

MAGISTERUPPSATS I BIBLIOTEKS- OCH INFORMATIONSVETENSKAP
VID BIBLIOTEKS- OCH INFORMATIONSVETENSKAP/BIBLIOTEKSHÖGSKOLAN
2003:58

GRÄNSSNITT, ANVÄNDBARHET OCH SÖKMOTORER ETT EXPERIMENT

Sara Larsson

© **Författaren/Författarna**

Mångfaldigande och spridande av innehållet i denna uppsats
helt eller delvis – är förbjudet utan medgivande av författaren.

Svensk titel: Gränssnitt, användbarhet, och sökmotorer: ett experiment

Engelsk titel: The User Interface, Usability and Search Engines: an Experiment

Författare: Sara Larsson

Kollegium: Kollegium 2

Färdigställt: 2003

Handledare: Johan Eklund

Abstract: This aim of this thesis is to measure the effect of the user interface in real searches on the Internet through a small experiment.

Theories on web and search usability and psychological factors concerning usability were first explored. Then three interfaces were constructed, one "good" that followed usability guidelines and had multiple search functions, one "bad" that completely disobeyed those guidelines and had only a simple and small search field, and one "intermediate" interface that was a mix of the two. The search field was set to different sizes, with the shortest in the "bad" interface, a medium field in the "intermediate" interface and several long fields in the "good" interface. All three interfaces were then attached to the same web search engine, Altavista. 15 students divided into 3 groups were given one interface each. They performed a list of searches and gave a judgement on the relevance of their result. Thus, precision was recorded. The time to formulate a query and the length of the queries were also recorded. At the end of the experiment a short questionnaire followed to measure the students' satisfaction on a Likert scale.

The results were analysed with a Mann-Whitney test and the conclusion was that the precision measure showed no significant differences between the interfaces. However, the time was significantly longer in the "good" interface compared to the other two and the length of queries was significantly longer in the "good" interface. Satisfaction was also significantly higher in the "good" user interface.

Nyckelord: Gränssnitt, sökmotor, användbarhet, utvärdering, MDI, Människa-Dator Interaktion, Internet

INNEHÅLLSFÖRTECKNING

1. INLEDNING.....	5
1.2 BAKGRUND	5
1.3 BEGREPPSDEFINITIONER OCH FÖRKORTNINGAR	6
2. PROBLEMFÖRMULERING.....	8
2.1 SYFTE	8
2.2 AVGRÄNSNINGAR	8
2.3 FRÅGESTÄLLNINGAR.....	8
3. LITTERATUR OCH TIDIGARE FORSKNING.....	9
3.1 KOPPLING TILL BIBLIOTEKS- OCH INFORMATIONSVETENSKAP	11
4. TEORI.....	12
4.1 ANVÄNDBARHET	13
4.2 RECALL OCH PRECISION.....	14
4.3 KORT HISTORIK	15
4.4 SÖKFUNKTIONER.....	16
4.4.1 En eller flera sökningar.....	18
4.4.2 Meny, knappar och ikoner.....	18
4.5 PSYKOLOGISKA FAKTORER.....	20
4.5.1 Olika typer av minne	20
4.5.2 Förförståelse.....	21
4.6 HJÄLPFUNKTIONER.....	22
4.7 GRAFIK OCH TYDLIGHET	23
4.7.1 Navigation	24
4.7.2 Estetik	25
5. METOD.....	27
5.1 KVANTITATIV OCH KVALITATIV METOD.....	27
5.2 METOD I DENNA UPPSATS.....	28
5.2.1 Design av gränssnitten.....	29
5.2.2 Gränssnittens utformning	30
5.2.3 Sökuppsdragen och enkätfrågorna	34
5.2.4 Testpersonerna.....	35
5.2.5 Antalet användare.....	36
5.2.6 Testmiljö.....	36
5.2.7 Statistiska förklaringar	38
6. REDOVISNING AV RESULTAT	41
6.1 PRECISION	42
6.2 TID.....	42
6.3 SÖKSTRÄNGENS LÄNGD	43
6.4 TILLFREDSSTÄLLELSE	44
7. ANALYS OCH DISKUSSION.....	46
7.1 PRECISION	46
7.2 TID.....	48
7.3 SÖKSTRÄNGENS LÄNGD	50
7.4 TILLFREDSSTÄLLELSE	51
7.5 SLUTSATSER.....	53
7.6 FÖRSLAG PÅ VIDARE FORSKNING.....	54
8. SAMMANFATTNING.....	56
KÄLLFÖRTECKNING.....	58
TRYCKTA KÄLLOR.....	58
ELEKTRONISKA KÄLLOR	59

BILAGOR	61
BILAGA 1. STENCILN SOM DELADES UT VID EXPERIMENTET	61
BILAGA 2. SÖKUPPDRAGEN	63
BILAGA 4. GRÄNSSNITTENS UTSEENDE	66
<i>Gränssnitt 1</i>	66
<i>Gränssnitt 2</i>	67
<i>Gränssnitt 3</i>	67
BILAGA 5. DE ENSKILDA TESTPERSONERNAS SÖKRESULTAT	69
<i>Precision</i>	69
<i>Tid</i>	70
<i>Söksträngens längd</i>	72
<i>Tillfredsställelse</i>	74

1. INLEDNING

Jag har valt att arbeta med gränssnitt i datorer bland annat på grund av mitt intresse för webbdesign och konsekvenserna av den. Jag har under min studietid läst bland annat webbutveckling och elektronisk publicering och tycker om att experimentera med HTML.

Under min tid på BHS har jag flera gånger reflekterat över tillkrånglade gränssnitt, som försvårat sökningar i olika databaser. Också på webben är det ibland oklart vilka operatörer som får användas och hur. Samtidigt har jag tänkt att det kanske inte är gränssnitten som är det verkliga problemet och att det vore intressant att veta hur stor del av sökresultatet och sökbeteendet som är relaterat till gränssnittet.

När Johan Eklund kom med förslaget att göra en empirisk undersökning av detta tyckte jag att det lät som en spännande idé som skulle ge konkreta forskningsresultat.

När man börjar intressera sig för gränssnitt och människa-dator interaktion märker man snabbt att *usability*, användbarhet, är ett utbrett begrepp. Det är denna användbarhet som diskuteras i litteraturen, och det är också det jag strävat efter att behandla i min uppsats.

1.2 BAKGRUND

Vad är egentligen ett gränssnitt? Ett hjälpmedel eller en barriär? Varför ska man forska i det?

Jonas Löwgren som forskar i området människa-dator interaktion beskriver hur ett gränssnitt kan vara något så vardagligt som en duschkrans utseende. En äldre dusch har ofta en varm *och* en kallvattenkran, något som fungerar utmärkt, men som kan vara obehagligt när man står i duschen och behöver justera temperaturen snabbt. I nyare duschar kan man justera temperaturen med hjälp av endast *ett* vred. Löwgren visar på att man vid normala förhållanden inte reagerar på att ett gränssnitt är klumpigt, men att det visar sig vid till exempel stress (när man skållar sig i duschen) och då kan bli ett hinder mellan människa och system.¹ Så fungerar det också vid datorn. Man tänker inte på gränssnittet så länge allt fungerar, men när man behöver lära sig något nytt, eller om datorn krånglar, kan det vara svårt att hitta rätt knapp eller hjälpfunktioner.

Denna uppsats har, genom ett experiment där användare utfört sökuppdrag i tre olika egenproducerade gränssnitt av varierande kvalitet som var kopplade till Altavistas² sökmotor, strävat efter att ge svar på frågor om sambandet mellan gränssnitt och sökprestation och sökbeteende.

Marchionini skriver att söksträngsformulering (eng. query formulation)³ är två olika typer av kartläggning: en semantisk kartläggning av informationssökarens vokabulär använt för att artikulera uppdraget till databasens vokabulär, och en funktions-

¹ Löwgren, Jonas, 1993. Human-Computer Interaction: What every system developer should know, s, 13ff.

² Altavista finns med i källförteckningen och kommer fortsättningsvis inte redovisas med noter.

³ Min strävan är att konsekvent erbjuda svenska alternativ till engelska termer, men då översättningarna många gånger tvingats avvika från originalets betydelse har de engelska termerna inkluderats i texten.

kartläggning av de strategier som informationssökaren bedömer är bäst för att konvertera uppdraget till de regler och finesser som gränssnittet medger.⁴ Att gränssnittet påverkar både söksträngsformulering och sökresultat är inte ett orimligt antagande.

Ett gränssnitt ska, enligt mig, hjälpa användaren att utnyttja systemet bakom. Ett idealiskt gränssnitt ska vara så självklart och lätt att använda att det nästan inte märks.

1.3 BEGREPPSDEFINITIONER OCH FÖRKORTNINGAR

Detta avsnitt är viktigt för att fortsättningen ska vara tydlig.

Jag har hittills använt ordet gränssnitt utan vidare problematisering, men vad är ett gränssnitt egentligen? I Nationalencyklopedin på Internet finns följande definition:

gränssnitt, eng. interface, i tekniska sammanhang en väldefinierad och formellt beskriven kontaktyta mellan två system eller enheter. Gränssnittet beskriver hur enheterna kan kommunicera, t.ex. hur signaler kan skickas på en uppsättning ledningar eller hur två programmoduler kan samarbeta. Ytterligare ett exempel är användargränssnitt, (graphical) user interface, som beskriver hur bilden på datorskärmen används för kommunikation med ett program i datorn.⁵

Vid en likadan sökning hos svenska datatermgruppen, som ger rekommendationer för svenskt språkbruk av datatermer, och har en webbplats som ligger under Kungliga tekniska högskolans, får jag följande resultat:

gränssnitt

kontaktyta mellan olika funktioner eller delar i ett system. Använd gränssnitt för eng. interface.

Kommentar:

Exempel på gränssnitt är gränssnittet mellan två program, mellan dator och modem samt mellan dator och skrivare. Användargränssnitt, eng. user interface, kallas det gränssnitt som möjliggör kommunikation mellan människa och dator och utgörs bl.a. av det man ser på bildskärmen. Det finns t.ex. grafiska användargränssnitt (eng. graphic(al) user interface, GUI) som bygger på fönster, ikoner (symboler) etc., och rent textbaserade användargränssnitt. Om man avser en hårdvaruenhet kan gränssnittskort för eng. interface board (interface card) vara en lämplig svensk term.⁶

Gary Marchionini, som fokuserar på gränssnittet för sökningar, beskriver gränssnittet som en kommunikationskanal mellan människa och maskin med fysiska och konceptuella komponenter. Gränssnittet avgör hur användbart, tillfredsställande och lätt att lära ett söksystem är. Gränssnittet kan därför påverka en söknings prestation.⁷

Jag har tidigare i denna uppsats visat hur ordet gränssnitt kan tolkas i vid mening, och denna vidare tolkning är bra att hålla i minnet när vi nu går vidare, eftersom det som gör

⁴ Marchionini, Gary, 1997. *Information Seeking in Electronic Environments*. S. 53.

⁵ Nationalencyklopedin på Internet, NE.se, sökning på "gränssnitt".

⁶ Sv. datatermgruppen: Ordlista (v20), sökning på "gränssnitt".

⁷ Marchionini, Gary, 1997. *Information Seeking in Electronic Environments*. S. 41.

något så enkelt som ett dörrhandtag bra eller dåligt, många gånger är samma saker som gör avancerade gränssnitt på webben bra eller dåliga. Det ska vara tydligt hur gränssnittet ska användas och det ska vara lätt att lära sig. I denna uppsats kommer jag dock framöver att mena grafiska användargränssnitt när jag talar om gränssnitt. Jag kommer nedan att förklara de vanligaste specifika förkortningarna och orden inom detta ämne och denna uppsats.

MDI: Människa-Dator Interaktion, motsvaras av engelskans:

HCI: Human-Computer Interaction.

webb: den svenska stavningen av funktionen som gör att man enkelt kan hämta sammanlänkad information i form av text, bild och ljud från Internet eller ett intranät.

webbsida: en sammanhållen sida som man kan nå helt utan att klicka på länkar.

webbplats: en samling sammanlänkade webbsidor som behandlar ett visst ämne, och/eller utgår från en viss organisation eller person.

förstasida: webbsida som är huvudingång till en webbplats.⁸

Fler definitioner inom området får Löwgren stå för:

System: detta är ett ord jag kommer att använda i betydelsen ”det bakomliggande systemet”, det som en människa interagerar med genom att använda gränssnittet. I mitt fall blir detta själva databasen av webbdokument.

Systemservice: Systemservicen är *vad* som kan göras av systemet, medan gränssnittet bestämmer *hur* det kan göras. Om man till exempel har en databas där författarnas födelseår finns inlagda, bör gränssnittet också göra detta sökbart för att användarna ska kunna nyttja systemservicen optimalt.⁹

Usability: Löwgren är en av många som argumenterar för att ordet användarvänlighet inte är korrekt eftersom en användare vill ha gränssnitt, datorer osv som effektivt utför de uppgifter som efterfrågas, vänlighet har inte med saken att göra. ”Användarvänligheten” ska ju också vara mätbar, men själva ordet gör detta svårt och det låter oseriöst. Därför har forskare i ämnet övergått till termen usability, användbarhet, och denna term kommer jag att använda i uppsatsen.¹⁰

Användbarhet: se ovan.

Sökmotor: En sökfunktion på webben, för att finna webbsidor, pdf-filer eller motsvarande. Till exempel Altavista och Google¹¹.

Dokument: Ordet dokument används i denna uppsats om de flesta slutprodukter för sökningar, det vill säga webbsidor, tryckta böcker, och fulltextartiklar i databaser.

⁸ Definitioner av webb, webbsida, webbplats och förstasida från Sv. datatermgruppen: Ordlista (v20).

⁹ Löwgren, Jonas, 1993. *Human-Computer Interaction: What every system developer should know*, s. 13ff. ”Översättningarna” är mina egna.

¹⁰ Ibid., s. 51.

¹¹ <http://www.altavista.com> och <http://www.google.com>. jag kommer i fortsättningen inte att skriva ut noter till dessa sökmotorer.

2. PROBLEMFÖRMULERING

Denna uppsats strävar efter att genom ett experiment få klarhet i om gränssnitt påverkar effektiviteten i sökningar, och om det i så fall påverkar i enlighet med nu rådande teorier. I detta avsnitt kommer en mer exakt förklaring på vad jag vill uppnå med uppsatsen att presenteras.

2.1 SYFTE

Syftet var att ta reda på om rådande teorier om användbarhet applicerade i tre olika gränssnitt påverkar verkliga sökningar på ett förväntat sätt. Precision, tidsåtgång för sökning, söksträngens längd och subjektiv tillfredsställelse var de fyra områden där jag förväntade mig att se skillnader i resultat.

2.2 AVGRÄNSNINGAR

Det finns en del litteratur om hur kulturella skillnader och handikapp bör påverka gränssnittets och utformning, men det behandlas inte här. Längre fram i uppsatsen kommer dessa ämnen att vidröras flyktigt, men då de inte inverkar på undersökningen har jag valt att inte fördjupa mig i detta.

I litteraturen betonas också ofta vikten av träfflistans utseende¹² vid sökningar och felmeddelandens¹³ utformning vid uteblivna relevanta träffar. Av tekniska skäl är det svårare och mer tidsödande att manipulera resultatsidor än gränssnittet för sökning. Jag har därför valt att låta resultatsidorna för de olika gränssnitten i experimentet genereras av sökmotorn och de har därför fått Altavistas standardutseende. Jag har inte heller redovisat för teorier kring träfflistor i någon större utsträckning.

2.3 FRÅGESTÄLLNINGAR

- Påverkas sökningars precision av gränssnittets utformning?
- Är tiden som går mellan läsning av sökuppdrag och formulering av färdig söksträng beroende av gränssnittets utformning?
- Påverkas söksträngens längd av storleken på sökfältet?
- Påverkas användarnas subjektiva tillfredsställelse och upplevelse av sökning av gränssnittets utformning?
- Kan de variationer som uppstår, då det gäller de fyra faktorerna: precision, tid, söksträngslängd och tillfredsställelse, förklaras med slumpen eller är skillnaderna signifikanta?
- Påverkas resultatet enligt förväntningarna och relevant litteratur?

¹² Se t ex Hearst, Marti A, 1999. User Interfaces and Visualization, ingår i *Modern Information Retrieval*. s. 289-306 och Head, Alison J, 2000. *Design wise: A guide for evaluating the interface design of information resources*. s. 62.

¹³ Se t ex Allwood, Carl Martin, 1991. *Människa-datorinteraktion: ett psykologiskt perspektiv*. s. 80ff.

3. LITTERATUR OCH TIDIGARE FORSKNING

För att finna lämplig litteratur har jag använt mig av den så kallade ”snöbollsmetoden”. Jag började med att leta upp lämpliga magisteruppsatser som berörde gränssnitt och användbarhet, för att få en överblick över området och finna lämpliga termer. Jag tittade i källförteckningar och sökte vidare på namn som återkom ofta. De uppsatser som låg närmast mitt eget uppsatsämne gav många idéer till min disposition och vilka aspekter av gränssnitt jag skulle ta upp, men då de använde sig av kvalitativa metoder, genom att intervjua webbansvariga¹⁴, dela ut enkäter till användare¹⁵ och rena litteraturstudier¹⁶, fick jag söka mig till andra källor för min metod. Kursböckerna *Modern Information Retrieval* och *Information Seeking in the Online Age* har översiktliga kapitel med användargränssnitt och sökmotorer, där fler viktiga namn togs upp. Med utgångspunkt från dessa sökte jag sedan upp tryckta böcker på biblioteken. I enkla sökningar i Google och Science Direct¹⁷ fann jag ytterligare intressant information, bland annat av Jakob Nielsen, som jag redan förstått var ett stort namn inom området användbarhet. En stor fördel är att detta ämne lämpar sig för webbpublicering och att många forskare är måna om att publicera sig själva på webben, med maximal användbarhet. ”Interface”, ”design”, ”HCI”, ”usability” och ”search” var sökord som ledde in på rätt spår, och jag använde dessa också för att finna ytterligare böcker i Borås högskolebiblioteks katalog.

De namn som ständigt återkommit är Jakob Nielsen, en filosofie doktor med användbarhet på webben som specialitet, Ben Shneiderman, professor i datavetenskap och grundare och chef för Human-Computer Interaction Laboratory, Gary Marchionini, professor i biblioteks- och informationsvetenskap och medlem av Human-Computer Interaction Laboratory, Jonas Löwgren, som skrivit en humoristisk och lättfattlig bok om användargränssnitt, och Carl Martin Allwood som är docent i psykologi och har fokuserat på de psykologiska aspekterna av MDI-området. Magnus Lif från Uppsala universitet har skrivit en avhandling med namnet *Adding Usability: Methods for Modelling, User Interface Design and Evaluation* som också visat sig mycket användbar.

Efter samtal med bland andra doktorand Preben Hansen, som forskar i informationsvetenskap, framgick att min typ av experiment är ovanligt och förmodligen det första i sitt slag i Sverige. Detta har gjort det svårt att hitta litteratur med omfattande metodförslag. Byeong-Min Yu och Seak-Zoon Roh har i en artikel i *Journal of the American Society for Information Science and Technology (JASIST)* koncentrerat sig på vad de kallar meny-design i söksystem som påminner om Yahoo!, där information och länkar är strukturerade och användaren tar sig fram genom att klicka på länkar snarare än att söka i fält. De har sedan jämfört tre olika webbplatser med samma information men olika design.¹⁸ Denna undersökning har till viss del varit en inspirationskälla till den statistiska beräkningen och mätningen av tillfredsställelse, men är i övrigt alltför olik denna uppsats undersökning för att användas som mall.

