade på placering av objekt och information medan endast en heuristisk testare gjorde detta kan bero på att de heuristiska testarna inte själva är användare av det nuvarande systemet Struqtur och därmed saknar praktisk erfarenhet av vilka objekt som i störst utsträckning används vid dagligt arbete.

5.1.3. Visualisering av information

Inom denna kategori upptäckte de heuristiska testarna dels övergripande brister och dels detaljerade brister gällande specifika element. Testare H1 och H2 uppmärksammar exempelvis att statusindikatorerna på sidan för projektöversikt är otydliga och borde designas annorlunda alternativt förklaras mer utförligt av hjälpfunktionalitet i gränssnittet. Under observation av en användare noterades även att det dröjde ett tag innan denne lyckades utläsa information från dessa statusindikatorer. Användaren reagerade därmed på symptomen av den brist som testare H1 och H2 uppmärksammade, vilket stödjer resonemanget av Doubleday et al (1997) gällande att användare i större utsträckning upptäcker brister som är en symptom av en annan brist. I det här fallet ledde avsaknaden av hjälpfunktionalitet till att användaren upplevde det som omständligt att utläsa information från statusindikatorerna.

Testare H1 uppmärksammar två mer övergripande brister i form av att projektsidan innehåller för mycket information för att det ska vara lättillgängligt samt att det saknas ett introduktionsavsnitt för förstagångsbesökare:

För mycket information på projektsida, tar tid att få grepp om informationen (testare H1)

Introduktion / hjälpavsnitt bör finnas för förstagångsbesökare när systemet startas (testare H1)

I förhållande till övriga kategorier var antalet brister inom visualisering av information lågt och endast en av bristerna upptäcktes av mer än en testare.

5.2. Användarinvolverad utvärdering

Resultatet av den användarinvolverade utvärderingen var att 30 olika brister upptäcktes. Av dessa brister upptäcktes en av samtliga användare och tre brister upptäcktes av två användare. Till skillnad från de testare som utförde den heuristiska utvärderingen har användarna varit mer styrda vad gäller sin navigering i prototypen och samtliga användare har fått genomföra samma scenarion i gränssnittet. Därav kan den stora variationen i upptäckter inte förklaras med att användarna skulle ha utfört olika uppgifter i systemet och därmed fokuserat på varierande aspekter av det. Två potentiella förklaringar till varför variationen är så stor är dels möjligheten att användare inte har nämnt allt de upplevt som problematiskt med gränssnittet, alternativt att observatörerna inte lyckats uppmärksamma alla brister i samband med observationen. En tredje möjlig förklaring kan vara att antalet brister i gränssnittet är många till antalet och att användarna fokuserat sin uppmärksamhet på olika typer av brister. Hade studien genomförts med ett större antal användare hade det eventuellt varit möjligt att observera en brytpunkt där sannolikheten att ytterligare en deltagande användare upptäcker nya brister är låg.

5.2.1. Intuitivitet

Inom denna kategori upptäcktes ett antal brister av relativt specifik art gällande olika objekt och element i gränssnittet. Datumväljaren uppmärksammades exempelvis av två användare som upplevde att den var svår att identifiera i gränssnittet vilket ledde till interaktionssvårigheter vid inmatning av datum. Användare A2 upplevde även inmatningsfälten för tidrapportering som något bristfälliga och beskrev detta på följande vis under intervjun:

Det som är konstigt här som jag nämnde att det ligger projekt förvalda och att man inte förstår "dagar att rapportera", gör inte jag i alla fall. (intervju, användare A2)

Vid observation av användare A3 uppmärksammades att denne försökte arbeta med tidrapporteringar så som det utförs i det existerande gränssnittet vilket ledde till att dessa scenarion tog längre tid. Att försöka arbeta enligt det nuvarande gränssnittet (Struktur) samt jämföra och referera till hur det gränssnittet fungerar var vanligt förekommande bland användarna. Detta tyder på att användarna har en inövad kunskap om det nuvarande gränssnittet och att det kan vara olämpligt att genomföra omfattande förändringar vad gäller namnen på, och utformningen av, vissa funktioner.

5.2.2. Layout

De brister som upptäcktes inom denna kategori var av förhållandevis allvarlig grad. En av bristerna kan anses som särskilt allvarlig då den leder till att användaren inte kan utföra nödvändiga uppgifter i systemet. Bristen berör funktionen för att generera tidrapporter och inbegriper att det i prototypen endast är möjligt att generera en rapport för samtliga anställda eller för en anställd åt gången, därmed saknas möjligheten att generera en rapport för ett urval av anställda. Användare A3 uppmärksammade detta och anser att det är nödvändigt att åtgärda för att funktionen ska vara användbar och påpekar även att denna funktionalitet existerar i det nuvarande gränssnittet:

Ja precis. För det kan vi på den som vi har som rapportgenerator idag. (intervju, användare A3)

Denna brist påverkar i stor utsträckning systemets användbarhet och upptäcktes inte av den heuristiska utvärderingen. Det ger ett visst stöd till Liljegren (2006) som menar att användarinvolverad utvärdering beaktar de mest relevanta delarna av användbarhet.

En av bristerna i denna kategori upptäcktes av samtliga användare:

Liten knapp på projektöversikten för att hantera projekt som dessutom ligger för sent (för långt till höger) i tabellen. (användare A1, A2, A3)

Denna brist ledde till att samtliga användare försökte klicka på andra platser på projektöversikten för att kunna hantera ett projekt. Det är också den enda brist som upptäcktes av samtliga användare. Bristen upptäcktes även vid den heuristiska utvärderingen av testare H1.

Ett flertal andra brister berör mer specifika element så som enskilda knappar, texter och färgval samtidigt som ett par brister berör mer övergripande problem. Detta går i linje med det resultat Doubleday et al (1997) och Tan, Liu och Bishu (2009) kommit fram till gällande att användarinvolverad utvärdering i större utsträckning upptäcker detaljerade brister.

5.2.3. Visualisering av information

En brist inom denna kategori anses vara av mycket allvarlig art och syftar till att användaren upplevde det som omständligt att hitta specifik information gällande ett projekt då informationen var omfattande. Övriga brister var av en mer specifik art och berör exempelvis utseendet på arbetsaktiviteter som användare A2 önskar varit mer tydligt och uppdelat i separata faser:

[...] att kunna lägga in faser, se när nästa fas ska påbörjas att man har hela det här flödet i projektet på ett annat sätt. (intervju, användare A2)

Antalet brister inom denna kategori var lågt i förhållande till de övriga kategorierna och ingen av bristerna upptäcktes av fler än en användare.

6. Diskussion

De bägge utvärderingarna har, som går att utläsas av tabell 12, tillsammans genererat 44 stycken brister. Upptäckterna har till största del varit varierande både inom de respektive utvärderingarna och mellan de två utvärderingarna. Av de 44 bristerna är endast 10 stycken upptäckta av både den heuristiska utvärderingen och den användarinvolverade utvärderingen. Detta visar på att utvärderingarna har upptäckt olika typer av problem. Den tydligaste skillnaden är att de heuristiska testarna upptäckt mer strukturella och övergripande brister medan den användarinvolverade utvärderingen resulterade i mer detaljerade brister kring specifika element. Därav står denna studies resultat mer i linje med Doubleday et al (1997) och Tan, Liu och Bishu (2009) snarare än Liljegren (2006).

Tabell 12. Fördelning av brister

	Samtliga brister		Intuitivitet		Layout		Visualisering av info.	
Gemensamma	10	22.73%	2	10.52%	6	35.29%	2	25%
Heuristiska	14	31.82%	7	36.84%	4	23.52%	3	37.50%
Användare	20	45.45%	10	52.63%	7	41.18%	3	37.50%
Total:	44		19		17		8	
Andel av total:			43.18%		38.64%		18.18%	

Om man bortser från vilka typer av brister som upptäckts var den heuristiska utvärderingen både enklare och mindre tidskrävande att utföra än den användarinvolverade utvärderingen.

Förberedelserna inför den användarinvolverade utvärderingen krävde noggrann planering av vilka scenarion som deltagarna ska genomföra, hur observationen bör gå till, vår roll som observatörer i samband med observationen samt vilka frågor som bör ställas till användarna vid intervjun.

Detta var ej nödvändigt för att den heuristiska utvärderingen skulle kunna genomföras och

därmed framstår metoden som mer kostnads- och tidseffektiv gentemot användarinvolverad utvärdering vilket även Rogers, Sharp och Preece (2011) hävdar att den är.

Vid den användarinvolverade utvärderingen refererade och jämförde användarna prototypen med systemet Struktur som de arbetar i idag. De heuristiska testarna har inte arbetat i det nuvarande gränssnittet och har därmed inte haft möjligheten att hänvisa till detta. Detta kan till viss del ligga till grund för de skillnader i upptäckter som gjorts mellan utvärderingarna då användare stött på problem när de försökt arbeta i gränssnittet på samma vis som de gör i det nuvarande.

I studien av Tan, Liu och Bishu (2009) upptäckte de heuristiska testarna ett större antal brister än användarna gjorde, i denna studie har dock användarna upptäckt flest antal brister. I studien av Tan, Liu och Bishu (2009) är utvärderingarna genomförda på färdigutvecklade gränssnitt till skillnad från denna studie där en prototyp utvärderas. Prototypen i denna studie är därtill en ny design av ett gränssnitt som användarna i nuläget använder. Möjligtvis kan detta vara en anledning till att användarna upptäckte fler brister då de har kunskap och erfarenhet av praktiskt arbete med systemet. Detta kan exemplifieras med den brist som upptäcktes av användare A3 gällande tidrapportering och att det i prototypen inte är möjligt att se tidrapporter för mer än en användare i taget. Eftersom att användare A3 använder sådan funktionalitet i sitt dagliga arbete reagerade användaren på att detta saknas i prototypen. Denna brist upptäcktes inte av de heuristiska testarna och sannolikt är det en brist som är svår att upptäcka utan erfarenhet kring exakt vad för funktionalitet som krävs för att systemet ska vara användbart.