¹⁴ Johansson, Ann-Catrin & Kangas, Rita, 2001. *Folkbibliotekens webbplatsers grafiska användargränssnitt: kvalitativa intervjuer med webbansvariga.*

¹⁵ Lindgren, Karin, 2000. *Användarvänlig design av bibliotekskataloger på webben.*

¹⁶ Ramstedt, Martina & Ågren, Erika, 2001. *Fritt val! Om individanpassning av sökmotorers gränssnitt.*

¹⁷ <http://www.sciencedirect.com/>

¹⁸ Yu, Byeong-Min & Roh, Seak-Zoon, 2002. *The Effects of Menu Design on Information-Seeking Performance and User's Attitude on the World Wide Web.*

Löwgren beskriver den historiska utvecklingen för gränssnittsforskning och gör en indelning av olika framställningsmetoder av gränssnitt, från lågt användardeltagande till högt. På 70- och tidiga 80-talet gjordes experiment för att etablera teorier om psykologi och generella riktlinjer för design. Man använde sig av studenter och många experiment var grovt förenklade för att få statistiskt intressanta resultat.¹⁹

På mitten av 80-talet gick intresset över till mer praktiska tillämpningar och idéer och användaren placerades i centrum. Forskare iakttog verkliga användare och deras arbete och designade gränssnitt därefter. Problemet var att designers frågade användarna vad de ville ha i ett gränssnitt. Användarna kände inte till de tekniska förutsättningarna och efterfrågade då sådant de redan kände till. Resultatet blev att användarna uppfattades som veliga och att gränssnitten inte utvecklades. Detta kallar Löwgren teori-baserad (eng. theory-based) design.²⁰

Användbarhetskonstruktion (eng. usability engineering) var nästa metod, som gick ut på att formulera mätbara användbarhetsmål som uppnåddes genom att designen av ett gränssnitt tillverkades, utvärderades, modifierades, utvärderades osv. Detta efterföljdes av en än mer användarorienterad teknik som Löwgren kallar kontextdesign (eng. contextual design). Forskare iakttog och spenderade tid med användarna och gränssnitten utvecklades och utvärderades i samarbete med användarna. Detta ledde till ett ökat intresse för användarna och inlärningen av de nya gränssnitten gick då också bättre. Deltagande design (eng. participatory design) är det senaste steget där utvecklare av gränssnitt och användare arbetar tillsammans och lär av varandras områden. Den grundläggande inställningen är att gränssnitten inte blir färdiga utan ständigt kan förbättras.²¹

Löwgren lägger inga värderingar i de olika metoderna, utan anser att den aktuella situationen får avgöra vilket tillvägagångssätt som passar.²²

Marti Hearst beskriver i den några år yngre boken *Modern Information Retrieval* att forskning om gränssnitt har förändrats mycket på grund av att omständigheterna förändrats. Webben som är gratis och lättillgängligt är helt annorlunda mot att söka i dyra databaser som förr var det enda alternativet. Hearst förklarar:

- Nu söker användaren oftare i fulltextdatabaser (till exempel på webben) i stället för i katalogposter.
- Moderna system använder sig av statistisk rankning av träffar, vilket är mer effektivt i fulltextdatabaser, medan gamla undersökningar fokuserat på booleska sökningar, där träffarna kommer i slumpvis ordning bara de uppfyller sökfrågans krav.
- Idag utförs sökningar av slutanvändare, de som ska använda sig av resultatet. Förr utfördes sökningar av professionella förmedlare av information.

¹⁹ Löwgren, Jonas, 1993. *Human-Computer Interaction: What every system developer should know*, s. 35ff.

²⁰ Ibid.

²¹ Ibid.

²² Ibid.

- ”Nära skjuter visst hare.” Uttrycket ”nära skjuter ingen hare” stämmer inte i hypertextens tid. Ett icke-relevant dokument kan ligga mycket nära ett relevant. Idag kan ett sådant dokument leda vidare till ett relevant dokument med hjälp av länkar. Till exempel: en användare söker efter information om prisutvecklingen för el 2000, men finner bara prisutvecklingen för 2002. Det är ändå troligt att det finns en länk på den webbsidan som leder till en elprisets historik. Förr var ett sådant icke-relevant dokument i en träfflista inte ett steg närmare målet, utan bara ett misslyckat sökresultat.
- Sökning i en databas kan vara mycket dyr och man betalar per minut och/eller sökning. Detta påverkar också sökprocessen och användarna fick förr uppfinna metoder som gjorde sökningar snabba och/eller effektiva eftersom det var dyrt att göra om och förfina sin sökning.²³

3.1 KOPPLING TILL BIBLIOTEKS- OCH INFORMATIONSVETENSKAP

Det kan tyckas att gränssnitt, design och frågeoperatorer ligger långt bort från den traditionella biblioteksvetenskapen, men så är inte fallet idag. Under utbildningen i biblioteks- och informationsvetenskap i Borås behandlas närliggande ämnen bland annat inom kursen dokumentrepresentation, men framför allt inom kursen i Information Retrieval, ett stort forskningsområde inom informationsvetenskapen som sysselsätter sig med just sådant som informationssökning i databaser.

Gränssnitt är kanske främst en del av MDI-området som behandlar interaktion mellan människa och dator. Ämnet är alltså inte renodlat biblioteksanknutet, i denna uppsats finns också element av psykologi och datavetenskap. Magnus Lif skriver i sin avhandling också att MDI-området innefattar datavetenskap, sociologi, ergonomi mm.²⁴ Mina egna kunskaper i HTML, som används i konstruktionen av experiment-gränssnitten har jag också främst vunnit i andra utbildningar.

Alison Head, som forskar i biblioteks- och informationsvetenskap, spekulerar i framtiden för gränssnitt och sökning. Hon förutsäger att bibliotekarien och andra informationsförmedlare i framtiden återigen kommer att ägna sig åt framställandet av redskap som underlättar kunskapsorganisation och förvärvande av information, liksom när klassificeringssystemen en gång tillverkades. Bibliotekarien och informationsförmedlaren kommer att bidra med viktiga pusselbitar i tillverkningen av effektiv och användbar design för gränssnitt.²⁵ Detta är en spännande tanke och den ger ämnet ytterligare en dimension.

²³ Hearst, Marti A, 1999. User Interfaces and Visualization, ingår i *Modern Information Retrieval*. s. 266f.

²⁴ Lif, Magnus, 1998. *Adding Usability: Methods for Modelling, User Interface Design and Evaluation*. s. 3.

²⁵ Head, Alison J, 2000. *Design wise: A guide for evaluating the interface design of information resources*. s. 172f.

4. TEORI

Det finns flera teorier om vad som gör ett gränssnitt användbart eller inte. Uppsatsen tar här upp begreppet användbarhet, begreppen recall och precision förklaras, en kort historik över gränssnittens utveckling, från kommandobaserade till grafiska, ges och de viktigaste punkterna inom sökfunktioner, psykologiska faktorer, hjälpfunktioner och grafik kommer att granskas.

Avsnitten om användbarhet: kapitel 4.1, samt recall och precision: kapitel 4.2, har inspirerat till metoden för genomförandet och utvärderingen av experimentet, men jag har ändå valt att placera dem i teoridelen på grund av deras relativt teoretiska innehåll.

Kapitel 4.4, 4.5, 4.6 och 4.7 behandlar delvis samma aspekter av gränssnittsdesign men har olika infallsvinklar på ämnet. Teorier inom dessa områden inspirerade till gränssnittens design. Alla punkter inverkade inte på designen, men är ändå relevanta för uppsatsen, för att ge en övergripande bild av ämnesområdet. Fokus ligger dock på de punkter som prövades i min undersökning.

Ben Shneiderman har formulerat åtta regler (och ett nionde tillägg till dem som för uppsatsen är central) för effektiva sök-gränssnitt:

1. **Konsekvens:** Vanliga ord för till exempel ”sök” på sök-knappen bör användas. I experiment har man visat att en förändring av termen ”search” till ”query” märkbart försämrade användares förståelse och effektivitet.
2. **Genvägar:** Vana användare bör belönas med enkel navigering. Ett exempel är snabbkommandon i Windows.
3. **Feedback:** Användaren ska vara informerad om hur systemet behandlar en sökning, vilka operatorer som kan användas osv. När sökningen är klar ska det vara tydligt vad som har hämt och varför. Om sökningen har gått dåligt bör systemet ge tips om hur resultatet ska förbättras.
4. **Designa för tillslutning.** Det ska vara tydligt för användaren om han/hon har undersökt samtliga aktuella dokument.
5. **Enkla felmeddelanden:** om något är fel bör användaren bli informerad i klartext och inte med programmerarspråk.
6. **Lätt återvändo:** Det ska alltid vara lätt att backa ett steg om man ångrar något. På webben sker detta naturligt med webbläsarens bakåt-knapp
7. **Användarkontroll:** Användaren bör känna att han/hon har kontroll över sökningen. Fält ska till exempel fyllas i valfri ordning och inte påtvingas.
8. **Korttidsminnet:** Om information är utspridd över flera webbsidor, blir det svårt för användaren att hålla kvar den i minnet. Långa webbsidor som behöver ”scrollas” är också svåra att få överblick över. Trots att en kompakt presentation är svårläst, är det ändå lättare att överblicka än flera separata webbsidor.

9. (**Stora sökfält:** Sökfälten bör göras stora så att användaren fyller i långa söksträngar, vilket ökar precisionen.)²⁶

Punkter 2, 3, 4 och 5 är inte aktuella i denna uppsats men samtliga nio punkter tjänar ändå tjänar bra som en introduktion till ämnesområdet. De övriga punkterna kommer att behandlas i mer detalj nedan.

Shneiderman nämner sökfältets storlek som en parentes här, men detta betyder inte att det inte är värt att undersöka om det finns ett samband mellan sökfältets storlek och söksträngens längd.

4.1 ANVÄNDBARHET

Dumas och Redish ger i sin bok *A practical guide to usability testing* en tydlig bild av begreppet användbarhet i vid bemärkelse. Boken behandlar tekniken användbarhetskonstruktion med kvantitativa användbarhetsmål och iterativ design, det senare innebär att gränssnitten kontinuerligt evalueras och designas om. Boken fokuserar inte på gränssnitt utan tar också upp användbarhet för telefonsvarare, hissknappar, patientmonitörer för sjukhus osv.²⁷ Dumas och Redish förklarar att inlärningskurvan för typiska produkter stiger från att användaren inte behärskar produkten alls (0%) upp till ca 20% av produktens totala kapacitet och sedan stannar där. De flesta användare lär sig alltså aldrig att utnyttja en produkts fulla potential. Det spelar alltså ingen roll vad systemservicen kan göra om användarna inte orkar eller kan lära sig hur man gör det. Användare har låg tolerans mot komplicerade inläringar och nöjer sig ofta med att lära sig en produkt upp till en låg men acceptabel nivå och sedan stanna där.²⁸ (En illustration till detta skulle kunna vara programmet Word som de flesta av oss använder, utan att kunna förklara eller använda ens hälften av ikonerna och funktionerna som faktiskt finns inbyggda.)

Därför är hög användbarhet nödvändig för produkter för att de ska användas alls. I det kommersiella perspektivet tjänar ett företag på att göra sina produkter användbara eftersom de då säljer mer, stärker sitt varumärke genom att tillverka en bra produkt, och reducerar support- och inlärningskostnader.²⁹

Dumas och Redish delar in användbarhetstester i fem punkter. Dessa punkter har i varierande grad inspirerat till denna uppsats metod.

1. Huvudmålet är att öka användbarheten. Det ska också finnas specifika mål som ska uppnås.
2. Testpersonerna ska representera verkliga användare.
3. Testpersonerna ska utföra verkliga uppgifter.
4. Vad testpersonerna gör och säger ska observeras och bevaras.

²⁶ Shneiderman, Ben, 1997. Clarifying Search: A User-Interface Framework for Text Searches. (Vid utskrift s. 7ff)

²⁷ Dumas, Joseph S & Redish, Janice C, 1994. *A practical guide to usability testing*. s. 26ff.

²⁸ Ibid., s. 4ff.

²⁹ Ibid., s. 14ff.

5. Analys av data. Problemområden ska identifieras och förslag på lösningar görs.³⁰

Dumas och Redish verkar instämma i den metod som Löwgren ovan kallade användbarhetskonstruktion.

Jonas Löwgren definierar ett sätt att mäta användbarhet i fyra punkter som han kallar för REAL:

- Ett gränssnitts **R**elevans: hur väl passar gränssnittet användarens behov?
- Hur **E**ffektivt kan användarna utföra uppgifter med gränssnittet?
- Vilken **A**ttityd har användarna till gränssnitt och system?
- Hur lätt är gränssnittet att **L**ära sig?³¹

Det är med utgångspunkt från dessa definitioner flera av frågeställningarna i denna uppsats uppkommit och definierats.

4.2 RECALL OCH PRECISION

Recall och precision är två vanliga mått för mätning av sökresultat som beskrivs i *Modern Information Retrieval*. Recall är ett mått på hur stor andel av relevanta dokument i en databas som återfinns vid en sökning. Om det finns 20 relevanta dokument i en databas och man vid en given sökning finner 5 av dem är recall alltså 25%. Precision är ett mått på andelen relevanta återfunna dokument vid en given sökning. Om man vid en sökning får 10 träffar och 5 av dem är relevanta är alltså precision 50%.³² Det finns dock en viss problematik med dessa mått, vid stora databaser som webben är antalet dokument inte så begränsat att det går att veta exakt hur många relevanta dokument som existerar totalt, och därför är måttet recall svårt att tillämpa.³³ Detta gäller för denna undersökning och därför är det precision som mäts här.

I samma bok berör Marti Hearst ämnet utvärdering flyktigt, men med en del intressanta aspekter. Hon diskuterar också nackdelarna med recall-måttet då de flesta användare är mer intresserade av ett fåtal relevanta dokument än att finna samtliga 1000 relevanta dokument i en databas. Hearst påpekar också att inlärningstid, grad av fel och tid för att uppnå frågekonstruktion är viktigt att mäta. Problemet är att dessa saker är svåra att mäta med signifikanta statistiska resultat och Hearst refererar till Nielsen som i sin tur intar en informell inställning till utvärderingar och lägger tyngdpunkten på andra värden än statistik.³⁴

Det är alltså ovanligt att göra en kvantitativ undersökning av det slag som denna uppsats behandlar, men också intressant just eftersom det är så ovanligt.

³⁰ Dumas, Joseph S & Redish, Janice C, 1994. *A practical guide to usability testing*. s. 22ff.

³¹ Löwgren, Jonas 1993. *Human-Computer Interaction: What every system developer should know*, s. 52. För att Realkonceptet skulle fungera också på svenska kan översättningen ha blivit något otymplig.

Begreppet attityd är i denna uppsats översatt till tillfredsställelse, men har i grunden samma betydelse.

³² Baeza-Yates, Ricardo & Ribeiro-Neto, Berthier, 1999. *Modern Information Retrieval*. s. 75ff.

³³ *Ibid.*, s. 81.

³⁴ Hearst, Marti A, 1999. *User Interfaces and Visualization*, ingår i *Modern Information Retrieval*. s. 261f.

4.3 KORT HISTORIK

Agneta Gulz skriver i *Human IT* att gränssnitten var kommandobaserade fram till mitten av 80-talet. Användaren fick kommunicera med datorn med hjälp av bokstavskoder och sifferkoder. Koderna var ofta slumpartade och användarna var tvungna att inte bara känna igen dem, utan också aktivt kunna producera dem ur minnet. Detta stämmer dåligt med människans psykologi som kommer att beskrivas närmare i kapitel 4.5.

De grafiska användargränssnitten förändrade detta. Med hjälp av färger, former och bilder blev gränssnitten mer anpassade till det mänskliga minnet.

Gulz betonar dock att det fortfarande är viktigt att utforma varje gränssnitt så att det passar sina användare.³⁵

Magnus Lif beskriver i sin avhandling om användbarhet att det i datorernas barndom inte fanns något behov av god gränssnittsdesign. Möjligheterna var mycket begränsade när det gällde design, och bilder och ljud gick inte att använda alls. Användare och programmerare var ofta samma person och det fanns ingen anledning att göra ett system användbart för andra människor. Idag däremot ska datorprogram ofta användas av användare utan kunskaper i programmering och kanske också begränsad erfarenhet av datorer överhuvudtaget. Lif betonar också att det med dagens oändliga designmöjligheter är viktigt att anpassa gränssnitt till sina användare.³⁶ Överbelastade och röriga grafiska gränssnitt kan vara sämre än kommandobaserade.³⁷

Large, Tedd och Hartley tar också upp utvecklingen från kommandobaserade gränssnitt till grafiska men är inte odelat positiva till det senare. De ger en detaljerad beskrivning av databasvärden Dialogs³⁸ gränssnitt där den erfarna användaren kan göra mycket snabba, detaljerade och effektiva sökningar eftersom varje kommando-ord ersätter tidsödande kommunikation med ett grafiskt gränssnitt. Det går att kombinera flera instruktioner, till exempel för både sökning och presentation av resultat, i en sträng. Kommandospråket gör också att ett fåtal termer går att kombinera till mycket preciserade sökningar.³⁹ Nackdelarna med kommandobaserade gränssnitt är dock också många. Användaren måste lära sig det aktuella kommandospråket, praktisera det ofta för att minnas det och hålla isär det från andra kommandospråk som används i andra databaser. Om en sökning ska utföras i flera databaser måste användaren behärska samtliga kommandospråk som används i dem. Kommandobaserade gränssnitt förutsätter att användaren behärskar kommandospråket och funktionerna i systemet, och erbjuder sällan hjälp med formuleringar och navigering. Användare som en gång lärt sig använda ett kommandospråk kan dock föredra ett sådant gränssnitt.⁴⁰

³⁵ Gulz, Agneta, 1999. "Does one size fit all?" - kognitiv variation och IT.

³⁶ Lif, Magnus, 1988. *Adding Usability: Methods for Modelling, User Interface Design and Evaluation*. s. 3.

³⁷ Ibid., s. 16.

³⁸ Dialog Web Overview. En databasvärd är lite förenklat ett gemensamt gränssnitt för flera databaser. Just Dialog är värd för 600 databaser och används på utbildningen Bibliotekshögskolan.

³⁹ Marchionini bekräftar att experter ofta föredrar kommandobaserade gränssnitt på grund av deras effektivitet. Marchionini, Gary, 1997. *Information Seeking in Electronic Environments*. S. 43.

⁴⁰ Large, Andrew, Tedd, Lucy A & Hartley, Richard J, 2001. *Information Seeking in the Online Age: Principles and Practice*. s. 202ff.

4.4 SÖKFUNKTIONER

Vilka operatörer som kan användas och hur de ska fyllas i är avgörande för sökresultat och sökupplevelse. I detta avsnitt kommer gränssnitt att behandlas med avseende på just sökprocessen.