I denna studie leder både den användarinvolverade utvärderingen och den heuristiska utvärderingen till att det presenteras förslag på hur vissa brister ska åtgärdas. Enligt Doubleday et al (1997) ger en heuristisk utvärdering ett större antal förslag på lösningar, något som denna studie inte kan bekräfta. En möjlig anledning till att även användare bidrog med förslag på lösningar är att de har utfört de olika uppgifterna i prototypen tidigare, men i ett annat gränssnitt, då mycket av deras förslag berörde det nuvarande gränssnittet och hur de olika uppgifterna utförs där.

Tabell 13 redovisar bristernas fördelning utefter allvarlighetsgrad och vid vilken utvärdering de upptäcktes. Exempelvis upptäckte de heuristiska testarna två stycken brister av allvarlighetsgrad 1 inom kategori intuitivitet och användarna upptäckte tre stycken. Ingen av bristerna upptäcktes av bägge utvärderingarna. Gällande visualisering av information upptäcktes en mycket allvarlig brist av vardera utvärdering. Dessa två brister gäller samma problem och anses därför vara upptäckt av bägge utvärderingarna.

Tabell 13. Antal brister utefter allvarlighetsgrad och utvärdering där de upptäcktes

Allvarlighetsgrad	Typ	Intuitivitet	Layout	Vis. av information
1 (Mycket allvarlig)	Heuristiska	2	3	1
	Användare	3	5	1
	Varav gemensamma	0	3	1
2 (Allvarlig)	Heuristiska	3	6	2
	Användare	4	6	3
	Varav gemensamma	1	3	1
3 (Mindre allvarlig)	Heuristiska	4	1	2
	Användare	5	2	1
	Varav gemensamma	1	0	0

Samtliga brister av mycket allvarlig art som upptäcktes av de heuristiska testarna inom kategori layout upptäcktes även av användare. Användare upptäckte dock brister av mycket allvarlig art som de heuristiska testarna inte upptäckte. Detta förekom inte för någon annan kategori eller grad av allvarlighet vilket betyder att bägge utvärderingar har gjort upptäckter som den andra inte har gjort. Därmed har bägge utvärderingar bidragit med unika upptäckter som bör åtgärdas och att användarinvolverad utvärdering är mest effektiv, vilket Liljegren (2006) påstår, kan ej bekräftas av denna studie.

Bägge utvärderingsmetoderna har utvärderat prototypens användbarhet och lyckats uppmärksamma brister som bör åtgärdas för att uppnå en hög grad av användbarhet. Det existerar dock skillnader i vad för typ av brister som till största del uppmärksammas av utvärderingsmetoderna. Denna studies resultat tyder på att en heuristisk utvärdering av ett webbgränssnitt i prototypstadie upptäcker ett större antal generella och strukturella problem samtidigt som en användarinvolverad utvärdering upptäcker mer detaljerade brister kring olika element och funktioner. Detta tyder på att resultaten som Doubleday et al (1997) och Tan, Liu och Bishu (2009) kommit fram till även är applicerbara på webbgränssnitt i prototypstadie. Därmed bör, om möjligt, bägge utvärderingsmetoder användas för att upptäcka en så stor mängd brister som möjligt och på så vis kunna skapa webbgränssnitt med hög användbarhet.

Denna studie har genomförts med hjälp av tre heuristiska testare och tre användare vilket bör tas i beaktande vid tolkning av detta resultat. Hertzum och Jacobsen (2003) rekommenderar att fyra till fem användare deltar vid en användarinvolverad utvärdering och Nielsen (1995b) menar att fem heuristiska testare bör upptäcka 70 procent av bristerna i ett gränssnitt. Med det något lägre antal deltagare som denna studie baseras på ökar sannolikt risken för det Hertzum och Jacobsen (2003) menar är metodernas nackdel, att personliga preferenser och erfarenheter hos både studiens deltagare och de som genomför studien påverkar slutresultatet.

7. Slutsats

Studien har resulterat i ett antal insikter gällande heuristisk utvärdering och användarinvolverad utvärdering vilka ligger till grund för svaret på tidigare givna frågeställning; *“vilka typer av designbrister upptäcks vid heuristisk respektive användarinvolverad utvärdering av ett webbgränssnitt i prototypstadie?”*

De insikter som erhållits av studien är:

- Både heuristisk utvärdering och användarinvolverad utvärdering leder till unika upptäckter av brister vilka bör åtgärdas för att uppnå en hög grad av användbarhet i ett gränssnitt.
- Upptäckterna skiljer sig något mellan de olika utvärderingarna då heuristisk utvärdering ledde till mer övergripande och strukturella upptäckter. Användarinvolverad utvärdering uppmärksammar i större grad detaljerade brister gällande specifika objekt och element i ett gränssnitt.
- Användarinvolverad utvärdering var mer effektiv gällande att upptäcka brister rörande det praktiska utförandet av olika uppgifter då användarna har erfarenhet av att arbeta i det tidigare systemet samt erfarenhet av byggbranschen.
- Användarinvolverad utvärdering ledde till ett större antal upptäckta brister.
- Den heuristiska utvärderingen var mindre tidskrävande och enklare att utföra i förhållande till den användarinvolverade utvärderingen.
- Tidigare forskning gällande dessa utvärderingsmetoder är till stor del applicerbar även på webbgränssnitt i prototypstadiet.

Dessa insikter visar på att det förekommer skillnader mellan de olika utvärderingsmetoderna. Bägge metoder framstår dock som nödvändiga att utföra för att uppnå högsta möjliga användbarhet hos ett gränssnitt. Därmed rekommenderas gränssnittsutvecklare att i den mån det är möjligt genomföra både heuristisk utvärdering och användarinvolverad utvärdering på prototyper av ett gränssnitt för att kunna åtgärda en så stor mängd brister som möjligt. Sannolikt bör bägge dessa utvärderingar leda till unika upptäckter av brister och gränssnittsutvecklare får därmed ett större underlag att ta i beaktande vid fortsatt utvecklande av ett gränssnitt. Saknas möjligheten att genomföra bägge utvärderingsmetoder bör gränssnittsutvecklaren vara medveten om att heuristisk utvärdering kan leda till att brister rörande arbetssätt och praktiskt utförande av uppgifter kan gås om miste, samtidigt som en användarinvolverad utvärdering är mer tidskrävande och tenderar att i större utsträckning fokusera på detaljer i gränssnittet.

7.1. Studiens relevans och överförbarhet

Med uppsatsen har vi velat studera likheter och skillnader mellan metoder för att mäta användbarheten hos gränssnittsprototyper och har fokuserat på att redovisa vilka olika typer av brister som hittas under heuristisk utvärdering och användarinvolverad utvärdering. Då tidigare forskning i första hand studerat metoderna vid utvärdering av färdiga system som redan varit i bruk tror vi att resultatet av denna studie kan bidra till en ökad kunskap hos utvecklare som vill utvärdera gränssnitt i tidiga utvecklingsstadier. Vi tror även att de upptäckter som gjorts och som har redovisats i uppsatsen kan ge läsaren information om utvärderingsmetodernas styrkor, svagheter och karaktärsdrag både i det givna sammanhanget och inom andra situationer.

7.2. Förslag till fortsatt forskning

För studien har således ett förslag på ett nytt gränssnitt för det administrativa verktyget Struktur utvärderats med respektive metod. Tre personer har deltagit för att heuristiskt utvärdera

gränssnittet och tre personer har ställt upp som deltagare i den användarinvolverade utvärderingen. En utökad studie där metoderna testas på ett större antal gränssnitt med ett större antal deltagare hade kunnat producera mer data och troligtvis ett utökat resultat med fler insikter om respektive metod.

Då även funktionstunga system idag kan användas på handhållna enheter ställer det höga krav på väldesignade system med en hög grad av användbarhet på mindre skärmyta. Därför tror vi att det skulle vara intressant att genomföra en liknande studie där olika utvärderingsmetoder testas på tidiga prototyper som är specifikt ämnade att användas på handhållna enheter. Det hade varit intressant att se om resultatet av exempelvis heuristiska utvärderingar och observationer av användare vid en sådan studie skulle skilja sig från denna studie där ett traditionellt system som främst är ämnat att användas på datorer utvärderats.