Ben Shneiderman skriver om att sökmotorer på webben ofta är inkonsekventa och att det är oklart hur de ska användas. Dålig design orsakar osäkerhet, felaktiga antaganden, långsamhet och oförmåga att finna relevanta dokument för användaren. Shneiderman visar i ett exempel hur en söksträng: "Hall effect" kan betyda ett tiotal olika saker beroende på sökmotor. Den kan bland annat behandlas som en fras, med eller utan skiftlägeskänslighet⁴¹ och den kan generera olika booleska operatörer som AND och OR mellan orden. Det är också ofta oklart för användaren hur resultaten rankas och vad sambandet sökfråga/träfflista är.⁴²

Shneiderman delar upp sökprocessen i fyra delar, nämligen:

1. **Formulering** (eng. formulation): Det som händer innan användaren börjar sin sökning, val av strategi.
2. **Handling** (eng. action): Användaren börjar sin sökning.
3. **Bedömning av resultat** (eng. review of results): Vad användaren läser ut av sitt resultat.
4. **Finjustering** (eng. refinement): Användaren justerar sin sökning utifrån de resultat han/hon fått och gör eventuellt en ny sökning.

Steg ett och två behandlas vidare nedan, medan steg tre blir mer aktuellt i kapitel 5.2. Steg fyra kommer inte att behandlas vidare.

Formulering kan i sin tur delas upp i fyra punkter:

- **Källor** (eng. sources): Användaren bestämmer vilka källor han/hon ska använda, det vill säga i vilken databas eller sökmotor sökningen ska ske. Detta blev inte aktuellt i denna uppsats då valet av sökmotor var förbestämt.
- **Fält** (eng. fields): Vilka fält ska fyllas i? I bibliotekskataloger går det ofta att söka efter titel, författare och ämnesord. I sökmotorer finns ibland fält för land, år, språk och liknande.
- **Söksträng** (eng. what to search for):⁴³ Användaren får här välja hur söksträngen ska vara formulerad. En fräsökning, där orden som skrivs in måste återfinnas i samma ordning och närhet, ökar precision men minskar recall. Det finns också en mängd operatörer som kan infogas, till exempel trunkeringstecken, som kan vara *, ?, \$,

⁴¹ Översättning av engelskans case-sensitive/case-insensitive. I detta fall åsyftas sökmotorns matchning eller icke-matchning av gemener och versaler. Ordet skiftlägeskänslig är godkänd av Sv. datatermgruppen.

⁴² Shneiderman, Ben, 1997. Clarifying Search: A User-Interface Framework for Text Searches. (Vid utskrift s. 1ff)

⁴³ Shneidermans begrepp "What to search for", är något vagt, jag har därför valt en översättning som stämmer mer med innehållet i den efterföljande texten.

eller något annat tecken, beroende på sökmotor, ett tecken som kan ersätta en eller flera bokstäver i ett ord. Genom att skriva *el** får man då träffar på *el*, eldistribution, elräkning osv. Booleska operatorer som AND, OR och NOT är användbara men inte alltid så logiska som de verkar. (En vanlig användare som söker information om till exempel katter och hundar skriver följande söksträng: *katter AND hundar* utan att inse att han/hon på detta sätt begränsar sökningen till dokument som behandlar bägge djuren i stället för att återvinna dokument som bara handlar om katter *och* dokument som bara handlar om hundar.⁴⁴ Bruket av operatorn AND upplevs som en utökning av sökningen och användaren tror att det ska ge fler träffar, när det i själva verket ger färre.⁴⁵) Det viktiga här är att systemet tolkar söksträngen på samma sätt som användaren avsett. Är det oklart hur till exempel *trunkeringstecknet* ska se ut gör användaren lätt fel och resultatet blir då förmodligen noll träffar.

- **Varianter** (eng. variants): Systemet kan tillåta varianter av sökord, synonymer, förkortningar osv. Vid sökning på "eldistribution" kan systemet ge träff på "kraftdistribution" som också är relevant. Det är också här viktigt att gränssnittet tydligt anger om varianter tillåts och hur de behandlas.

Handling går traditionellt till så att användaren klickar på en "sök-knapp" och inväntar ett resultat. Shneiderman diskuterar alternativ till denna metod, där ett ideal vore att användaren kan söka i en mer dynamisk miljö där man omedelbart får se resultat av en modifierad söksträng och kan finjustera sin sökning stegvis på några få sekunder.⁴⁶

Nielsen förklarar att användare i genomsnitt använder två ord i sina söksträngar och inte behärskar booleska operatorer eller avancerade sökningar. Därför bör en sökmotor kunna ge högrelevanta resultat även för mycket korta söksträngar. Avancerad sökning med booleska operatorer, parametersökning (till exempel att begränsa sökning till dokument från ett visst datum), bör erbjudas men inte vara förvalt.⁴⁷

Large, Tedd och Hartley beskriver de två typer av färdigheter en användare måste behärska för att framgångsrikt använda sig av ett datorprogram eller tjänst. Dels krävs *syntaktiska* färdigheter: förmåga att hantera musen, radera ett ord, scrolla, öppna ett dokument och så vidare. Det krävs också *semantiska* färdigheter: förmåga att förstå vilka möjligheter ett program eller tjänst har, och hur man bäst utnyttjar dessa möjligheter för att nå ett önskat resultat. Det finns användare med olika nivåer av syntaktiska och semantiska färdigheter och gränssnitt bör designas för nybörjare, experter och de däremellan. Ett sätt att göra detta på är att erbjuda både enkla och avancerade sökfunktioner i samma gränssnitt eller att, som Nielsen föreslår ovan, erbjuda ett alternativt avancerat gränssnitt.⁴⁸

Head har konstruerat guider över frågor om gränssnittet som är lämpliga att ställa. I motsats till vanliga guider av *designers* för *designers* är dessa guider utvecklade av *användare* för *användare*. Det finns tre guider uppdelade i uppdragsstöd (eng. task support), användbarhet och estetik (eng. aesthetics). Den övergripande frågan när det

⁴⁴ Nielsen, Jakob, 2001. *Användbar webbdesign*. s. 227.

⁴⁵ Hearst, Marti A, 1999. User Interfaces and Visualization, ingår i *Modern Information Retrieval*. s. 279.

⁴⁶ Shneiderman, Ben, 1997. Clarifying Search: A User-Interface Framework for Text Searches. (Vid utskrift s. 2ff)

⁴⁷ Nielsen, Jakob, 2001. Search: Visible and Simple (Alertbox May 2001).

⁴⁸ Large, Andrew, Tedd, Lucy A & Hartley, Richard J, 2001. *Information Seeking in the Online Age: Principles and Practice*. s. 199f.

gäller uppdragsstöd är om gränssnittet uppfyller användarens förväntningar och stöder användaren i hans/hennes uppdrag. Men det finns också flera ”mindre” frågor som är väl värda att besvara. Här återges ett urval:

- Passar gränssnittet användarens nivå (nybörjare/expert) och uppdrag? Är designen flexibel nog att stödja olika nivåer och uppdrag?
- Vilket slags arbete är gränssnittet designat för, och är det samma arbete som gränssnittet i praktiken kommer att användas för?
- Vilka sökfunktioner finns att använda? Trunkering, booleska kommandon, sökhistorik?
- Vem gör det största jobbet, användaren eller gränssnittet?⁴⁹

4.4.1 EN ELLER FLERA SÖKNINGAR

Det amerikanska företaget User Interface Engineering, som forskar i användbarhetsfrågor för gränssnitt, har visat att användare finner vad de söker med hjälp av sökmotorer i endast 34% av fallen. Av de som misslyckades med sin sökning var det 47% som bara sökte en gång och sedan gav upp.⁵⁰ Det är alltså vanligt att användaren ger upp efter en sökning och därför bör gränssnittet och systemet prestera ett godtagbart resultat vid första försöket.

Jakob Nielsen har undersökt samma sak men kom fram till något annorlunda siffror. Vid första försöket får 51% av användarna relevanta träffar, vid andra försöket 32% och vid tredje bara 18%. Med andra ord, lyckas inte användaren med sin första söksträng, blir resultatet inte bättre vid omformulering. Nielsen menar därför också att det är första sökningen som måste förbättras.⁵¹

4.4.2 MENYER, KNAPPAR OCH IKONER

Enligt Large, Tedd och Hartley var det första steget, mot att förenkla de kommando-baserade gränssnitten, att införa menyer. Menyer är effektiva eftersom de minskar minnesbördan för användaren då det är lättare att känna igen ett kommando än att aktivt minnas det⁵². Menyer visar också upp fler alternativ för användaren som annars begränsar sig till de kommandon han/hon redan känner till. Det blir också möjligt att visa kontextberoende hjälpsidor som relaterar till det aktuella stadiet i sökningen. Detta lämpar sig väl för den oerfarna användaren. Det finns fyra olika typer av menyer:

- **Fullskärmsmenyer** (eng. Full-screen menus): En hierarkisk uppbyggnad av menyer där användaren får klicka sig fram till olika skärmmenyer. Fördelen är att användaren blir guidad genom hela sökprocessen, nackdelen är att det blir tröttsamt att klicka sig igenom en lång sökprocess.

⁴⁹ Head, Alison J, 2000. *Design wise: A guide for evaluating the interface design of information resources*. s. 60ff.

⁵⁰ User Interface Engineering, 2001. *User Interface Engineering – People Search Once, Maybe Twice*.

⁵¹ Nielsen, Jakob, 2001. *Search: Visible and Simple (Alertbox May 2001)*.

⁵² se kapitel 4.5

- **”Dra-ner- och pop-upp-menyer”** (eng. Pull-down and pop-up menus): Den här typen av menyer guidar inte användaren utan ger mer kontroll och självständighet. ”dra-ner menyer” existerar i alla typer av mjukvara, inte bara i söksystem. Arkivmenyn i Windows är till exempel en ”dra-ner meny”.
- **Kommandomenyer** (eng. Command menus): Kommandomenyer fungerar som en blandning av grafiskt användargränssnitt och kommandobaserade gränssnitt. Kommandon och förklaringar till dem står på skärmen och användaren väljer ett eller flera kommandon från denna meny med hjälp av tangentbordet. Det går att reglera vilka kommandon som erbjuds i olika stadier av en sökning.
- **Formulär** (eng. on-screen forms): Detta är den vanliga typen av meny som de flesta sökmotorer på webben använder. Användaren fyller i fält på skärmen, med de termer som eftersöks och eventuella operatorer.⁵³

Marchionini gör en liknande uppdelning när han med inspiration från Shneiderman delar in interaktionen med system i fyra olika typer:

- **Kommandon** (eng. commands): rent kommandospråk.
- **Menyer** (eng. menus): Liknar det Large kallar fullskärmsmenyer. Användaren behöver inte behärska något kommandospråk men är också mycket begränsade i sina sökningar.
- **Formulär** (eng. form fill-ins): samma som Formulär enligt Large. Marchionini menar att formulär är en kombination av kommandon och menyer eftersom gränssnittet bestämmer *hur* boxarna/rutorna ska fyllas i, men inte *vad* som ska fyllas i.
- **Direkt manipulation** (eng. direct manipulation): Direkt manipulation innebär att interaktionen mellan användare och system är omedelbar på samma sätt som en rörelse av musen på en gång är synlig på skärmen. På samma sätt skulle resultat och feedback kunna ske simultant med sökning och skapa ett ”transparent” system där användaren kan koncentrera sig på sitt uppdrag till fullo utan att behöva tänka på gränssnittet. Detta är dock under utveckling och är ännu inget vanligt alternativ för sökmotorer.⁵⁴

Knappar, radioknappar och ikoner är vanliga i gränssnitt. Knappar är normalt rektangulära eller kvadratiska med text, till exempel ”sök-knappen”, radioknappar är små runda utrymmen som kan markeras och bli ifyllda med hjälp av musen och ikoner är en grafisk representation av en enhet. Det är lättare att minnas ikoner än knappar.⁵⁵

Det grafiska användargränssnittet är baserat på fönster, ikoner och knappar. Musen används för att direkt manipulera objekt på skärmen.⁵⁶

⁵³ Large, Andrew, Tedd, Lucy A & Hartley, Richard J, 2001. *Information Seeking in the Online Age: Principles and Practice*. s. 205-212.

⁵⁴ Marchionini, Gary, 1997. *Information Seeking in Electronic Environments*. S. 43.

⁵⁵ Large, Andrew, Tedd, Lucy A & Hartley, Richard J, 2001. *Information Seeking in the Online Age: Principles and Practice*. s. 212f.

⁵⁶ Ibid., s. 215.

4.5 PSYKOLOGISKA FAKTORER

Interaktionen mellan dator och människa påverkas av olika psykologiska faktorer. Allwood berättar i sin bok hur det mänskliga minnet fungerar och hur det bör påverka gränssnittens utformning. Boken präglas också av användbarhetsperspektivet.

Allwood förklarar att kognitionspsykologin, som intresserar sig för människans kunskaps- och tankeliv, är den psykologiska riktning som mest påverkat forskningsområdet MDI.⁵⁷ Det är också denna riktning som dominerar i hans bok.

4.5.1 OLIKA TYPER AV MINNE

Allwood skriver att det finns tre olika typer av minnessystem, sensoriska informationsbuffertar, korttidsminne och långtidsminne. Man skiljer också på tre olika faser, nämligen *inkodning* som sker när information utifrån når minnet, lagring som är *förvaring* av information och *framtagning* när minnet aktiveras och på något sätt används. De sensoriska informationsbufferterna är aktiva främst under inkodningen, men korttidsminnet och långtidsminnet används i alla tre faserna.⁵⁸

När information når en individ tas den upp i den sensoriska bufferten, ögats buffert kallas det ikoniska minnet och örats det ekoiska. Det ikoniska minnet håller information i ca en fjärdedels sekund och kan bli stört om ytterligare information tillkommer snabbt. När individen tolkar informationen och relaterar den till sina förkunskaper i långtidsminnet når den korttidsminnet.⁵⁹ Information behöver behandlas i korttidsminnet mellan 5-20 sekunder för att framgångsrikt lagras i långtidsminnet.⁶⁰ Långtidsminnet består av representationer av företeelser, föreställningar och färdigheter.⁶¹

Korttidsminnet är begränsat. Forskning har visat att en normal individ klarar att hålla fem till nio enheter, bitar av information som till exempel siffror, aktiva samtidigt. Detta medför att information i en ny skärmbild kan störa korttidsminnet och göra att minnet av den tidigare skärmbilden förstörs. Det är alltså viktigt att gränssnittsdesign tar hänsyn till korttidsminnets kapacitet.⁶² Lif beskriver korttidsminnet och långtidsminnet på samma sätt och påpekar också att det är viktigt att all information som behövs för att utföra en handling bör finnas på skärmbilden samtidigt.⁶³

Långtidsminnet är i princip obegränsat. Glömska beror inte på att långtidsminnet blivit fullt, utan på att det kan vara svårt att komma åt lagrad information. När en person söker i sitt minne kan ledtrådar göra arbetet enklare. Vid ett igenkänningstest får en testperson informationen presenterad och ska avgöra om det är rätt information. Till exempel: ”Var det den här ikonen du klickade på?” I ett återgivningstest däremot måste testpersonen själv komma fram till informationen. Till exempel: ”Hur såg ikonen ut?” Det är generellt lättare för en person att känna igen än att återge information.⁶⁴ Här kommer paralleller till Agneta Gulz påstående att grafiska användargränssnitt är lättare att lära

⁵⁷ Allwood, Carl Martin, 1991. *Människa-datorinteraktion: ett psykologiskt perspektiv*. s. 120f.

⁵⁸ Ibid., s. 15f.

⁵⁹ Ibid., s. 16f.

⁶⁰ Ibid., s. 28.

⁶¹ Ibid., s. 16f.

⁶² Ibid., s. 28ff.

⁶³ Lif, Magnus, 1998. *Adding Usability: Methods for Modelling, User Interface Design and Evaluation*. s. 4.

⁶⁴ Allwood, Carl Martin, 1991. *Människa-datorinteraktion: ett psykologiskt perspektiv*. s. 25f.

sig eftersom de bygger på igenkänning och att användaren inte aktivt behöver generera kommandon ur minnet.⁶⁵

4.5.2 FÖRFÖRSTÅELSE

Förförståelse är summan av en individs bakgrundserfarenheter och är en avgörande faktor för hur en individ tänker. Tänkande är till sin natur analogiskt. Det innebär att en individ utnyttjar sina förkunskaper för att förstå nya saker. Första gången en individ använder en dator kan han/hon relatera den till en skrivmaskin och på det sättet öka sin förståelse för datorn. Skrivmaskinen är då en *ursprungsdomän* och datorn en *tillämpningsdomän*. Allwood föreslår termen *analogiskt* eller *metaforiskt* tänkande när man på detta sätt låter kunskap från ett område påverka tänkandet inom ett annat. Det är viktigt att vara medveten om vilka ursprungsdomäner användarna av ett gränssnitt kommer att använda sig av. I exemplet med skrivmaskinen kan analogin hjälpa en användare att skriva in ord och meningar, men den kan också orsaka en missuppfattning av tangentbordets funktioner (mellanslagstangenten på en dator flyttar eller raderar text men på en skrivmaskin kan den användas för att förflytta sig på en rad) och begränsa användarens förväntningar på datorn (användaren kanske inte lär sig att göra text kursiv eller marginaljusterad eftersom det inte går på en skrivmaskin). Ursprungsdomäner kan säga både för mycket och för lite.⁶⁶

Det finns vissa ursprungsdomäner användarna med stor sannolikhet kommer att använda sig av:

- Andra delar av samma program.
- Andra program användaren känner till.
- Konventioner i naturligt språk.
- Konventioner i kommunikation ansikte-mot-ansikte.⁶⁷

Om ett gränssnitt följer konventionerna blir det lättare för användare att förstå och lära in det.⁶⁸ Om en sökmotor påminner om andra sökmotorer blir det alltså lätt för användaren att använda den. Men det finns också en risk att det uppstår fel om samma kommandon inte kan användas, och nya funktioner i den nya sökmotorn kanske inte används därför att användaren inte förväntar sig dem.

Man kan se hur analogier används i datorsammanhang för att göra program lättillgängliga. Till exempel kallas det man ser på skärmen när datorn startas för "skrivbordet" där man kan se "mappar" som kan öppnas och stängas, och "papperskorgen" där man slänger gamla dokument.

Lif beskriver hur en individ bearbetar information och utför uppgifter på olika kognitiva nivåer. Uppgifter som utförs på rutin sker på en låg kognitiv nivå, till exempel den vana bilföraren som växlar utan att tänka på det, medan mer krävande uppgifter sker på en

⁶⁵ Gulz, Agneta, 1999. "Does one size fit all?" - kognitiv variation och IT.

⁶⁶ Allwood, Carl Martin, 1991. *Människa-datorinteraktion: ett psykologiskt perspektiv*. s. 32-40. Se även Löwgren, Jonas, 1993. *Human-Computer Interaction: What every system developer should know*. s. 69ff.

⁶⁷ Allwood, Carl Martin, 1991. *Människa-datorinteraktion: ett psykologiskt perspektiv*. s. 37.

⁶⁸ Ibid., s. 37.

hög kognitiv nivå. När en individ ligger på en medveten kognitiv nivå, som till exempel vid studerandet av text kan han/hon bara bearbeta en process i taget. Ett gränssnitt bör vara designat så att det krävs en låg kognitiv nivå att utföra uppgifter med det och så att en medveten och hög nivå reserveras för viktigare beslut.⁶⁹ I fallet sökmotorer bör alltså formuleringen av söksträngar ske på en medveten nivå, medan sättet att skriva in dem bör ske på en låg, automatisk nivå.