Referenslista

- Alshamari, M. A. (2015). Improving the Effectiveness of Developed Heuristic Evaluation Method. *Journal of Applied Sciences*, 15(6), ss. 917.
- Batra, S. & Bishu, R. (2007). Web usability and evaluation: issues and concerns. I *Usability and Internationalization. HCI and Culture*. Peking, Kina 22-27 juli 2007, ss.243-249. DOI: 10.1007/978-3-540-73287-7_30
- Bevan, N. (1995). Measuring usability as quality of use. *Software Quality Journal*, 4(2), ss. 115-130.
- Delice, E. K. & Güngör, Z. (2009). The usability analysis with heuristic evaluation and analytic hierarchy process. *International Journal of Industrial Ergonomics*, 39(6), ss. 934-939.
- Doubleday, A., Ryan, M., Springett, M. & Sutcliffe, A. (1997). A comparison of usability techniques for evaluating design. I *Proceedings of the 2nd conference on Designing interactive systems: processes, practices, methods, and techniques*. Nederländerna 18-20 augusti 1997, ss. 101-110. DOI: 10.1145/263552.263583
- Fernandez, A., Insfran, E. & Abrahão, S. (2011). Usability evaluation methods for the web: A systematic mapping study. *Information and Software Technology*, 53(8), ss. 789-817.
- Hertzum, M. & Jacobsen, N. E. (2003). The evaluator effect: A chilling fact about usability evaluation methods. *International journal of human-computer interaction*, 15(1), ss. 183-204.
- Kurečić, M. S. (2005). Improvement of Web Design using the Heuristic Evaluation Method. *Acta Graphica*, 17(1-4), ss. 7-14.
- Liljegren, E. (2006). Usability in a medical technology context assessment of methods for usability evaluation of medical equipment. *International Journal of Industrial Ergonomics*, 36(4), ss. 345-352.
- Maan, J. & Mantha, N. (2014). Rich Internet Applications, Platforms and Tools - A Paradigm Shift in Web User Experience. *Computer Science & Information Technology*, 4(3), ss. 121-129.
- Nazrul Islam, M. & Tétard, F. (2014). Exploring the impact of interface signs' interpretation accuracy, design, and evaluation on web usability: a semiotics perspective. *Journal of Systems and Information Technology*, 16(4), ss. 250-276.
- Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. I *Proceedings of the SIGCHI conference on Human factors in computing systems*. Seattle, USA 1-5 april 1990, ss. 249-256. DOI: 10.1145/97243.97281

Nielsen, J. (1992). Finding usability problems through heuristic evaluation. I *Proceedings of the SIGCHI conference on Human factors in computing systems*. Monterey, USA 3-7 maj, ss. 373-380. DOI: 10.1145/142750.142834

Nielsen, J. (1994). *Usability engineering*. Elsevier.

Nielsen, J. (1995a). *10 Usability Heuristics for User Interface Design*.
<https://www.nngroup.com/articles/ten-usability-heuristics/> [2016-04-05]

Nielsen, J. (1995b). *How to Conduct a heuristic evaluation*.
<https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/> [2016-04-14]

Patel, R. & Davidson, B. (2011). *Forskningsmetodikens grunder: Att planera, genomföra och rapportera en undersökning*. 4. uppl., Lund: Studentlitteratur.

Rogers, Y., Sharp, H. & Preece, J. (2011). *Interaction Design: Beyond Human-Computer Interaction*. 3. uppl., Chichester: Wiley.

Shackel, B. (2009). Usability–Context, framework, definition, design and evaluation. *Interacting with Computers*, 21(5-6), ss. 339-346.

Tan, W., Liu, D. & Bishu, R. (2009). Web evaluation: Heuristic evaluation vs. user testing. *International Journal of Industrial Ergonomics*, 39(4), ss. 621-627.

The International Organization for Standardization (ISO) (1998). *9241-11 Ergonomic requirements for office work with visual display terminals (VDTs)*. ISO.

Usability Partners. (u.å.). *ISO-standarder*. <http://www.usabilitypartners.se/om-anvandbarhet/iso-standarder.php> [2016-04-10]

Bilagor

Bilaga 1 - Instruktioner för heuristisk utvärdering

Instruktioner - Heuristisk utvärdering

Denna utvärdering är en del av ett examensarbete som utförs av Patrik Nilsson och Pontus Tengnäs på Göteborgs universitet. Syftet med arbetet är att kartlägga och jämföra två olika metoder för utvärdering av gränssnitt, heuristisk samt användarinvolverad utvärdering. Därmed kommer resultatet av denna utvärdering att analyseras med detta i åtanke snarare än att användas för att förbättra den design som utvärderas.

Du som utvärderare kommer att vara fullständigt anonym och endast benämnas vid en sifferkod i uppsatsen. En kortare presentation av din erfarenhet och bakgrund kan presenteras vid ditt godkännande. Du har i så fall rätt att läsa denna presentation för att godkänna den innan arbetet lämnas in. Vårt examensarbete kommer vid ett godkänt betyg att publiceras på Göteborgs universitets digitala arkiv för uppsatser. Detta arkiv är tillgängligt för allmänheten.

Prototypen som ska utvärderas är ett designförslag på ett projekthanteringssystem som har tagits fram på uppdrag av Vendium AB. Vendium AB:s kunder är byggföretag av mindre storlek och det är också dessa som systemet riktar sig mot. Dessa användare kan ha varierande grav av datorvana. Systemet är webbaserat och körs därmed i valfri webbläsare. Prototypen har ingen funktionalitet implementerad men det är möjligt att navigera runt bland de sidor som tagits fram för detta tillfälle.

Utvärderingens upplägg

I denna zip-fil finns följande innehåll:

- *Heuristiker.pdf*; Dokument innehållande de heuristiker som ska granskas under utvärderingen.
- */Prototyp/*; Denna mapp innehåller den prototyp som utvärderingen ska genomföras på.

Premissen för heuristisk utvärdering är att testaren utifrån ett antal på förhand bestämda riktlinjer/heuristiker bedömer ett användargränssnitt genom att fritt navigera och interagera i detta. Vi föreslår att du börjar med att öppna dokumentet *heuristiker.pdf* och studerar de riktlinjer och regler som finns listade där. Det är sedan dessa aspekter du kommer bedöma gränssnittet utifrån, om det går att finna några överträdelser/brister inom designen för dessa heuristiker. Upptäcker du brister som inte täcks in i de bifogade heuristikerna ber vi dig att ange dessa som brister av kategori *Övrig designbrist*. I fall din feedback inte går att matcha till vissa heuristiker över huvud taget spelar detta ingen roll. Vi uppskattar all feedback och det är inget problem i fall mycket av dina upptäckter hamnar under kategorin *Övrig designbrist* eller i fall vissa heuristiker inte omnämns över huvud taget.

Nästa steg är förslagsvis att köra prototypen som du gör genom att öppna filen *startside.html* med valfri webbläsare. Övriga sidor i mappen kan nås genom att navigera i prototypen. Hur Du genomför utvärderingen är upp till dig själv. Du väljer om du vill genomföra utvärderingen i en session eller dela upp det i fler och du avgör själv hur du vill granska prototypen. Viktigt är dock att utvärderingen utgår från de heuristiker som finns bifogade, vi hoppas att dessa kan ge någon form av inriktning till din granskning.

Du fyller i de överträdelser av heuristikerna du upptäcker i ett formulär som finns på följande länk: <http://goo.gl/forms/8vhQMvQHl3>. Du anger då vilken heuristik som överträds samt ger en kommentar angående vad det är som orsakat problemet. Du anger också vilken grad av allvarighet som överträdelser innebär på en skala mellan 1-3 (1 = Mycket allvarligt, 2 = Allvarligt, 3 = Mindre allvarligt). Överträdelser av grad 1 och 2 är sådant som måste åtgärdas snarast medan grad 3 innebär mildare överträdelser som inte har så stor påverkan på användarupplevelsen.

Om du har några frågor rörande hur du ska genomföra utvärderingen, eller undrar någonting annat så är vi tillgängliga.

Patrik Nilsson

Tel: 0737-594917

Mail: patriknil90@gmail.com

Pontus Nilsson Tengnäs

Tel: 0762-084011

Mail: pontusnt@gmail.com

Bilaga 2 - Medgivande om ljud och skärminspelning vid observation samt intervju

Medgivande om inspelning

Under denna utvärderingssession kommer ljud samt skärmbilder att spelas in, detta inkluderar att den intervju som genomförs också spelas in med ljud. Detta för att materialet ska kunna analyseras av oss vid senare tillfälle och användas i vår uppsats. I uppsatsen kommer du som testdeltagare att behandlas anonymt och benämnas vid den id-siffra du tilldelas vid detta tillfälle. En kortare presentation om din bakgrund och erfarenhet av liknande system publiceras i uppsatsen vid ditt godkännande. Du har möjlighet att få läsa och godkänna denna presentation innan uppsatsen lämnas in.

Vårt examensarbete kommer vid ett godkänt betyg att publiceras på Göteborgs universitets digitala arkiv för uppsatser. Detta arkiv är tillgängligt för allmänheten.

Vänligen skriv under nedan vid samtycke.

Jag godkänner att detta utvärderingstillfälle spelas in med ljud samt skärmbilder och att detta material används av Patrik Nilsson och Pontus Nilsson Tengnäs i deras examensarbete vid Göteborgs universitet vårterminen 2016.

Signatur: _____

Namnförtydligande: _____

Datum: _____

Bilaga 3 - Scenarion vid observation

1. Startside: Gå till projekt 1004.
2. Avsluta aktivitet Rivning och påbörja aktivitet Målning.
3. Pontus Tengnäs har köpt 3st vinkelslipar för 800kr st. Registrera detta i projektet
4. Alan Carlson har också köpt material. 2st cementblandare för 3000kr st, till projekt 1002. Registrera detta och skriv sedan ut projektet.
5. Visa en översikt för alla inrapporterade tider för alla projekt
6. Visa Torbjörn Nilssons inrapporterade tider med arbetsmoment för mars
7. Rapportera en tid för Torbjörn Nilsson på projekt 1003 för ett arbete som ligger utöver offerten. Antal timmar och övriga värden kan du fritt fylla i med vad som helst.
8. Gå tillbaka till startsidan.
9. Ta dig till projektöversikten för att kolla hur många projekt som gått över deadline.
10. Utan att lämna projektöversikten, kolla hur många dagar det är kvar till deadline på projekt 1010.
11. Lägg till sista en sista tidrapport i projekt 1002. Användaren Håkan Mild har arbetat 16 timmar med att Slipa under 2 dagar med start den 4e april. Han har även kört 2 mil med sin privata bil för detta arbete.
12. Projektet är nu färdigt, avsluta projekt 1002.