4.6 HJÄLPFUNKTIONER

För att användaren ska få hjälp med att använda ett program eller tjänst finns ofta hjälpresurser i form av tryckta manualer, felmeddelanden, on-line manualer och backa-funktioner. I detta avsnitt kommer hjälpfunktioner i gränssnittet att behandlas.

Enligt Allwood ska en bra hjälpfunktion hjälpa användaren att ta sig ur aktuella svårigheter och att effektivt använda sig av programmets funktioner. Hjälpfunktionen ska svara på frågor om vilka kommandon som är tillgängliga vid ett givet tillfälle, förklara programmets frågor och felmeddelanden, och ge exempel på korrekta kommandon.⁷⁰

Dagens hjälpfunktioner är inte alltid korrekt anpassade till användaren. Speciellt nybörjare kan ha svårt att använda och förstå hjälpfunktionen. Det är svårt att hitta rätt hjälp, svårt att förstå den ofta alltför tekniska texten, och många gånger behandlar texten inte användarens specifika problem ändå. Hjälptexterna utgår ofta från programmets funktioner i stället för användarens uppgifter.⁷¹

Ett annat vanligt fel är att hjälptexten ersätter den tidigare skärmbilden. Detta medför att användaren måste minnas skärmbilden som används, läsa hjälptexten, återvända till den aktuella skärmbilden där kommandot ska ges, och där minnas vad som stod i hjälptexten. Detta ger en hög belastning på användarens korttidsminne. En lösning är att presentera hjälpfunktionen i ett särskilt fönster. På det sättet är hjälpfunktion och uppgiftssituation tillgängliga samtidigt. Nackdelen är att en ovan datoranvändare kan ha svårt att hantera olika fönster, och att situationen ytterligare försvåras.⁷²

Large, Tedd och Hartley beskriver att hjälpfunktioner som kräver att användaren tar sig igenom flera skärmbilder för att hitta fram till rätt hjälp oftast leder till att användaren ger upp. De påpekar också att nivån på hjälpen måste vara anpassad till sina användare, vilket kan vara extra svårt när användargruppen är så spridd som den på webben, då ord som "trunkering", och "boolesk" inte är självklara för alla. Det är vanligt att sökmotorers hjälpfunktioner håller en för hög och komplicerad nivå vilket förvirrar användaren.⁷³

Allwood ger ett exempel på enkla medel med vilka en hjälpfunktion kan förbättras. Genom att undvika datorjargong och förklara enkelt, fokusera på inläring och ge konkreta förslag och exempel på korrekta kommandon i stället för att med tekniska

⁶⁹ Lif, Magnus, 1998. *Adding Usability: Methods for Modelling, User Interface Design and Evaluation*. s. 5.

⁷⁰ Allwood, Carl Martin, 1991. *Människa-datorinteraktion: ett psykologiskt perspektiv*. s. 68.

⁷¹ Ibid., s. 68f.

⁷² Ibid., s. 68f.

⁷³ Large, Andrew, Tedd, Lucy A & Hartley, Richard J, 2001. *Information Seeking in the Online Age: Principles and Practice*. s. 229f.

termer förklara hur programmet fungerar kan hjälpfunktionen bli verkligt hjälpsam och minimera felnivåer.⁷⁴

En backa- eller ångra-knapp gör att en användares misstag inte får allvarliga konsekvenser, då det blir lättare att komma ur felsituationer. Detta leder till mindre datorrädsla och att användaren vågar experimentera och lär sig mer om ett program. Om ett kommando är oåterkalleligt bör detta anges tydligt. Ett exempel på ett sådant kommando är "send" i mail-program.⁷⁵ På webben har själva webbläsaren både stopp- och backa-knapp vilket gör ytterligare liknande knappar i gränssnittet onödiga.

Språket i gränssnittet är också viktigt. Galitz som är expert på grafisk gränssnittsdesign förklarar att texten i exempelvis felmeddelanden måste vara skriven med ett språk som inte är typiskt inom datorvärlden, och därför svårt att förstå för den genomsnittlige användaren. Det får ej heller vara ett språkbruk som betyder något annat utanför datorvärlden. Meningarna måste vara, korta, tydliga, icke-auktoritära, icke-straffande och icke-hånfulla. Viktiga saker som användaren bör lägga på minnet bör också komma tidigt i texten. Språket får gärna vara personligt och påminna om talspråk när det gäller enkelhet.⁷⁶

4.7 GRAFIK OCH TYDLIGHET

I detta kapitel dominerar Jakob Nielsen, vilket beror på att han är det största namnet inom användbarhet på webben. I nästan samtliga böcker om gränssnitt nämns han som en expert. Niensens bok *Användbar webbdesign* behandlar konstruktionen av användbara webbplatser för företag. Trots detta är många av de punkter han tar upp relevanta också för den typ av gränssnitt för sökningar som denna uppsats behandlar. Nielsen har en välbesökt webbplats där han publicerar artiklar om användbarhet, gränssnitt, webben, intranät och andra närliggande ämnen. Vissa av dessa artiklar kommer också att behandlas här.

Bildskärmen utnyttjas generellt dåligt enligt Nielsen. 31% av skärmens punkter utgörs av operativsystemets och webbläsarens kontroller. Reklam och navigeringsmenyer tenderar också att ta upp mycket plats. Helt outnyttjat utrymme kan däremot tjäna en funktion genom att leda ögat och motverka en alltför kompakt framställning. Nielsen anser att det verkliga innehållet bör fylla minst 50% av en webbsida.⁷⁷

Webbsidor måste också fungera oberoende av upplösning, skärmstorlek, webbläsare och typ av dator. En webbsida som designats för en färgskärm med 1024x768 upplösning kommer att se helt annorlunda ut i en liten svartvit hand-dator till exempel.⁷⁸ Det är också viktigt att inte använda sig av nya icke-standardiserade koder. Det finns HTML-kod som fungerar i Netscape men inte Internet Explorer och tvärtom. Om man använder sig av de nyaste versionerna av HTML och CSS⁷⁹ går det dessutom inte att se sidorna korrekt i äldre versioner av webbläsare. Att använda sig av kod som inte

⁷⁴ Allwood, Carl Martin, 1991. *Människa-datorinteraktion: ett psykologiskt perspektiv*. s. 72f.

⁷⁵ Ibid., s. 82f.

⁷⁶ Galitz, Wilbert O, 1997. *The essential guide to user interface design: an introduction to GUI design principles and techniques*. s. 540-551.

⁷⁷ Nielsen, Jakob, 2001. *Användbar webbdesign*. s. 18ff.

⁷⁸ Ibid., s. 25ff.

⁷⁹ Cascading Style Sheets, eller *formatmallar* på svenska, en teknik som används för att få större kontroll över det typografiska utseendet på en webbsida och för att separera utseende och innehåll på webben.

fungerar i de vanligaste webbläsarna stänger ute många användare. Om koden dessutom gör att datorn "kraschar" kan besökare bli helt avskräckta från webbplatsen.⁸⁰ Nielsen förklarar att 10% av användarna använder gamla program och datorer, det är tillräckligt många för att det ska vara onödigt att stänga dem ute.⁸¹

4.7.1 NAVIGATION

Länkar är en av webbens centrala byggstenar och de kan leda till webbsidor inom webbplatsen, *strukturella navigeringslänkar*, eller ut till andra relaterade webbsidor och webbplatser, *associativa länkar*. Länkad text kan fungera som nyckelord för användaren att fästa ögonen på när han/hon ögnar igenom texten på en webbsida. Det är därför viktigt att dessa texter inte blir för långa och att de är informationsbärande. "För information om elpriset, [klicka här](#)" är en dålig länk eftersom det inte säger något om var länken leder för en användare som skummar en webbsida. "Vi har ytterligare information om [elpriset](#)" drar däremot blickarna till elpriset och leder användaren rätt. Det är också bra att förtydliga ytterligare var länken leder genom en bit förklarande text, till exempel "Där kan du se förra årets statistik i tabeller". Användare vill inte klicka på länkar som inte säkert leder till något av intresse, särskilt inte om svarstiderna kan bli långa. Det går också att lägga till ett TITLE-attribut i A-taggen. Det innebär att lägga in en beskrivning av en länk inuti koden för länken, vilket gör att beskrivningen syns som en liten ruta när användaren för muspekaren över länken. Lämpliga beskrivningar kan vara namnet på webbplatsen länken leder till, en förklaring till varför länken är relaterad till den aktuella sidan, eventuella varningar om problem som kan uppstå på den länkade webbsidan, som användarregistrering eller tung grafik.⁸²

För interna, strukturella länkar är det viktigt att vara tydlig med vart länken leder, och att på den länkade sidan vara tydlig med den information som har fått användaren att klicka på länken från början.⁸³

Länkfärger är en annan navigeringshjälp som inte bör underskattas. Standardfärger för länkar är blått, medan besökta länkar visas i rött eller lila. Detta gör att det är lätt att se vilka länkar användaren redan besökt. Om länkfärgerna byts ut tappar användaren lätt greppet om vilka delar han/hon utforskat och kan lämna webbplatsen i förtid, eller få svårt att hitta tillbaka till en länk som han/hon funnit värdefull.⁸⁴ Det är så vanligt med okonventionella länkfärger idag att användare blir förvirrade av vanliga understrykningar i text och tror att det är länkar.⁸⁵ Jag har också märkt att det är vanligt att webbdesigners helt tagit bort understrykningar från länkarna vilket också försvårar navigering.

Svarstider, tiden från att man klickar på en länk eller sök-knapp till att nästa sida är färdigladdad, är avgörande för användbarheten. För att en användare ska kunna röra sig obehindrat på webben måste svarstiden vara under en sekund. För att en användare ska kunna hålla kvar uppmärksamheten, och inte börja göra något annat, bör svarstiden inte

⁸⁰ Nielsen, Jakob, 2001. *Användbar webbdesign*. s. 31f.

⁸¹ Ibid., s. 97.

⁸² Ibid., s. 51-62.

⁸³ Ibid., s. 65f.

⁸⁴ Ibid., s. 62f.

⁸⁵ Nielsen, Jakob, 1999. "Top Ten Mistakes" Revisited (*Alertbox May 1999*).

överskrida 10 sekunder. Med bredband är inte detta något stort problem, men för användare med modem kan ”tung” webbsidor vara mycket frustrerande.⁸⁶

Detta är en stor anledning till att vara försiktig med bilder på webben, eftersom de är stora filer som orsakar långa svarstider. Det är bättre att göra webbsidan grafiskt tilltalande med hjälp av färger och typsnitt och eventuellt vissa enkla och återkommande bilder (som inte behöver laddas flera gånger) enligt Nielsen.⁸⁷

Head tar i sin guide om användbarhet också upp frågan om navigering och ställer bland annat följande frågor:

- Hur komplex är navigationen?
- Är navigationens layout uppenbar och omedelbar?
- Finns det visuella ledtrådar som berättar vad man kan göra och hur man ska navigera?
- Går det att ångra steg i navigationen?
- Är det för många steg/klick för att kunna genomföra en uppgift?⁸⁸

4.7.2 ESTETIK

Design och utseende är en smaksak, men det finns ändå vissa riktlinjer som bör efterföljas för att ett gränssnitt ska vara användbart.

Heads guide tar upp några aspekter av den rena estetiken och börjar med det övergripande kriteriet att gränssnittet ska kommunicera visuellt med användaren, hjälpa honom/henne med sin uppgift samtidigt som gränssnittet ska hålla informationsflödet på en lagom nivå. Mer specifika frågor att ställa är:

- Passar designen sitt syfte?
- Gör färg och layout att användaren kan fokusera, eller blir han/hon mer splittrad?
- Är skärmen överlastad med grafik och ”plotter”?
- Är typografin (storlek och typsnitt) effektiv, är det lätt att läsa?
- Är det roligt att använda och titta på gränssnittet?⁸⁹

Large, Tedd och Hartley är än mer specifika i sina anvisningar. Det får inte finnas för mycket information på en skärmbild, de refererar till Cherry och Cox som föreslår en maximal ”tyngd” av 600 tecken per skärmbild. Användaren bör inte heller tvingas scrolla då det finns risk att han/hon då tycker att det känns obekvämt och blir desorienterad. Ett sätt att begränsa informationen på skärmbilden är att använda

⁸⁶ Nielsen, Jakob, 2001. *Användbar webbdesign*. s. 42ff.

⁸⁷ Ibid., s. 46ff.

⁸⁸ Head, Alison J, 2000. *Design wise: A guide for evaluating the interface design of information resources*. s. 63.

⁸⁹ Ibid., s. 64.

förkortningar och fackspråk men detta kan ske på bekostnad av tydligheten och leda till att nybörjare blir utestängda.⁹⁰

Typsnittsstorleken bör vara stor, speciellt för att underlätta för äldre användare, samtidigt får det plats mer information på den enstaka skärmbilden om typsnittsstorleken är något mindre. Balansen är viktig. Typsnitt kan delas in i serif och sans-serif. Serifer är de små strecken längst ner på bokstäverna som hjälper till att leda ögat vid läsning i tryckta medier. På webben passar dock sans-serif bättre.⁹¹ Eftersom upplösningen på bildskärmar generellt är betydligt sämre än för tryckt text är det viktigt att teckenformerna är enkla och tydliga.⁹² Galitz rekommenderar att aldrig använda fler än två olika typsnitt, två stilar (som normalt och kursivt), två vikter (som normalt och fetstil) eller tre storlekar. Kursiv stil fungerar generellt dåligt på en skärm och blir svårsläst. Blinkande och rörlig text bör bara användas om det är viktigt att användaren uppmärksammar något omedelbart, det är annars tröttsamt och irriterande för användaren.⁹³

Large, Tedd och Hartley varnar också för multimedia-inslag, som animeringar och musik, i gränssnittet, som kan distrahera användaren.⁹⁴ Nielsen påpekar också att animerade bilder och blinkande text är mycket störande för det mänskliga sinnet och gör det svårt att läsa den vanliga texten.⁹⁵

Färg kan göra ett gränssnitt trevligare och lättare att arbeta med, men det finns också fallgropar. Galitz skriver att ett dåligt färgval kan leda till långsammare läsning, missförstånd och trötthet hos användare. En dålig användning av färg är värre än ingen färg alls enligt honom. När det gäller färg bör antalet färger begränsas till det absolut nödvändiga, räcker det med två färger ska inte fler användas. Som tumregel gäller att det aldrig ska vara fler än 4 färger på samma skärmbild. Det är viktigt att uppnå en hög kontrast mellan text och bakgrund. Mörk text mot mörk bakgrund är svårt att läsa. För att anpassa ett gränssnitt till användare som kan vara färgblinda, ha svartvita skärmar eller arbetar i dålig ljussättning bör alla gränssnitt först designas i svartvitt och först efter det färgläggas. Vissa färger har också kulturella betydelser som designern måste ta hänsyn till. Rött betyder vanligen fara, medan grönt betyder okej, men det finns kulturella skillnader mellan olika nationaliteter och också yrkesgrupper som inte bör ignoreras. En viss färg kan betyda vitt skilda saker för en präst, en bankman och en sjuksköterska.⁹⁶

⁹⁰ Large, Andrew, Tedd, Lucy A & Hartley, Richard J, 2001. *Information Seeking in the Online Age: Principles and Practice*. s. 221f.

⁹¹ Sans är franska för ”utan”, och sans-serif är alltså typsnitt utan serifer. I denna uppsats kan man se exempel på serif-typsnitt i brödtexten, och sans-serif-typsnitt i rubrikerna.

⁹² Sandin, Åsa, 2000. *Stilguide för webb*.

⁹³ Galitz, Wilbert O, 1997. *The essential guide to user interface design: an introduction to GUI design principles and techniques*. s.129ff och 191ff .

⁹⁴ Large, Andrew, Tedd, Lucy A & Hartley, Richard J, 2001. *Information Seeking in the Online Age: Principles and Practice*. s. 225f.

⁹⁵ Nielsen, Jakob, 1996. *Top Ten Mistakes in Web Design (Alertbox May 1996)*.

⁹⁶ Galitz, Wilbert O, 1997. *The essential guide to user interface design: an introduction to GUI design principles and techniques*. s. 486-500, regeln om fyra färger står på s. 193.

5. METOD

Nedan följer en beskrivning av metoden som används i denna uppsats. För att motivera och förklara denna metod kommer också en del relevant litteratur att behandlas.

5.1 KVANTITATIV OCH KVALITATIV METOD

Det har generellt varit svårt att finna material om den här typen av undersökning där resultatet till största delen mäts kvantitativt, men utan att föreslå förbättringar till modifierade gränssnitt och där det inte finns något kommersiellt intresse bakom undersökningen. Det är också svårt att hitta underlag till en statistisk metod. Om det därför finns brister i metoden är detta en förklaring.

Flera magisteruppsatser utvärderar gränssnitt på kvalitativa sätt genom att fråga webb-användare och designers om deras upplevelser.⁹⁷ I nästan samtliga böcker finns också en kommersiell utgångspunkt, som utgår från att framställaren av gränssnitt arbetar på uppdrag av ett företag som vill förbättra sina anställdas produktivitet med ett användbart gränssnitt, eller ett företag som vill utveckla ett program eller en service med hög användbarhet för försäljning. Jakob Nielsen har i sin bok *Användbar webbdesign* skrivit om hur kommersiella webbplatser ska testas internationellt för att förhindra kulturella och språkliga missförstånd⁹⁸. Böckerna av Dumas och Redish, Marianne Rudisill, Magnus Lif och Alison Head är också märkbart kommersiellt inriktade vilket stundtals har gjort dem svåra att applicera på denna uppsats.

Lif förklarar att användarprestation kan mätas kvantitativt genom att låta en grupp användare utföra förbestämda uppgifter och mäta tid och felnivåer. Fördelen är att resultatet då kan mätas i siffror vilket gör det lätt att jämföra olika gränssnitt. Nackdelen är att det kan vara svårt att ”läsa mellan raderna” när resultatet består av enbart siffror. Då är enkäter en metod för att få ett mer kvalitativt mått på användarnas subjektiva tillfredsställelse och anledningar till problem. Det finns fördelar och nackdelar med båda metoderna och för en heltäckande evaluering av gränssnitt bör de kombineras.⁹⁹ Så har också skett i denna uppsats.

Nielsen skriver i en artikel på webben att han avråder från kvantitativa metoder i gränssnittsevaluering men när det används är det ofta för att mäta följande tre kvaliteter:

- Tiden det tar att utföra en uppgift.
- Felfrekvens.
- Användarens subjektiva tillfredsställelse.¹⁰⁰

⁹⁷ Jag tänker här främst på Lindgren, Karin, 2000. *Användarvänlig design av bibliotekskataloger på webben* och Johansson, Ann-Catrin & Kangas, Rita 2001. *Folkbibliotekens webbplatser grafiska användargränssnitt: kvalitativa intervjuer med webbansvariga*.

⁹⁸ Nielsen, Jakob, 2001. *Användbar webbdesign*, s. 315-344.

⁹⁹ Lif, Magnus, 1998. *Adding Usability: Methods for Modelling, User Interface Design and Evaluation*. s. 19f och s. 26.

¹⁰⁰ Nielsen, Jakob, 2001. *Usability Metrics (Alertbox Jan. 2001)*.

Metoden i denna uppsats strävar efter svar på dessa tre frågor där felfrekvens får motsvaras av precision.

5.2 METOD I DENNA UPPSATS

Efter en kort genomgång av uppsatsens metod kommer mer utförliga förklaringar och motiveringar till de olika delarna av metoden.