Bilaga 4 - Intervjumall

1. Vad är din övergripande uppfattning om interaktionen i systemet?
 - a. Fråga varför deltagaren tycker som den tycker
2. Vad upplevde du var bra med gränssnittet?
3. Vad upplevde du var mindre bra?
 - a. Be deltagaren peka ut specifika delar och förklara vad som hade kunnat göras bättre
4. Vilka scenarion/instruktioner ansåg du var svåra att genomföra? Varför?
5. Vad anser du om hur informationen presenteras för dig som användare? Var den översiktlig och lätt att ta till sig?
 - a. Tabeller
 - b. Startside
 - c. Projektöversikt
 - d. Projekt
 - e. Tidrapport
6. Är gränssnittet behagligt?
 - a. Estetiskt?
 - b. Att använda?
7. Upplever du gränssnittet som sammanhängande och logiskt?
 - a. Exempelvis vid växlande av projekt, känns det naturligt och praktiskt att göra på det viset?

Bilaga 5 - Skärmdumpar från prototyp

Välkommen (namn)!

Projekt - April

10	1
Aktiva projekt	Projekt nära deadline
2	3
Avslutade projekt	Nya projekt

Tidrapporter - April

23	2
Rapporterade timmar	Projekt utan tidrapportering
2016-04-11	4
Senaste tidrapport	Övertidstimmar

Intern forum

Här visas inlägg skrivna av användare.

Skapa nytt inlägg

Skriv din text här:

Kommande händelser

Typ	Datum	Projekt
Projektstart	2016-04-25	325 - Mats Green
Inventering	2016-05-03	Almänt
Möte	2016-05-10	500 - Dennis
Materieberedning	2016-05-12	500 - Dennis
Projektstart	2016-05-18	410 - Vadsörens Plan

Mail

Titel	Status	Öppnat
Chefen	Viktigt	Ja
Bosse	Normal	Nej
Skattevecket	Normal	Ja

Startsida

Projektöversikt

Sök pågående projekt...

Pågående | Avslutade | Nära deadline

ID	Beställare (Förstagagnsopp)	Ansvarig (Kontakt)	Målnöjning	Status	Notering	Senaste aktivitet	Har ett projekt
1001	GÅS	Roland Nilsson	Målning	■	0	3 dagar sedan	🗄
1002	IFK Göteborg	Mats Gren	Kökrenovering	■	0	2 dagar sedan	🗄
1003	Örgryte IS	Alfonso Alves	Kökrenovering	■	0	1 dag sedan	🗄
1004	BK Häcken	Petronella Getherdsson	Takläggning	■	0	2 dagar sedan	🗄
1005	Västra Frölunda IF	Niklas Skoog	Ny ugn	■	0	2 dagar sedan	🗄
1006	Övding BK	Tobias Sana	Fasadrenovering	■	0	1 dag sedan	🗄
1007	Öddevold BK	Percy Barnesik	Byta av fönster	■	0	2 dagar sedan	🗄
1008	Västra Frölunda IF	Hasse Blomqvist	Kalla badrum	■	0	Idag	🗄
1009	IFK Göteborg	Hjalmar Jonsson	Snöcken träpaneler	■	0	4 dagar sedan	🗄
1010	Kortedda IF	Daniel Alfredsson	Målning	■	0	2 dagar sedan	🗄
1011	IFK Göteborg	Mats Gren	Golvläggning etage	■	0	4 dagar sedan	🗄
1012	BK Häcken	Muhammed Abubakari	Limning av paneler	■	0	3 dagar sedan	🗄
1013	IFK Göteborg	Petta Frej	Tömning av papperskorg	■	0	7 dagar sedan	🗄
1014	GÅS	Dime Jankulovski	Riva gäsgården	■	0	4 dagar sedan	🗄

Projektöversikt

struktur

Home Personallista Tidrapport Offert Dato Projekt Fakta Läs Översikt

Översikt Skapa projekt Kalender Sammanställning Aktiv

Projekt 1002 [Välj annat projekt](#)

Beställare

Förnamn: IFK Göteborg
 Adress: Halmstadsvägen 22, 421 21 V. Halmstad
 E-mail: mats.gren@ifkgbrotorp.se
 Mobiltelefon: 031245455
 Telefon: 031545451

Information

Datum: 2016-03-23 — 2016-03-31
 Huvudman: Mats Gren
 Öffentligt an. nummer: 28 300 (en-vel)
 Kunden ordernr.: 111322

[Fyll ut planerade](#) [Fyll ut avslutade](#)

[Generera projekt ID](#) [Faktureras](#)

[Skriv ut](#) [Ta bort projekt](#)

Tidrapportera

Välj användare: Glenn Hysén

Dagar att rapportera: 1

Datum: Åååå-mm-dd

Timmar: Arbetsade timmar

Mål: Mål tjänstefordon

Privata mil: Antal privata mil

Utfört arbete: Beskrivning

Ligger arbetet som utförts utöver aktuell offert?

[Rapportera](#)

Utfört arbete

Datum	Utfört arbete	Fakturerade timmar	Fakturerade mil
2016-04-02	Finnsaker kök	16	0
2016-04-01	Muret	8	0
2016-03-29	Kallat badrumsgg/vederläggning	8	4
2016-04-02	Finnsaker kök	16	0
2016-04-01	Muret	8	0
2016-03-29	Kallat badrumsgg/vederläggning	8	4
2016-04-02	Finnsaker kök	16	0
2016-04-01	Muret	8	0
2016-03-29	Kallat badrumsgg/vederläggning	8	4

Arbet dagar: 3
 Fakturerade timmar: 80 Fakturerade mil: 2
 E) fakturerade timmar: 14 E) fakturerade mil: 2

Material / Övriga kostnader

Beställning	Antal	Faktureringspris	Summa
Material			
KLOOR 4150 FINNÅSÄMMING 2	10	132,18	1 321,80
KABEL ARCTIC VIT V400 A0354	200	5,5	1 100,00
Takläggning/gulv 0x4 20x	1	2638,4	2 638,40
Totalsk			4 891,2
Makskostnad			
Bogsering	2	500	1 100,00

Pågående Aktiviteter

Rivning	Slipning	Tapetsering
Planerad start: 2016-03-29 Färdigt start: 2016-03-31 Planerad slut: 2016-04-22 Avsluta aktivitet	Start: 2016-04-05 Planerad slut: 2016-04-20 Avsluta aktivitet	Planerad start: 2016-04-05 Färdigt start: 2016-04-08 Planerad slut: 2016-04-22 Avsluta aktivitet

Planerade Aktiviteter

[Målning](#) [Öppning](#) [Städning](#) [Transport](#)

Projektsida

struktur

Home Personallista Tidrapport Offert Dato Projekt Fakta Läs Översikt

Tidrapport Översikt Projekt Rapportgenerator

Tidrapport

Tidrapportera

Välj pågående projekt: 1001 I GAIS

Välj användare: Glenn Hysén

Dagar att rapportera: 1

Datum: Åååå-mm-dd

Timmar: Arbetsade timmar

Mål: Mål tjänstefordon

Privata mil: Antal privata mil

Utfört arbete: Beskrivning

Ligger arbetet som utförts utöver aktuell offert?

[Rapportera](#)

Rapporterade tider - Översikt

Användare: Alla användare Projekt: Alla projekt Fil: Åååå-mm-dd Till: Åååå-mm-dd [Visa](#)

Visa med kommentar Visa med arbetsmoment [Spara som PDF \(0\)](#)

Datum	Projekt	Timmar	Mil	Privata mil	Utfört arbete
2016-04-02	1001	2	2	0	Ng
2016-04-02	1002	4	1	1	Ng
2016-04-23	1001	7	1	0	Ja
2016-04-20	1004	8	2	0	Ng
2016-04-20	1003	4	0	0	Ng
2016-04-20	1002	6	1	0	Ja
2016-04-19	1001	7	0	2	Ng
2016-04-19	1004	5	2	0	Ng
2016-04-18	1003	9	0	0	Ng
2016-04-18	1002	8	1	0	Ng
2016-04-18	1001	7	1	0	Ng
2016-04-17	1003	4	1	1	Ng
Totalt för vald period		68	12	4	

[Visa detaljerad information för vald tidrapport](#)

Senaste 40 tidrapporterna

Användare	Projekt	Datum	Utfört arbete	Timmar	Mil
Bror Johnson	1001	2016-04-28	Spelut	6	1
Karin Gustafsson	1003	2016-04-28	Rev en vägg	6	0
Håkan Hellström	1003	2016-04-27	Släppte en singel	7	2
James Herfeld	1001	2016-04-27	Kallade	8	0
Kalle Mörus	1002	2016-04-27	Spelat	3	1
Lef Olv	1004	2016-04-26	Läste ett mysterium	5	0

Frånvaro

Användare	Datum	Timmar	Arbetslag
Rick James	2016-01-03	8	Sjukvård
Rick James	2016-01-02	8	Sjukvård
Lars Ullrich	2016-01-01	1	Försvning

Tidrapporteringssida