Genom att designa tre olika gränssnitt och koppla dem till Altavista skapades ett experimentmaterial. Tre grupper på 5 personer vardera utprovade gränssnitten genom att utföra sökuppdrag av olika slag i dem. Varje grupp fick varsitt gränssnitt. 20 sökuppdrag inom vitt skilda ämnen utfördes en utsatt dag och tid¹⁰¹. Testpersonerna fick sedan subjektivt relevansbedöma de 10 första träffarna för att möjliggöra en mätning av precision. Tiden som gick mellan att testpersonerna fick sina sökuppdrag till färdigformulerad söksträng mättes. Söksträngarna registrerades också, vilket möjliggjorde en mätning av deras längd. Testpersonerna svarade också på en mindre enkät för att förklara hur de upplevde sökningen och gränssnittet. Testpersonernas tryckta instruktioner, sökuppdragen och enkäten är bifogade som bilaga 1, bilaga 2 och bilaga 3.

Jakob Nielsen skriver att det viktigaste är att första träffsidan visar relevanta träffar eftersom användare sällan tittar längre än så.¹⁰² Eftersom vanliga sökmotorer visar tio träffar på första sidan, är det alltså de tio första träffarna som är centrala för att sökningen ska tillfredsställa en användare. Precisionen beräknades alltså vid DCV 10 (DCV står för Document Cutoff Value).

Ben Shneiderman skriver i sin artikel *Clarifying Search: A User-Interface Framework for Text Searches* att en vanlig strategi för användare vid sökuppdrag är att först göra en snabb sökning för att göra en uppskattning av hur mycket material det finns och vilka sökord som är lämpliga, för att sedan göra en finjusterad andra sökning¹⁰³. Den här användarmetoden skulle gjort min undersökning otymplig eftersom det då skulle vara svårt att mäta tiden från läsning av sökuppdrag till färdig söksträng. Precisionen blir också närmast irrelevant då den snarare skulle handla om tålamod än om gränssnittets utformning. Därför mättes endast den första söksträngens resultat. På detta sätt uppstod också ett mått på gränssnittets omedelbarhet, ett bra gränssnitt ska gå att använda och förstå utan att användaren ska behöva pröva sig fram. I kapitel 4.4.1 beskrevs också hur viktig den första sökningen är.

För varje enskilt sökuppdrag fick testpersonen fylla i sin relevansbedömning i form av kryssrutor på den webbsida jag tillhandahållit med sökuppdragen. Användarna gjorde alltså sina egna relevansbedömningar, eftersom undersökningen strävar efter att vara så naturlig som möjligt och efterlikna verkligt webb-beteende. En användare måste ju i verkliga livet alltid själv avgöra om en sökning varit lyckad eller inte. En praktisk aspekt är också att det skulle varit omöjligt för mig att personligen relevansbedöma närmare 3000 träffar (15x20x10), och också en sådan relevansbedömning hade varit subjektiv.

¹⁰¹ 2003-03-28, klockan 09.00-12.00.

¹⁰² Nielsen, Jakob, 2001. *Search: Visible and Simple (Alertbox May 2001)*.

¹⁰³ Shneiderman, Ben, 1997. *Clarifying Search: A User-Interface Framework for Text Searches*. Detta kommer att diskuteras vidare under sökmöjlighet/sökspråk.

Relevansbedömning är ett omdebatterat ämne inom Information Retrieval, något som Mizzaro visar i en artikel om relevansbedömningens historia. Han menar dock att systemrelevans, en matematiskt beräknad relevansbedömning som bygger på likheter mellan sökfråga och dokument, inte är realistisk.¹⁰⁴ Mizzaro menar att relevans är subjektivt och att det finns olika sätt att mäta den på. Vanliga metoder är att låta användaren gradera relevans på olika typer av skalor, men den binära metod, som används i denna uppsats, där användaren gör en bedömning av relevant/icke-relevant, är standard.¹⁰⁵

Efter denna undersökning fick testpersonerna också fylla i en mindre enkät med frågor om själva gränssnittet. Testpersonerna fick här fylla i radioknappar för att markera sina ställningstaganden. Allra sist kom en textruta där testpersonerna fick fylla i allmänna kommentarer.

Precision, tid, personlig tillfredsställelse och söksträngslängd för varje gränssnitt mättes.

5.2.1 DESIGN AV GRÄNSSNITTEN

Mycket av den litteratur som skrivits i ämnet gränssnittsdesign bygger på att man undersöker en användargrupp och sedan utvecklar ett gränssnitt i samråd med sin aktuella grupp. Carl Martin Allwood är en av de som förespråkar en sådan arbetsmetod, som han kallar iterativ design.¹⁰⁶ Löwgren har också visat hur utvecklingen gått mot ett allt större deltagande av användarna.¹⁰⁷ Av naturliga skäl fungerar dock inte denna metod för mig då jag ska tillverka gränssnitten själv, utvärdera och sedan avsluta mitt arbete, i stället för att gå vidare och förbättra mina gränssnitt efter testpersonernas önskemål.

Lif delar upp sättet att betrakta design i fyra varianter:

- **Hantverk** (eng. craft). Det finns ingen generell metod eftersom varje designprojekt är unikt. Det krävs en bra designer för att göra ett bra gränssnitt.
- **Kognitiv konstruktion** (eng. cognitive engineering). Teorier om informationsbearbetning och problemlösning appliceras på gränssnittsdesign. Man tar till exempel hänsyn till hur lång tid det tar att hitta en viss tangent på tangentbordet.
- **Teknisk stil** (eng. technologists). För att befria programmerare från designproblem förses de med automatiska verktyg för gränssnittsutveckling.
- **Fördjupad mjukvarukonstruktion** (eng. enhanced software engineering). Här används uppgiftsanalys (eng. Task Analysis), som är en metod för analys av användarnas arbete där man delar upp och beskriver deras uppgifter i hierarkiska termer, för att utveckla gränssnittet.¹⁰⁸

¹⁰⁴ Mizzaro, Stefano, 1997. *Relevance: The Whole History*. s. 812

¹⁰⁵ *Ibid.*, s. 814f

¹⁰⁶ Allwood, Carl Martin, 1991. *Människa-datorinteraktion: ett psykologiskt perspektiv*. s. 116.

¹⁰⁷ Löwgren, Jonas 1993. *Human-Computer Interaction: What every system developer should know*, s. 35ff.

¹⁰⁸ Lif, Magnus, 1998. *Adding Usability: Methods for Modelling, User Interface Design and Evaluation*. s. 15, Task Analysis står beskrivet på s. 11.

Min ambition var att göra så professionella gränssnitt som möjligt och utgå från hantverk, kognitiv konstruktion och fördjupad mjukvarukonstruktion i den mån det var möjligt. Jag är dock medveten om att jag inte är professionell och att det är möjligt att jag begått misstag som kommit att påverka undersökningen. "Hantverk-synen" på design säger att det är svårt för mig att lyckas framställa fullgoda gränssnitt. Den överlägset vanligaste metoden i litteraturen för att designa gränssnitt är den så kallat iterativa.¹⁰⁹ Lif rekommenderar också att gränssnittsdesign utförs av ett team av experter med stor erfarenhet.¹¹⁰ Inget av detta var en möjlighet i denna uppsats. Gränssnitten designades dock med utgångspunkt från innehållet i kapitel 4.4, 4.5, 4.6 och 4.7.

Lif poängterar att estetiken i ett gränssnitt inte är helt oviktig även om fokus ligger på användbarhet. En användare måste trivas med utseendet och känslan i ett gränssnitt för att välja det bland konkurrerande alternativ. Utseendet bör också passa funktionen. En sökmotor för popmusik bör se annorlunda ut än en databas för informationsvetenskap.¹¹¹ Jag har lagt viss vikt vid den estetiska aspekten i min design.

5.2.2 GRÄNSSNITTENS UTFORMNING

För att få inspiration till gränssnitten undersökte jag olika populära sökmotorer. En omedelbar iakttagelse är att många gränssnitt inte är särskilt användbara. Det är vanligt med överlastade och svåröverskådliga gränssnitt och hjälpfunktionerna är bristfälliga.

Min målsättning har varit att tillverka ett "dåligt" gränssnitt, ett "bra", och ett "mellan" som lånar från de bägge ytterligheterna. Målet har också varit att göra det "dåliga" gränssnittet begränsat att söka i, utan att göra det så begränsat att vissa sökuppdrag bara går att utföra i det "bra" gränssnittet. Alla operatörer och funktioner ska kunna användas i samtliga tre gränssnitt. Altavista har tre olika lägen för sökning: "enkel sökning", "mer precision" och "avancerat". Dessa tre har inspirerat till uppsatsens tre gränssnitt. Variablerna för tid, domän och filtyp som förekommer i Altavistas avancerade gränssnitt har dock helt tagits bort, för att göra de tre gränssnitten jämlika. Mitt resonemang har varit att jag ville mäta designens inverkan på sökningarna. Om ett gränssnitt också skulle ha fler möjligheter till god precision skulle resultatet ha blivit uppgjort och ointressant.

"Dåliga gränssnittet"¹¹²

Det "dåliga gränssnittet" kallas härnäst gränssnitt 1. Här valde jag att låtsas ett godis-tema för att få tillräckligt mycket irrelevant information och reklam som fyllde webbsidan och gjorde den dåligt användbar. Detta kan tyckas överkligt men efter min undersökning av olika söktjänster föreföll designen inte så otänkbar. På till exempel Spray-portalen¹¹³ finns en mängd information och länkar till olika tjänster som inte har med söktjänsten att göra, och själva sökrutan är liten och i skymundan.

¹⁰⁹ Se *Litteratur och tidigare forskning* i denna uppsats

¹¹⁰ Lif, Magnus, 1998. *Adding Usability: Methods for Modelling, User Interface Design and Evaluation*. s. 18.

¹¹¹ *Ibid.*, s. 19.

¹¹² Bilder på de tre gränssnitten finns i Bilaga 3.

¹¹³ <http://www.spray.se>

Shneiderman betonar konsekvens och tydlighet. Han har inspirerat till en minimal sökruta i gränssnitt 1, som enligt honom uppmuntrar till korta ineffektiva söksträngar.¹¹⁴ Söksträngens förhållande till sökfältet kommer ju också att undersökas. För att justera sökfältets storlek med HTML används ett attribut kallat "SIZE". Detta värde ska motsvara antalet tecken som får plats i rutan. I detta gränssnitt sattes sökfältets storlek till 20. Detta betyder dock inte att det finns en begränsning av hur många tecken som kan skrivas in. Det påverkar bara hur många tecken användaren kan se samtidigt.

Vidare inspirerat av Shneiderman är konsekvensen obefintlig, sök-knappen kallas "kör!" och länken till hjälpfunktionen ligger längst ner till höger, långt ifrån själva sökfunktionen, vilket inte är vanligt i andra sökmotorer. Hjälpfunktionen är svår-åtkomlig och meningen var att det skulle vara ottydligt för användaren hur en sökning skulle utföras och vad konsekvensen av olika operatörer blev.

Nielsen påpekar också att användare generellt gör korta söksträngar vilket ger dålig precision. Han skriver också att skärmutrymmet ofta utnyttjas dåligt. I gränssnitt 1 har därför det centrala, själva sökretan, gjorts mycket liten, medan oviktig information och bilder tar upp en absolut majoritet av skärmutrymmet. Niensens regel om att webbsidan ska fungera på olika skärmar och upplösningar är också bruten mot. Gränssnitt 1 är anpassad till upplösningen 1024x768, vilket gör den otymplig på mindre skärmar, upplösningen på sidan vid experimentet var låst till 800x600. Ingen hänsyn har heller tagits till att färgerna ska fungera i svartvitt läge. För att öka svarstiden är webbsidan överlastad med bilder.

Nielsen betonar navigationens tydlighet. Tvärtemot hans rekommendationer är länkarna formulerade som "Klicka här" och "läs mer". Webbsidan designades så att det skulle vara svårt att förstå vart olika länkar leder. Den viktigaste länken som leder användaren till hjälpfunktionen, är designad för att vara ottydlig och svår att hitta, ljusgrå mot vit bakgrund, och är också placerad i ett sammanhang av godisreklam i stället för övrig relaterad information. Länkarna är angivna i okonventionella färger och alla understrykningar är borttagna.

Också här kan det tyckas "överkligt" med en så otillgänglig sökfunktion, men ett så populärt exempel som Google, som har ett avskalat och snyggt gränssnitt, har mycket otillgängliga hjälpfunktioner. För att få hjälp angående vilka sökoperatörer som kan användas, krävs det att användaren klickar först på "Allt om Google", sedan på "Google finesser", sedan "sökhjälp" och sist "Google grunder" eller "förfinna din sökning", sammanlagt fyra klick. Hjälpen presenteras sedan på en lång sida som är svår att överblicka. Brittiska Looksmart förefaller sakna hjälpfunktion helt.¹¹⁵

Large, Tedd och Hartley och Nielsen varnar för animeringar. Det finns tre animerade reklambilder och två snurrande logotyper nära sökfältet i gränssnitt 1. De varnar också för överflöd av information och scrollning, båda dessa fel återfinns i gränssnittet.

Galitz som förordar enhetlighet har inspirerat till användandet av sex olika typsnitt varav tre är serif-typsnitt som är svåra att läsa på skärm. Det finns ett överflöd av färger i starka nyanser, typsnitten har flera stilar och vikter, kontrasterna i texten är stundtals

¹¹⁴ Eftersom alla faktorer som påverkar gränssnittets användbarhet är genomgångna i teori-kapitlet finns det inga repetitioner av fotnoter här. Jag har dock försökt att vara tydlig med att hålla isär författarna i olika stycken.

¹¹⁵ <http://www.looksmart.co.uk>

mycket dåliga, storleken på texten är liten och svårläst. Det förekommer också kursiv stil.

Allwood har visat hur en användare bör ha hjälpfunktionen tillgänglig på samma skärm som uppgiften ska utföras i. Att placera hjälpfunktionerna på en ny skärmbild belastar användarens korttidsminne för mycket. Enligt Lif kan det också tänkas att en sådan krävande hjälpfunktion skulle ta användarens medvetna kognitiva nivå i anspråk och göra formuleringen av effektiva söksträngar svårare.

Hjälpfunktionerna ligger tre klick bort i gränssnitt 1, och presenteras på nya skärmbilder. Det finns inte heller någon länk tillbaka till sökningen, utan användaren måste själv hitta tillbaka genom att backa. Detta gör att användaren måste klicka minst sex gånger för att ta sig till hjälpfunktionen och tillbaka till sin sökning. Hjälpfunktionerna presenteras i samma design som förstasidan, med förvirrande typsnitt, färger och bilder. Hjälpen är också teknisk och förklarar inte termer som ”boolesk” och ”trunkering”, vilket enligt Allwood gör att den blir svår att förstå för användaren. Jag har använt mig av Altavistas egna hjälptexter, men har tagit bort de illustrerande exemplen och bara behållit de praktiska anvisningarna, vilket gör att hjälptexten fokuserar på systemet och inte på användarens inläring. Till exempel NEAR förklaras i Altavista så här: ”Hittar dokument som innehåller båda de specificerade orden eller fraserna inom ett avstånd av tio ord från varandra. Sill NEAR potatis hittar dokument om rätten sill och potatis, men förmodligen inte dokument om potatis i andra sammanhang”. I gränssnitt 1 är bara den första meningen med i hjälpfunktionen, medan exemplet som förtydligar funktionen är borttaget.

Enligt Allwood kommer användaren att använda andra liknande programs konventioner som ursprungsdomän, därför har jag strävat efter att göra gränssnitt 1 olikt konventionella sökmotorer.

Jag har också strävat efter att få negativa svar på Heads olika guiders frågor. Där kan man se att gränssnitt 1 inte är flexibelt, eftersom det bara finns en slags sökning, det är också användaren som får göra det största jobbet genom att formulera booleska och liknande söksträngar, i stället för att fylla i lämpliga fält som kan ge stöd åt sökformuleringen. Navigationen är också oklar, eftersom hjälpfunktionerna ligger långt bort och har ologiska länknamn, och det är för många klick för att utföra en uppgift. Designen passar inte sitt syfte: att hjälpa fram till en god sökstrategi, i stället fokuserar den på reklam och information om godis vilket är helt irrelevant. Skärmen är överlastad med grafik och ”plotter”, typografin är ineffektiv och användaren kan bli splittrad i stället för fokuserad.

”Bra gränssnittet”

Det bra gränssnittet är numrerat som Gränssnitt 3.¹¹⁶ I detta gränssnitt ville jag tillhandahålla både enkel och avancerad sökning eftersom Nielsen, Head, Shneiderman och Large, Tedd och Hartley betonat vikten av att låta användaren arbeta på sin egen nivå, med möjlighet att välja mellan olika grader av enkla och avancerade gränssnitt. Användaren ska känna att han/hon har kontroll. Det avancerade gränssnittet är förvalt men med en tydlig länk till det enkla gränssnittet.

¹¹⁶ Gränssnitten presenteras i ”fel” ordning eftersom ”mellangränssnittet” är en kombination av gränssnitt 1 och 3. Det skulle vara svårt att motivera valen till ”mellangränssnittet” utan att först presentera gränssnitt 3.

Det avancerade gränssnittet består av fyra sökfält där användaren kan fylla i att ”alla dessa ord”, ”exakt den här frasen”, ”vilket som helst av dessa ord” och/eller ”inget av dessa ord” ska förekomma i ett relevant dokument. Syftet var att användaren skulle kunna komponera kraftfulla söksträngar utan att behöva lära sig och använda olika operatörer. För den som ändå ville det fanns ett väl tilltaget fält för booleska sökningar som alternativ. Fälten kunde fyllas i valfri ordning för att ge användaren en känsla av kontroll.

Hjälpen ”poppar upp” som separata små rutor med hjälp av Javascript och finns på skärmbilden samtidigt som sökrutan/sökrutorna vilket ger låg belastning på användarens korttidsminne enligt Allwood. Den mest grundläggande hjälpen, där utseendet på booleska operatörer visas ligger direkt på skärmen bredvid det booleska sökfältet. Den enkla varianten av gränssnittet fick mer plats över på skärmen och ”grundläggande tips och hjälp” kunde läggas upp direkt på förstasidan. Eftersom jag bedömde Altavistas hjälptexter som bra och tydliga, med neutral ton och konkreta exempel, som Allwood och Galitz rekommenderar, kopierade jag dem.

Eftersom testpersonerna förmodligen använder sig av sökmotorer som Altavista som ursprungsdomän och gränssnitten i experimentet som tillämpningsdomän, var det lämpligt att konventionerna efterföljdes, enligt Allwoods teorier. Gränssnitt 3 är därför mycket likt Altavistas original som kan vara välbekant för vissa användare, med knappar och länkar på samma ställe. Majoriteten av hjälplänkarna ligger längst upp till höger, vilket är vanligt i bland annat sökmotorer och webbutiker. Lifs ord om olika kognitiva nivåer gav mig tanken att detta skulle låta användarens medvetna kognitiva nivå sysselsättas med formuleringen av söksträngar och inte att förstå gränssnittets utformning.

Shneiderman betonar flera gånger att tydligheten för användaren är viktig, målet var därför att det skulle var enkelt för användaren att förstå vilka möjligheter som fanns och vad de innebar för sökresultatet. Flera fält erbjöds med tydliga anvisningar, vilket gjorde att användaren inte behövde känna till booleska operatörer, fräsökning med mera för att använda systemservicen maximalt. Samtliga sökfält är långa för att uppmuntra till ett stort antal termer och operatörer. De fyra översta sökfältens storlek sattes till 70, fältet för boolesk sökning anges på ett annorlunda sätt men var lite förenklat 53 bred och 6 rader hög. I den enkla varianten av gränssnitt 3 var sökfältets storlek satt till 90.

För att underlätta navigering enligt Nielsen är länkarna traditionellt blå och understrukna, de heter ”grundläggande webbsökningstips” och liknande logiska namn. De är också försedda med TITLE-taggar. För att tydligheten ska bestå även efter att användaren klickar på länken har hjälprutorna en titel som korresponderar med länknamnet.

Galitz teorier om färger ledde till att antalet färger minimerades till en dekorativ färg, ytterligare två till logon, och texten är svart mot vit bakgrund för maximal kontrast, vilket gör att webbsidan fungerar också i svartvitt. Samma typsnitt, det webbanpassade ”Verdana” användes genomgående, utan åtskillnad i stil och vikt. Två storlekar användes, en större för de viktigaste instruktionerna och en mindre för hjälpen för att undvika överdrivet scrollande. Hjälpen visades med gröna tabellfärger för att användaren skulle få en känsla av ”okej” och ”du är på rätt väg”.

Efter Nielsens idéer upptar huvudfunktionen sökning det mesta av skärmbilden i det avancerade gränssnittet, och i det enkla dominerar den enda sökkrutan tillsammans med den grundläggande hjälpen. Den enda estetiska eftergiften är färgen på tabellen, och den lilla logotypen i gif-format, något som går snabbt att ladda hem även med modem.

Svaren på Heads frågor bör vara övervägande positiva: navigationen är överskådlig och ingen hjälp ligger mer än ett klick bort, layouten är avskalad och gör det lätt för användaren att fokusera.

”Mellangränssnittet”

”Mellangränssnittet” är följaktligen gränssnitt 2. Det är ett försök till kompromiss mellan gränssnitt 1 och 3, och lånar av båda ytterligheterna.

Hjälpfunktionen ligger två klick bort, efter ett klick på ”Hjälp” kommer användaren till en sida där han/hon får välja typ av hjälp, och sedan kommer själva hjälpen som presenteras på ny skärmbild. Den är dock identisk i text med den ”bra” hjälpen i gränssnitt 3.

Länkarna är i grått och rosa, och har inga understrykningar eller TITLE-taggar. De är däremot relativt tydligt beskrivna.

Gränssnittet är avskalat och likt den enkla varianten av gränssnitt 3, men med en samling irrelevanta länkar nedanför sökkrutan.

Sökfältets storlek sattes till 40.

5.2.3 SÖKUPPDRAGEN OCH ENKÄTFRÅGORNA

Sökuppdragen konstruerades så att användarna skulle klara sig med enbart informationen som gavs i dem eftersom det annars kunde bli en fråga om allmänbildning. Ämnena var medvetet enkla och lättillgängliga, och rörde ämnen som pop, litteratur, resor och politik.¹¹⁷

Jag har utgått från Hansagi och Allebeck i mitt arbete med enkäternas utformning.¹¹⁸ Jag var intresserad av hur försökspersonerna upplevde gränssnittet, hur de använde det, om de använde sig av hjälpfunktionerna och hur testpersonerna själva upplevde sin precision och tid. Jag vill också veta hur de upplevde sökuppdragen, eftersom dåliga och svårförståeliga sökuppdrag kunde påverka undersökningen negativt.

Öppna frågor är lätta att ställa, men svåra att besvara och analysera¹¹⁹, därför har jag starkt begränsat antalet öppna frågor, till endast en ruta med ”övriga kommentarer”. Tonvikten ligger på slutna frågor med fasta svarsalternativ, men den öppna frågan krävdes också för att få mer specifik information om vad som påverkat sökprestationen och tillfredsställelsen.¹²⁰

¹¹⁷ Sökuppdragen går att läsa i bilaga 2.

¹¹⁸ Jag är medveten om att deras bok främst behandlar hälso- och sjukvård, men den är ändå så allmänt formulerad att den lägger en grund för alla former av enkäter.

¹¹⁹ Hansagi, Helen & Allebeck, Peter, 1994. *Enkät och intervju inom hälso- och sjukvård: handbok för forskning och utvecklingsarbete*. s. 38-43.

¹²⁰ Det går att läsa enkätfrågorna och se hur enkäten såg ut i bilaga 3.

Likert-skalan användes för de slutna frågorna. Detta innebär att användarna fick instämna eller ta avstånd från positiva påståenden om gränssnittet. Svaren graderades sedan med siffror 1-5 och det gick att mäta hur stark eller svag inställning testpersonerna har.¹²¹ I den tidigare nämnda artikeln av Byeong-Min Yu och Seak-Zoon Roh används också en Likert Skala, vilket gör att skalan förefaller extra lämplig eftersom denna uppsats har paralleller till deras undersökning.¹²²

5.2.4 TESTPERSONERNA

I studien har valde jag att arbeta med studenter på BHS, av flera anledningar:

- Enligt Dumas och Redish bör tester utföras på verkliga användare.¹²³ Verkliga användare till en sökmotor på webben är dock svåra att identifiera då de kan ha vilka egenskaper som helst. BHS-studenter är dock i hög grad verkliga användare till en sådan tjänst.
- Gruppen bestod av personer med liknande förutsättningar. Då det inte fanns möjlighet att göra en omfattande studie med alla typer av användare, gav denna metod med samma typer av användare åtminstone en möjlighet att kunna generalisera över skillnaderna i gränssnittet. Datavanan hos de olika användarna var ungefär lika stor (ingen är nybörjare, ingen är expert).
- BHS-studenter befinner sig också på en nivå mellan ”vanliga” oerfarna användare och informationsexperter. De har ofta studerat databaser, om än olika mycket, och använt dem, men har ännu ingen lång arbetslivserfarenhet från området.
- Eftersom jag i min framställning av gränssnitt arbetade teori-baserat, vilket enligt Löwgren är ”gammeldags”, och inte hade någon möjlighet att utarbeta gränssnitt i samarbete med mina användare, skulle jag inom tidsramen för denna uppsats haft svårt att hinna undersöka en användargrupp jag inte kände till sen tidigare.
- Av rent praktiska skäl var det också lämpligt att arbeta med studenter som kunde utföra sina sökkupdrag på Borås högskola, eftersom det då gick att använda en gemensam datasal, som förenklar utförandet av undersökningen. Mer om detta i kapitel 5.2.6.

Urvalet gick till så att jag beskrev experimentet i korthet för studenter jag kom i kontakt med och bad dem att delta i experimentet. Bortfallet blev stort på grund av att många var bortresta dagen för undersökningen, vissa var upptagna med annat, jag hade ingen möjlighet att ersätta studenter bosatta i Göteborg för resan, andra var stressade över sina egna uppsatsarbeten och ansåg sig inte ha tid att närvara. Naturligtvis hade ett helt slumpmässigt urval varit bättre men det var inte praktiskt möjligt, bortfallet hade då också förmodligen blivit än större. När det gäller datavana och intresse av webbsökning finns det dock inget som tyder på att urvalet på något sätt var riktat. Alla fyra kollegium på BHS var representerade i undersökningen. Tre testpersoner från kollegium 1, sju testpersoner från kollegium 2, tre testpersoner från kollegium 3 och två testpersoner från

¹²¹ Hansagi, Helen & Allebeck, Peter, 1994. *Enkät och intervju inom hälso- och sjukvård: handbok för forskning och utvecklingsarbete*. s. 53f.

¹²² Yu, Byeong-Min & Roh, Seak-Zoon, 2002. The Effects of Menu Design on Information-Seeking Performance and User's Attitude on the World Wide Web. S. 930

¹²³ Dumas, Joseph S & Redish, Janice C, 1994. *A practical guide to usability testing*. s. 22ff.

kollegium 4. Det var tre män och tolv kvinnor med en ålder som varierade mellan 22 och 35. 14 testpersoner läste på D-nivå och en testperson läste på B-nivå.

Testpersonerna fick själva välja var de ville sitta i datasalen. Datasalen hade tre kolumner och dessa tilldelades varsitt gränssnitt. På detta sätt blev fördelningen av användare på gränssnitten slumpmässig.

5.2.5 ANTALET ANVÄNDARE

Jag har valt att arbeta med endast 5 användare per gränssnitt, något som delvis har den praktiska förklaringen att tidsramen inte tillåter fler och att det var svårt att få studenter att ställa upp, men som också har en mer vetenskaplig grund i Nielsens uttalande om att 3-5 användare ger omfattande information om ett gränssnitt vid användbarhetstester. Nielsen förklarar att han vid en tidigare studie kommit fram till att redan med en testperson finner man en tredjedel av användbarhetsproblemen, och med 5 personer har man funnit de allra flesta, ca 85%. Efter detta ser man ungefär samma beteendemönster återkomma upp till den 15:e testpersonen då man har funnit 100% av användbarhetsproblemen. Eftersom 5 personer ändå räcker för att ge en tydlig bild av ett gränssnitt och dess problemområden rekommenderar Nielsen att hellre ta tre grupper om fem än att satsa all tid på 15 personer i en grupp.¹²⁴ Nielsen skriver visserligen att han för kvantitativa utvärderingar rekommenderar 20 användare per gränssnitt¹²⁵, men det är något som tyvärr inte är möjligt i en uppsats av den här storleken.

Nielsen skriver om användbarhetstester och har en kommersiell infallsvinkel, han räknar också med att testerna ska bli så kostnadseffektiva som möjligt eftersom hans tester kostar pengar i form av betald arbetskraft.¹²⁶ Uppsatsens förhållanden är annorlunda, men trots det tillräckligt lika för att hans teori om fem användare ska vara tillämpliga.

5.2.6 TESTMILJÖ

En datasal på högskolan användes för utförandet av experimentet. Detta underlättade insamlingen av testpersonernas relevansbedömning, tidtagningen, söksträngslängden och enkätfrågorna. Det förenklade också arbetet med bearbetningen av resultatet. Eftersom testet utfördes schemalagt under en enda dag, minskade också risken för att testpersonerna inte skulle delta eller lämna in sina uppgifter för sent. En mycket viktig aspekt är också den föränderliga webben där nya webbsidor tillkommer och försvinner varje minut. Eftersom testpersonerna utför sina sökningar samtidigt minskar risken att samma söksträng ger olika träfflistor, något som förmodligen skulle inträffa om sökningarna gjordes med flera dagars mellanrum.

Lif beskriver skillnaden mellan att testa i laboratoriemiljö och fältstudier där testpersonerna vistas i sin naturliga arbetsmiljö. I mitt fall var datasalen något av ett mellanting av laboratorium och normal arbetsmiljö.

Fördelarna med laboratoriemiljön är att mätning av resultat förenklas (som i mitt fall), och att forskaren (i detta fall jag) har full kontroll över testsituationen och dess variabler. Upplösning på skärm, uppkopplingshastighet, datorprestanda osv inverkar

¹²⁴ Nielsen, Jacob, 2000. *Test With 5 Users (Alertbox Mar. 2000)*.

¹²⁵ Nielsen, Jakob, 2001. *Usability Metrics (Alertbox Jan. 2001)*.

¹²⁶ Nielsen, Jacob, 2000. *Test With 5 Users (Alertbox Mar. 2000)*.

inte på resultatet eftersom testpersonerna använde samma typ av datorer med samma upplösning, uppkoppling och prestanda. Gränssnitten öppnades i upplösning 800x600 och webbläsaren var Internet Explorer 5.5. Det gick också att utesluta störande element som kunde inverka på testet. Om denna uppsats undersökning skulle utförts i högskolans datasalar när de inte var bokade för detta syfte, eller i hemmet, skulle användarna kunnat bli störda av kurskamrater, vänner och familjemedlemmar.

Nackdelarna med en laboratoriemiljö är att användaren upplever miljön som onaturlig och kan ha svårt att uppträda naturligt och spontant.¹²⁷ Min förhoppning är att de flesta BHS-studenter utför delar av sitt skolarbete i datasalarna, och att det därför var en bekant miljö.

Altavista valdes som databas eftersom Altavista har en sökmotor som ger många möjligheter till precisa sökningar med en stor mängd olika operatorer. Google som kanske är populärare har snarare specialiserat sig på att ge bra träfflistor trots korta söksträngar och har också tagit bort möjligheten till många booleska operatorer. Det är en databas och en systemservice som fokuserat på amatöranvändare och hade förmodligen givit bra träfflistor trots dåliga gränssnitt och hjälpfunktioner vilket hade gjort gränssnitten svåra att jämföra.

Miljön på skärmen bestod av en startsida som var länkad till en webbsida som med hjälp av ASP¹²⁸ hämtade upp sökuppdragen från en databas. Från denna sökuppdrags-sida öppnades ett nytt fönster med själva gränssnittet, HTML-filer med inslag av Javascript. Relevansbedömning utfördes på samma webbsida som sökuppdraget genom att användarna fick kryssa i fyrkantiga boxar i den ordning och antal som relevanta träffar visades på Altavistas träffsida. Det totala antalet träffar skulle också fyllas i för att ge ett korrekt värde för precision även vid färre träffar än tio. Instruktioner till detta gavs i tryckt form.¹²⁹ Tiden mättes från det att användaren öppnade gränssnittet från sökuppdrags-sidan, till det att användaren tryckte på "sök" (eller motsvarande) i det givna gränssnittet. Instruktionen var därför att sökuppdraget skulle läsas först efter att gränssnittet öppnats eftersom tiden som skulle mätas var konstruktionen och formuleringen av söksträngen, inte bara tiden för inskrivning av söksträng i sökfält. Efter sökuppdragen leddes testpersonerna vidare till en webbenkät, där de fick instämna i eller ta avstånd från nio påståenden genom att fylla i radioknappar. Det fanns också ett valfritt fält för kommentarer som flertalet deltagare fyllde i.

Alla data: IP-nummer, nummer på sökfråga, tid för söksträngsformulering, tid på dygnet, söksträng, nummer på gränssnitt, relevansbedömning, totala antalet träffar, enkätsvar och kommentarer lagrades direkt i den databas som ligger till grund för denna uppsats resultatredovisning.

¹²⁷ Lif, Magnus, 1998. *Adding Usability: Methods for Modelling, User Interface Design and Evaluation*. s. 26f.

¹²⁸ ASP är en programmeringsteknik som gör webbsidor dynamiska.

¹²⁹ Se bilaga 2.

5.2.7 STATISTISKA FÖRKLARINGAR

För att ta reda på om det finns signifikanta skillnader i de olika gruppernas resultat har jag använt mig av Mann-Whitney testet. Det är ett test som mäter hur stora skillnaderna mellan olika gruppers resultat är och vad den statistiska sannolikheten är för att dessa skillnader har uppstått av slumpen.

För en förklaring och motivering av de statistiska metoder som används i denna uppsats har böcker av Staffan Stukát, Robert R Pagano, Svante Körner och Lars Wahlgren använts. För att förstå statistik är en av grunderna att veta skillnaden mellan olika skalor. En *kvotskala* är en skala med lika långa skalsteg, det vill säga avståndet mellan 0,4 och 0,5 är lika långt som det mellan 0,2 och 0,3. Det ska också finnas en absolut nollpunkt.¹³⁰ Detta gäller för precision, tid och söksträng, även om scenariot att tid och söksträng i något verkligt fall skulle vara noll är osannolikt.

En *ordinalskala* har inte lika långa skalsteg. Däremot går det att rangordna värdena och det är meningsfullt att prata om "mer än" och "mindre än".¹³¹ I Likertskalan som användes för utvärderingen av den subjektiva tillfredsställelsen är det svårt att veta om avståndet mellan "tar helt avstånd" och "tar delvis avstånd" är lika stort som avståndet mellan "instämmer delvis" och "instämmer helt", men en rangordning av omdömena är möjlig och relevant.

Valet av statistisk metod är beroende av vilken skala som använts vid mätningen. För en kvotskala kan typvärde, aritmetiskt medelvärde eller medianvärde¹³² användas, för en ordinalskala ska typvärde eller medianvärde användas.¹³³

För att analysera resultaten har jag gjort en *hypotesprövning*, eller ett *signifikanstest* med ett annat ord. Jag har undersökt skillnaderna mellan de tre gruppernas resultat och har tagit reda på om skillnaderna är tillräckligt stora för att inte kunna förklaras med slumpen. Jag har beräknat *nollhypotesens*, hypotesen att det inte är någon skillnad mellan gränssnitten, sannolikhetsvärde. Ju mindre en sådan sannolikhet är, desto mer troligt är det att skillnaden är *signifikant* eller *statistiskt säkerställd* som det heter på statistiskt språk. Gränsen här är dock svår att dra. Vanligen anges signifikansnivån (som betecknas α) till 5% eller 1%. Signifikansnivån anger hur stor sannolikheten för att skillnaden är slumpmässig får vara. Om det är 6% sannolikhet att två olika skolklasser får 30 respektive 35 poäng i genomsnitt på ett prov är en signifikansnivå på 5% inte nådd. Det är då troligt att slumpen ligger bakom de skiftande resultaten och nollhypotesen får accepteras. Sannolikhetsvärdet kallas för p-värde (prob-value) och anges i decimaler, en sannolikhet på 5% motsvaras av p-värde 0,05.¹³⁴

Som exempel anges nu två olika hypoteser:

¹³⁰ Stukát, Staffan, 1993. *Statistikens grunder*. s. 29.

¹³¹ *Ibid.*, s. 30.

¹³² Typvärde är en enkel mätning av vilket värde som förekommer flest gånger. Aritmetiskt medelvärde är ett "vanligt" medelvärde, där summan av värdena delas med antalet personer eller motsvarande.

Medianvärde är en mätning av vilket värde som befinner sig precis "i mitten" av ett antal värden om de arrangeras i ordningsföljd. Ligger två värden i mitten är medianvärdet medelvärdet av dessa två.

¹³³ Stukát, Staffan, 1993. *Statistikens grunder*. s. 30.

¹³⁴ Stukát, Staffan, 1993. *Statistikens grunder*. s. 71ff och Körner, Svante & Wahlgren, Lars, 1998. *Statistiska metoder*. s. 122ff

H_0 = "Nollhypotesen." Det finns ingen skillnad mellan gränssnitten. De skillnader som finns beror på slumpen.

H_1 = "Mothypotesen." Det finns en verklig, signifikant skillnad mellan resultaten i de olika gränssnitten.

Det går aldrig att veta helt säkert om en fördelning är slumpmässig eller inte. Med en högre signifikansnivå ökar risken för att H_0 förkastas trots att den egentligen är sann, med en lägre signifikansnivå ökar risken för att H_1 inte förkastas trots att den är falsk. Signifikansnivån bestämmer vilken felrisk forskaren tar.¹³⁵ Det är alltså alltid nollhypotesen som prövas och accepteras eller förkastas. Det har införts en skala med stjärnor för att beteckna olika grader av signifikans, ett p-värde på 0,001 är trestjärnig signifikans, p-värde 0,01 är tvåstjärnig och p-värde 0,05 är enstjärnig. P-värde 0,05 är dock standardvärdet och det kommer också att användas här ($\alpha = 5\%$).¹³⁶

En hypotes kan vara av typen "a är signifikant större än b" ($a > b$) och mätningen görs då ensidigt och riktat, nollhypotesen är ju då att "a förhåller sig slumpmässigt till b" och "b är signifikant större än a". En tvåsidig hypotes däremot är "a är signifikant olik b" ($a \neq b$). Nollhypotesen är då bara "a förhåller sig slumpmässigt till b" och p-värdet blir dubbelt så stort eftersom den tvåsidiga mothypotesen då består av både "a är signifikant större än b" ($a > b$) och "b är signifikant större än a" ($b > a$).¹³⁷ I uträkningarna i denna uppsats är hypoteserna uteslutande ensidiga.

Ett vanligt statistiskt begrepp är *normalfördelning*. Fysiska, lätt mätbara egenskaper som människors längd är normalfördelade. Detta innebär att värdena är symmetriskt fördelade kring medelvärdet i en kurva där halvorna är spegelbilder av varandra. Det finns lika många människor som är kortare än medellängden som det finns de som är längre.¹³⁸ Om värdena för precision, tid och söksträngar var fördelade på det här sättet hade jag kunnat använda ett ANOVA-test (ANalysis Of VAriance), som användes i artikeln av Byeong-Min Yu och Seak-Zoon Roh¹³⁹, som är en metod att jämföra fler grupper än två på en gång. Detta är nu inte fallet och jag har i stället valt ett icke-parametriskt parvis test, nämligen Mann-Whitney. Det finns ingen anledning att gå in på de matematiska formlerna för detta test, men en kort genomgång av tanken bakom är relevant för att förstå beräkningarna och resultaten i kapitel 7.

Först kommer dock en kort presentation av begreppet variabler. En variabel är en mätbar egenskap eller ett objekt som kan anta olika värden. En oberoende variabel i ett experiment är den variabel som forskaren kontrollerar, i detta fall är den oberoende variabeln gränssnitten. Den beroende variabeln i ett experiment är den variabel forskaren mäter för att avgöra den oberoende variabelns effekt, i detta fall alltså precision, tid, söksträngslängd och tillfredsställelse.¹⁴⁰

Mann-Whitney testet är konstruerat för att jämföra två oberoende grupper med olika personer, i motsats till andra test där två grupper av resultat kan härstamma från samma

¹³⁵ Stukát, Staffan, 1993. *Statistikens grunder*. s. 74.

¹³⁶ Körner, Svante & Wahlgren, Lars, 1998. *Statistiska metoder*. s. 124ff

¹³⁷ Ibid., s. 127ff

¹³⁸ Stukát, Staffan, 1993. *Statistikens grunder*. s. 58f

¹³⁹ Yu, Byeong-Min & Roh, Seak-Zoon (2002). The Effects of Menu Design on Information-Seeking Performance and User's Attitude on the World Wide Web. s. 929.

¹⁴⁰ Pagano, Robert R, 1998. *Understanding Statistics in the Behavioral sciences*. s. 6.

personer, till exempel vid utvärdering av ”före och efter” en medicinsk behandling.¹⁴¹ Testet kräver inte heller normalfördelning eller kvotskala vilket gjorde det lämpligt för denna undersökning.¹⁴²

Mann-Whitney testet analyserar separationen mellan två grupper av värden: Om två sådana gruppers värden är lika skulle deras kurvor, om de ritades ut, överlappa varandra helt. Om det däremot är så att flera eller alla värden i den ena gruppen överstiger värdena i den andra, får man två kurvor som bara överlappar varandra till viss del. Ju mer separerade dessa kurvor är desto mer osannolikt är det att enbart slumpen är ansvarig för skillnaderna.¹⁴³

Jag provade de olika värdena parvis med detta test för att undersöka hur resultaten skiljer sig åt. Medelvärden och medianvärden för de olika sökuppdragen och enkätens påståenden jämfördes, grupperna emellan. Tre gränssnitt (oberoende variabler) ger att det var tre parkonstellationer (av beroende variabler) som kunde undersökas för precision, tid, söksträngslängd och tillfredsställelse.

Beräkningarna utfördes med *Analyse-it*, ett tillägsprogram till Excel.

¹⁴¹ Pagano, Robert R, 1998. *Understanding Statistics in the Behavioral sciences*. s. 241.

¹⁴² Lowry, Richard, 1998-2003. *Concepts and Applications of Inferential Statistics*. kap. 11a

¹⁴³ Pagano, Robert R, 1998. *Understanding Statistics in the Behavioral sciences*. s. 243f

6. REDOVISNING AV RESULTAT

På grund av tekniska missöden och missförstånd av instruktionerna utfördes vissa sökuppdrag flera gånger. Bland annat glömde vissa testpersoner att göra relevansbedömning och fick gå tillbaka och upprepa sin söksträng för att kunna göra den. Det hände också flera testpersoner att webbläsaren "kraschade" och vid ett tillfälle också själva datorn, de fick då göra om den sökning de var på när skadan inträffade. Vissa testpersoner verkar också ha missuppfattat instruktionerna och har gjort flera sökningar per sökuppdrag, främst till de första sökuppdragen. Trots att uppsatsens premisser kräver att det vid dessa fall ska vara den första sökningen som sparas har detta inte varit praktiskt möjligt. Tiden och relevansbedömning för varje sökuppdrag är i databasen matchade med den sista sökningen på varje sökuppdrag, och det har därför varit nödvändigt att kasta de poster som föregick den sista sökningen.

De upprepade sökningarna kan ha påverkat resultatet för undersökningen i olika riktningar. Vid utrensandet av de överflödiga posterna såg jag dock ingen tendens till övervikt för ett viss gränssnitt. Det vill säga, problemen med datorerna och missförstånden av instruktionerna verkar inte ha koncentrerat sig till en viss grupp utan låg relativt jämnt utspridda.

Av okänd teknisk anledning kom en testpersons relevansbedömning och tid för söksträngsformulering att bli identiska för de första tio sökuppdragen. Dessa resultat har också varit oanvändbara i undersökningen, när det gäller tid och precision, varför det i redovisningen av precision och tid kommer att förefalla som att det bara var fyra personer som utförde sökuppdrag i gränssnitt 2.

En nackdel med den här typen av undersökning är att omständigheterna kring sökningarna kan ha krävt för mycket av testpersonernas uppmärksamhet och kognitiva förmågor för att ge ett helt naturligt resultat. Som tidigare har diskuterats i teoridelen är växlandet mellan fönster inte att rekommendera, och det bästa hade förmodligen varit att lämna sökuppdragen i tryckt form. Detta var dock inte praktiskt möjligt på grund av tidmätningen. Flera deltagare påpekade efter undersökningen att situationen ändå kändes naturligt och att det var lätt att följa instruktionerna.

Det blev också nödvändigt att berätta om tidmätningen för att testpersonerna inte skulle glömma att växla snabbt mellan fönstren. Trots att jag betonade testpersonernas anonymitet och mitt intresse av det naturliga resultatet snarare än ett snabbt kan detta ha stressat vissa testpersoner.

Under experimentet fick några testpersoner problem som de bad om hjälp om: "backa-funktionen" försvann på vissa webbsidor de besökte, själva webbläsaren och datorn "kraschade" som beskrivet ovan, och det fanns också frågeställningar kring relevansbedömningen. Sammantaget inträffade dock inga stora problem och testpersonerna verkade väl införstådda med instruktionerna.

För att underlätta resultatredovisningen i uppsatsen och i bilagan har de olika testpersonerna givits namn som är relaterade till deras gränssnitt. Person G1:1 är den första testpersonen i gränssnitt 1, person G2:1 är första personen i gränssnitt 2, och G1:3 är tredje personen i gränssnitt 1, och så vidare. För varje grupp har också skapats en fiktiv medelperson som kallas G1, G2, och G3, för att enkelt redovisa medelvärden och

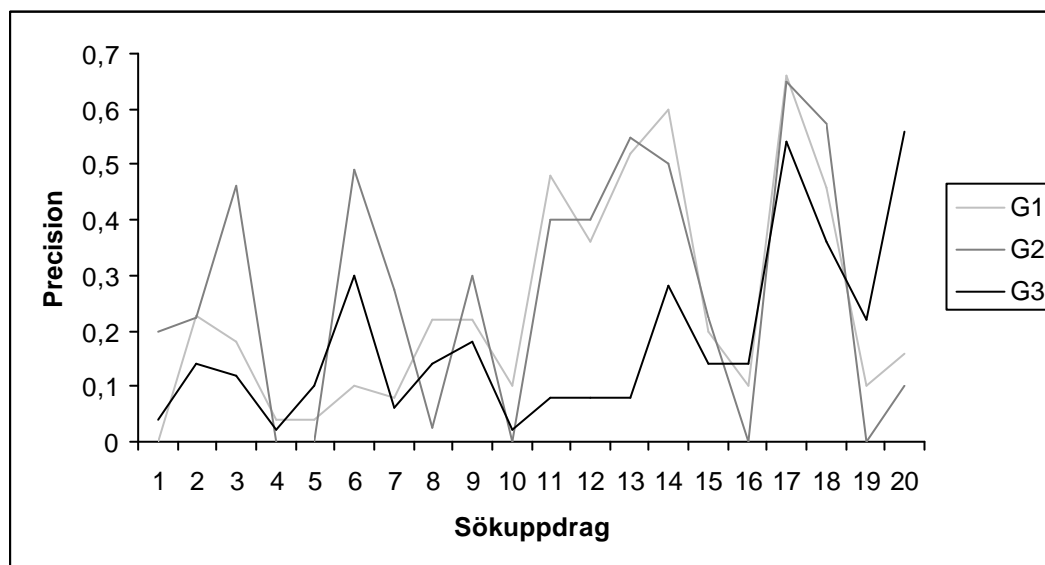
medianvärden för varje sökuppdrag och enkätsvar i de olika gränssnitten. Detta har då beräknats med aritmetiskt medelvärde för precision, tid och söksträngslängd, och medianvärde för tillfredsställelsen enligt de statistiska regler som beskrivs i 5.2.7. Eftersom Mann-Whitney testets beräkningar bygger på dessa medelvärden och medianvärden är det dessa som redovisas direkt i uppsatsen. Den fullständiga resultatredovisningen med de enskilda testpersonernas resultat återfinns i bilaga 5.

6.1 PRECISION

Precision har räknats ut så att antalet relevanta träffar har delats med tio, utom i de fall där det totala antalet träffar understigit tio, då det är den siffran som varit nämnare. Till exempel: två relevanta träffar med ett totalt antal träffar på 1162 ger precision 0,2, men en relevant träff på totalt fem träffar är också precision 0,2. Precision blir då alltid ett tal mellan 0 och 1.

G2:3 var den testperson vars resultat inte blivit korrekt införda och därför räknas inte dessa resultat med i medelvärdet.

För en översikt av resultaten för precision presenteras här de fiktiva medelpersonerna i ett diagram:



Figur 1: Diagram över medelvärden för precision

Det är tydligt att gränssnitt 3 inte har gett bäst precision, detta kommer att diskuteras i kapitel 7.

6.2 TID

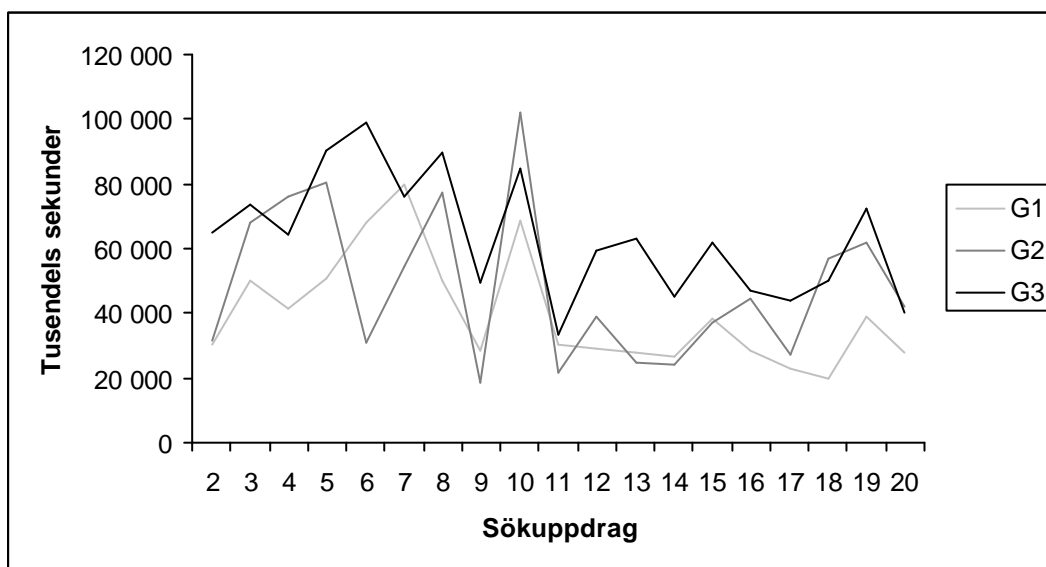
Tiden som mätts är tiden från att testpersonen börjar läsa sökuppdraget till det att testpersonen formulerat en söksträng och trycker på "sök" i det tilldelade gränssnittet förutsatt att testpersonerna följt instruktionerna. Måttenheten är tusendels sekunder.

Tiden är inte helt oproblematiserad i denna uppsats. Om precisionen hade varit bäst i gränssnitt 3 hade det varit intressant att veta om det också gick snabbare och var mer lättarbetat. En annan teori var att ett dåligt gränssnitt som gränssnitt 1 skulle vara snabbt

att arbeta med eftersom användarna kanske helt enkelt inte hittade hjälpfunktionerna och bara gissade sig fram i snabb takt med ett dåligt resultat som följd.

Trots denna problematik bedömde jag att tiden var en tillräckligt intressant aspekt för att mätas och resultatet redovisas nedan på samma sätt som i föregående avsnitt om precision.

Första sökuppdraget var ett extremfall. Eftersom det första sökuppdraget också var den första kontakten med själva experimentet uppstod en del frågor och problem kring detta. Detta ledde till att detta sökuppdrag blev extremt tidskrävande för en del testpersoner. Ett antal testpersoner fick göra om första sökuppdraget, och då det i sådana fall blev den andra sökningen som sparades i databasen fick dessa testpersoner ovanligt korta tider i stället. De fullständiga tabellerna i Bilaga 5 ger en tydlig bild av problemet. Första sökuppdragets extremvärden kan alltså påverka det aritmetiska medelvärdet på ett missvisande sätt och redovisas därför inte i diagrammet som följer:

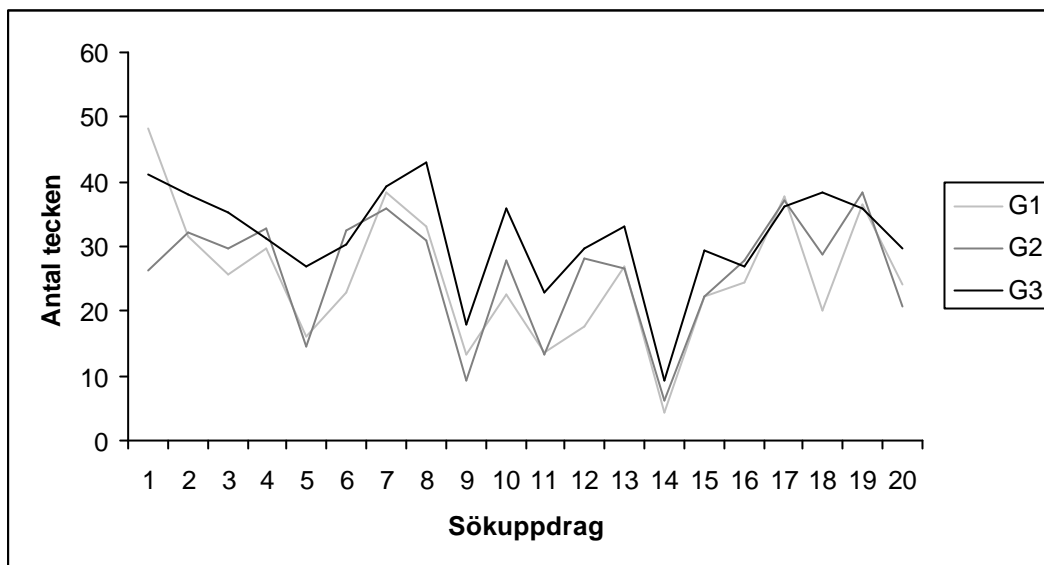


Figur 2: Diagram över medelvärden för tid

6.3 SÖKSTRÄNGENS LÄNGD

I beskrivningen av de olika gränssnitten i kapitel 5.2.2 beskrevs hur stora sökfälten gjordes. Extremvärdena när det gäller denna storlek var 20 respektive 90 vilket visar på att skillnaderna var stora. I gränssnitt 3 fanns det också fem olika fält att söka i vilket gav en sammanlagd längd som vida översteg den lilla rutan i gränssnitt 1. Enligt Shneiderman gör användare längre söksträngar om sökfältet är stort. Här redovisas resultatet av mina mätningar. Värdena är angivna som antal tecken inklusive mellanslag och operatorer.

Nedan visas medelvärdena för de olika sökuppdragen:



Figur 3: Diagram över medelvärden för söksträngslängd

6.4 TILLFREDSSTÄLLELSE

Mätningen av personlig tillfredsställelse bestod av nio påståenden som testpersonerna fick instämma med i olika grad. Påståendena går att läsa i bilaga 3.

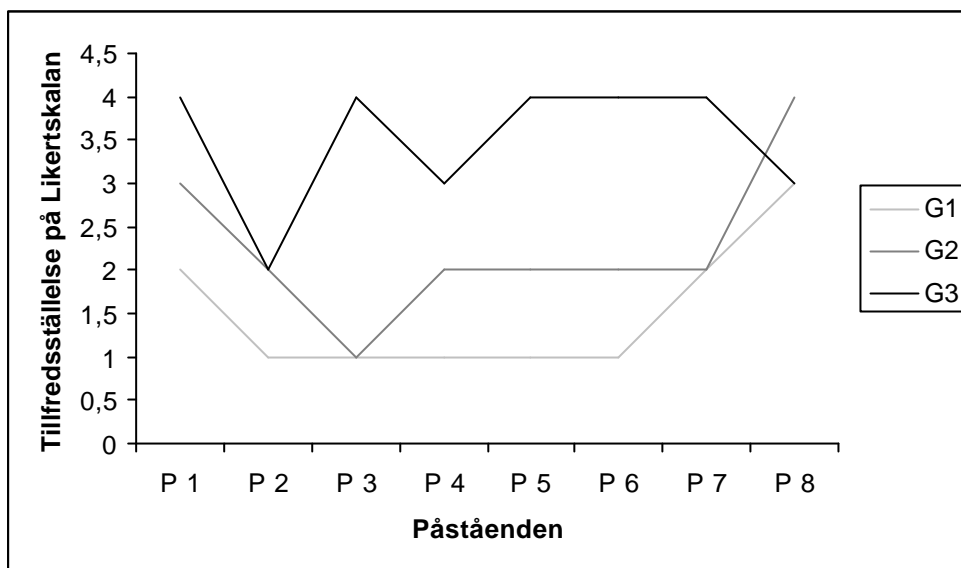
Alternativen var:

- tar helt avstånd
- tar delvis avstånd
- tar varken avstånd eller instämmer
- instämmer delvis
- instämmer helt

Dessa fem alternativ har graderats från 1 till 5, där 5 är det starkaste instämmandet, för redovisningen av resultatet. Påstående 9: ” Jag tyckte att sökuppdragen var tydliga och lätta att förstå” syftade till att utvärdera experimentet snarare än gränssnittet. Det var också ett påstående som inte påverkades av vilket gränssnitt testpersonerna blivit tilldelade. Det behövdes för att undersöka om lågt antal relevanta träffar kan ha orsakats av dålig förståelse av sökfrågorna. Medianvärdet på detta påstående blev 5 med lägsta instämmande 3, vilket tyder på att sökfrågorna i sig inte påverkat resultaten i undersökningen. Påstående 9 kommer därför fortsättningsvis inte redovisas.

Påståendena förkortas nedan med ”P”, P 1 är alltså påstående 1. Eftersom Likertskaalen är en ordinalskala har medianvärdet angivits för de fiktiva testpersonerna G1, G2 och G3.

Nedan presenteras dessa medianvärden i ett diagram:



Figur 4: Diagram över medianvärden för tillfredsställelse

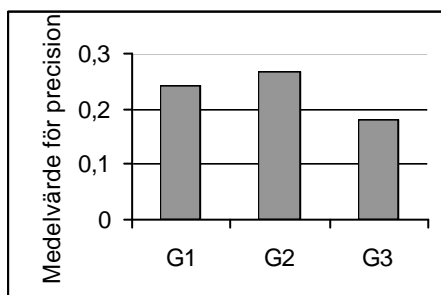
7. ANALYS OCH DISKUSSION

Det är nu dags för den statistiska analysen av resultaten och eftersom en uppdelning av denna analys och de diskuterande inlag som uppträder i samband med den vore otymplig har analys och diskussion kombinerats i denna uppsats.

Min hypotes i denna uppsats är att gränssnitt 3 kommer att ge högre precision, större tillfredsställelse och längre söksträngar, medan gränssnitt 1 kommer att ge ett sämre resultat och gränssnitt 2 någonstans däremellan. När det gäller tiden finns ingen hypotes utan en frågeställning om värdena skiljer sig åt på ett signifikant sätt och i vilken riktning de i så fall gör det.

7.1 PRECISION

Hypotesen var från början att gränssnitt 3 skulle prestera bättre än gränssnitt 2 som i sin tur skulle prestera bättre än gränssnitt 1. För att visa de faktiska skillnaderna kommer här ett litet diagram som visar medelvärdet för precision:



Figur 5: Ett förenklat diagram över medelvärden för precision.

Detta diagram visar att gränssnitt 2 presterat bäst, sedan gränssnitt 1, medan gränssnitt 3 har presterat sämst. Jag har då utgått från de faktiska skillnaderna som hypoteser för att se om skillnaderna var signifikanta. Nollhypotesen för de tre paren var dock samma:

H₀ Det finns ingen skillnad mellan gränssnitten när det gäller precision. De skillnader som uppstått beror på slumpen.

H_{1 nr 1} Gränssnitt 2 ger bättre precision än gränssnitt 1. (Gränssnitt 2 > gränssnitt 1.)

P-värdet blir för denna uträkning **0,4036**.

H_{1 nr 2} Gränssnitt 2 ger bättre precision än gränssnitt 3. (Gränssnitt 2 > gränssnitt 3.)

P-värdet blir för denna uträkning **0,1784**

H_{1 nr 3} Gränssnitt 1 ger bättre precision än gränssnitt 3. (Gränssnitt 1 > gränssnitt 3.)

P-värdet blir för denna uträkning **0,1714**.

Det är tydligt att inga av dessa värden befinner sig i närheten av signifikansnivån. Spridningen är förmodligen slumpartad och H_0 får antas i samtliga tre fall. Varför blev det så här? Det finns flera tänkbara förklaringar:

- En eller flera av teorierna som gränssnitten är byggda på är felaktiga.
- Gränssnitten borde ha provats ut enligt idéerna om iterativ design som flertalet teoretiker förespråkar. Jag är en alltför amatörmässig designer för att korrekt kunna applicera teorierna på gränssnitt.
- Gruppernas storlek var otillräcklig. Grupper om fem är i normalfallet inte tillräckligt att bygga statistik på. Om antalet testpersoner hade varit större hade resultaten skilt sig åt mer.
- De individuella skillnaderna kommer alltid att överstiga skillnaderna mellan gränssnitten. Det är större skillnad mellan en erfaren och intresserad användare och en oerfaren och ointresserad än det är mellan två gränssnitt av skiftande kvalitet. Gränssnitt är inte avgörande för sökresultat mätt i precision.

Jag ska nu behandla dessa punkter i tur och ordning.:

Det är möjligt att en eller flera teorier är felaktiga, men ingen forskare eller författare har konkret hävdats att sökresultats precision skulle påverkas av gränssnittets utformning, och därför är det svårt att säga att någon teori blivit direkt motbevisad. Hypotesen om precision och gränssnitt var högst experimentell.

När det gäller designen var det förmodligen ett misstag med gränssnitt 3 att göra flera fält. Trots tydliga anvisningar till sökfälten på samma skärmbild och en nedfällbar hjälpruta som förklarade hur fälten skulle användas försökte flera testpersoner använda citationstecken och booleska operatorer inuti de fält där de inte behövdes eller ens *fick* användas med nollresultat som följd. De olika sökfälten som skulle förenkla för ovana användare gjorde tyvärr att användarna gjorde fler misstag än normalt. Den enkla varianten av gränssnitt 3 har inte använts, något som gått att avläsa med hjälp av de sparade söksträngarna. I övrigt tyder dock enkätsvaren på att designen inte är misslyckad, utan att testpersonerna upplevt gränssnitten på det sätt som de var avsedda. Mer om detta i kapitel 7.4.

Nielsen och Large, Tedd och Hartley rekommenderade att placera ett enkelt gränssnitt som förval och erbjuda ett mer avancerat som alternativ. Jag gjorde i experimentet tvärtom eftersom jag bedömde att BHS-studenter skulle vara vana vid avancerade gränssnitt. Det förefaller dock som att det är bättre att låta användare börja med ett enkelt gränssnitt oavsett sökvana.

Eftersom p-värdet överskrider 0,05 så mycket tvivlar jag på att ett större urval hade gett signifikanta resultat. Det går dock inte att utesluta möjligheten.

Det är tydligt även vid en snabb överblick av tabellerna i bilaga 5 att de enskilda testpersonernas medelvärden för precision varierar mer än gränssnittens genomsnitt. Min förklaring är att individens bakgrund, förkunskaper, personlighet, intresse och motivation är avgörande för sökresultat. Vissa studenter kan ha fått mer träning i att söka information på webben och kanske också större insikter i hur en databas fungerar

och vad booleska kommandon kan hjälpa till med. Utan denna formella träning blir sökning svårare och intresset för att fördjupa sig i hjälpfunktioner mindre.

Det kan ju också vara så att konsten att söka framgångsrikt på webben är mer komplicerat än att ett gränssnitt kan lära ut det på ett par timmar. Det är möjligt att den här typen av undersökning, när det gäller precision, borde utföras under längre tid. Dumas och Redish beskrev hur användbarheten testas när det gäller helt nya maskiner och gränssnitt. När en ny funktion på en arbetsplats introduceras och testas, till exempel en ny slags monitor på ett sjukhus, skulle förmodligen testningen pågå under en längre tid för att se om de anställda lyckas lära sig den inom en rimlig tidsperiod. Med utgångspunkt från det resonemanget kanske det hade varit mer intressant att testa något som ingen av testpersonerna behärskade sedan tidigare.

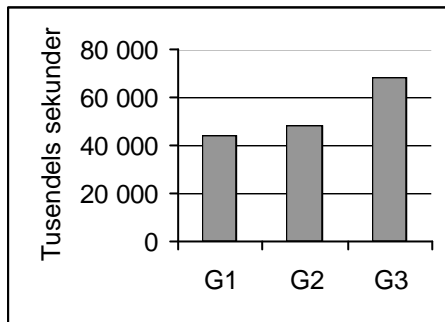
Ett rimligt antagande är att gränssnitt inte är avgörande för precisionen utan att användarnas intresse och erfarenhet, och även systemet och systemservicen, spelar en större roll. Altavista är inte den enda sökmotorn på webben och det är möjligt att testpersonerna var ovana vid just Altavistas operatorer och rankningstekniker.

Den subjektiva relevansbedömningen kan ha påverkat resultatet. Det är ju mycket troligt att olika personer kan ha bedömt samma dokument på olika sätt och den binära relevansbedömningen har gjort de ”små nyanserna” i relevansbedömningarna svåra att upptäcka. Det är dock svårt att avgöra i vilken riktning detta kan ha påverkat då det som tidigare nämnts förefaller vara större skillnader mellan individer än mellan gränssnitt. En undersökning med en testdatabas där forskaren på förhand vet vilka dokument som är relevanta skulle kunna ge svaret. Jag måste dock återigen hänvisa till att det är den subjektiva tillfredsställelsen som är viktig för verkliga användare vid verkliga sökningar. Det är möjligt att samma dokument bedömts på olika sätt, men då har förmodligen användaren också upplevt denna bedömning som riktig. En användare som uppnått precision 0,5 med sin söksträng är mer nöjd än en som har uppnått 0,0, även om de i själva verket har fått samma träffar. Precision är i denna uppsats ett sådant subjektivt mått.

Problemet med de felaktigt upprepade sökningarna kan också ha påverkat. Till största delen bestod dock dessa sökningar av identiska söksträngar eftersom de utfördes på grund av datorproblem eller att testpersonen glömt fylla i relevansbedömning. Detta gav då också identisk precision.

7.2 TID

Som tidigare nämnts var tiden en faktor som var svår att kalkylera på förhand. Idealet, att gränssnitt 3 skulle ge hög precision, långa söksträngar och en snabb tid verkade inte vara en realistisk kombination. Att gränssnitt 1 med dess dåliga hjälpfunktioner och förvirrande design skulle göra användaren stressad var troligare. Mycket riktigt är tiden för gränssnitt 1 kortare än för gränssnitt 2 som i sin tur är kortare än gränssnitt 3 som visas i diagrammet nedan:



Figur 6: Ett förenklat diagram över medelvärden för tid

Men är skillnaden signifikant? Precis som i redovisningens diagram har första sökuppdraget inte räknats med i diagrammet ovan, det är inte heller med i uträkningarna nedan.

H₀ Det finns ingen skillnad mellan gränssnitten när det gäller tid. De skillnader som uppstått beror på slumpen.

H₁ nr 1 Gränssnitt 1 ger en kortare tid per sökfråga än gränssnitt 2. (Gränssnitt 1 < gränssnitt 2.)

P-värdet blir för denna uträkning **0,1569**.

H₁ nr 2 Gränssnitt 2 ger en kortare tid per sökfråga än gränssnitt 3. (Gränssnitt 2 < gränssnitt 3.)

P-värdet blir för denna uträkning **0,0148**

H₁ nr 3 Gränssnitt 1 ger en kortare tid per sökfråga än gränssnitt 3. (Gränssnitt 1 < gränssnitt 3.)

P-värdet blir för denna uträkning **0,0003**.

För hypotes nr 1 antas H₀, men hypotes nr 2 och nr 3 ger högt signifikanta resultat med nära tvåstjärnig nivå på nr 2 och rent trestjärnig nivå på nr 3.

Här kan man tänka sig att den känsla av maktlöshet som ett dåligt gränssnitt ger enligt så många författare gjorde att testpersonerna i gränssnitt 1 gav upp och slentrianmässigt och planlöst gissade på olika söksträngar eftersom de upplevde att de inte hade kontroll över sina resultat. Det är också möjligt att de animerade bilderna hade en stressande effekt. Gränssnitt 3 däremot gav tydligare anvisningar om vilka möjligheter som fanns för sökning och kan ha utmanat testpersonerna att lägga mer tid på att formulera sina söksträngar.

Gränssnitt 3 var det gränssnitt som hade en mest uppenbar ursprungsdomän, nämligen Altravista. Gränssnitt 1 upplevdes däremot som underligt av testpersonerna, och en av dem frågade även i början av sökningen: "Är det här verkligen mitt gränssnitt?". Testpersonerna i gränssnitt 1 var förmodligen därför mer benägna att slarva igenom sökuppdragen, då gränssnittet kändes okonventionellt och förvirrande. Eftersom

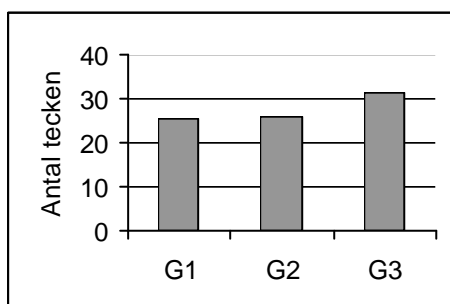
gränssnitt 3 påminde mer om ett seriöst och vanligt gränssnitt kan det ha motiverat testpersonerna att göra mer seriösa och tidskrävande sökningar. Sökrutornas storlek kan ha påverkat också här. Trots att antalet tecken var obegränsat i samtliga tre gränssnitt, har det förmodligen varit mer obekvämt att göra långa söksträngar i gränssnitt 1 och till viss del gränssnitt 2, eftersom man då inte kunde se hela söksträngen samtidigt, något som också kan ha påverkat sökningarnas tid.

Testpersonerna i gränssnitt 1 och gränssnitt 2 har inte heller använt sig av hjälpfunktionerna i någon större utsträckning eftersom de var svåra att hitta och använda¹⁴⁴. Detta har förmodligen också gett kortare söktider, då tiden testpersonerna ägnade åt hjälpfunktionerna är inkluderade i söktiden.

Något som kan ha påverkat tiden är problematiken med de upprepade sökningarna. Det första sökuppdraget har redan diskuterats och eliminerats från undersökningen. I övriga fall har det naturligtvis blivit så att de sökuppdrag som utfördes flera gånger har fått mycket korta uppmätta tider. Att detta förekom utspritt i de tre grupperna har förmodligen ändå gjort att resultaten inte påverkats i någon speciell riktning. Signifikansen är också så hög att det skulle krävas stora förändringar av värdena för att nollhypotesen skulle antas.

7.3 SÖKSTRÄNGENS LÄNGD

Hypotesen om söksträngslängd består i grunden av att söksträngen bör bli längre ju större sökfältet är. Resultaten i tabellform avslöjar att så är fallet:



Figur 7: Ett förenklat diagram över medelvärden för söksträngslängd.

Men är resultaten signifikanta?

H₀ Det finns ingen skillnad mellan gränssnitten när det gäller söksträngslängd. De skillnader som uppstått beror på slumpen.

H₁ nr 1 Gränssnitt 3 ger längre söksträngar än gränssnitt 2. (Gränssnitt 3 > gränssnitt 2.)

P-värdet blir för denna uträkning **0,0180**.

¹⁴⁴ Se kapitel 7.4 och bilaga 5.

H₁ nr 2 Gränssnitt 2 ger längre söksträngar än gränssnitt 1. (Gränssnitt 2 > gränssnitt 1.)

P-värdet blir för denna uträkning **0,3036**.

H₁ nr 3 Gränssnitt 3 ger längre söksträngar än gränssnitt 1. (Gränssnitt 3 > gränssnitt 1.)

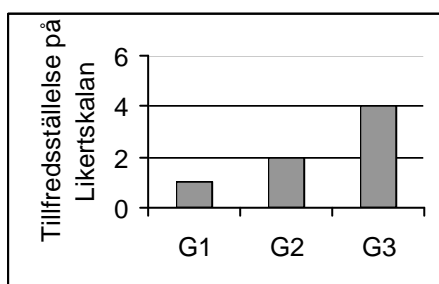
P-värdet blir för denna uträkning **0,0152**.

Både hypotes nr 1 och hypotes nr 3 ger nära tvåstjärniga resultat medan hypotes nr 2 inte har gett någon signifikans. En förklaring kan vara att skillnaden mellan gränssnitt 1 och gränssnitt 2 bara bestod av ett fält storlek, medan gränssnitt 3 hade fem olika fält som uppmuntrade till långa och kombinerade söksträngar. Visserligen hade den enkla versionen bara ett fält men ingen testperson använde sig av denna variant.

Det är svårt att dra slutsatser utifrån detta begränsade material eftersom det bara är tre gränssnitt som jämförs och dessa tre är så olika. Det är möjligt att resultatet hade blivit annorlunda om alla tre gränssnitten hade ett fält var i olika storlekar, eller om alla tre hade haft flera fält men i olika storlekar. Grundat på denna undersökning är det dock ett rimligt antagande att storlek och/eller antal påverkar söksträngslängden i positiv riktning. Eventuellt är det också så att det är antalet sökfält som påverkar söksträngslängden mest.

7.4 TILLFREDSSTÄLLELSE

Tillfredsställelsen har gett det förväntade resultatet:



Figur 8: Ett förenklat diagram över medianvärden för tillfredsställelse.

Gränssnitt 3 är överlägset, sedan kommer gränssnitt 2 och sist gränssnitt 1.

H₀ Det finns ingen skillnad mellan gränssnitten när det gäller tillfredsställelse. De skillnader som uppstått beror på slumpen.

H₁ nr 1 Gränssnitt 3 ger högre tillfredsställelse än gränssnitt 2. (Gränssnitt 3 > gränssnitt 2.)

P-värdet blir för denna uträkning **0,0040**.

H₁ nr 2 Gränssnitt 2 ger högre tillfredsställelse än gränssnitt 1. (Gränssnitt 2 > gränssnitt 1.)

P-värdet blir för denna uträkning **0,2738**

H₁ nr 3 Gränssnitt 3 ger högre tillfredsställelse än gränssnitt 1. (Gränssnitt 3 > gränssnitt 1.)

P-värdet blir för denna uträkning **0,0040**.

Tillfredsställelsen har gett anmärkningsvärt hög signifikans, med tvåstjärnig nivå för hypotes nr 1 och nr 3. Precis som med tiden och söksträngslängden är det här gränssnitt 3 som utmärker sig medan gränssnitt 1 och gränssnitt 2 har fått mer närliggande resultat där H₀ får antas. Det är anmärkningsvärt att användarna av gränssnitt 3 är så nöjda trots att de i själva verket fått lägre precision (om än ej signifikant) och längre tid. Kanske är det som Lif beskriver det: "We can be happy with a toaster just because we like the look of it even if it burns the bread."¹⁴⁵ Känslan och utseendet kan påverka mer än det konkreta sökresultatet.

Som förväntat var användarna av gränssnitt 1 missnöjda med gränssnittet och hjälpfunktionerna. G1:1 är övertygad om att det inte fanns hjälpfunktioner: "Avsaknaden av hjälpfunktioner var en klar brist eftersom man inte visste vad som var möjligt att infoga i söksträngen". Det är lätt att här dra paralleller till Shneiderman och andra författare som betonat vikten av att det bör vara tydligt för användaren vilka möjligheter som finns vid sökformulering. En annan kommentar som bekräftar Shneiderman är G1:5: "Jag tyckte inte gränssnittet var särdeles bra eftersom man knappt såg söksträngen för allt godis. Dessutom tyckte jag att söksträngen¹⁴⁶ var alldeles för liten." Jag har redan visat att sökfältets storlek och antal påverkar söksträngens längd och det är intressant att användarna också är medvetna om problematiken med för små sökfält.

Det är också förvånande att så få hittade hjälpfunktionerna i gränssnitt 2 trots att de var placerade precis intill sök-knappen. Av de fem användarna av gränssnitt 2 har fyra lämnat kommentarer på att det var svårt att finna hjälpfunktionen och ett par verkar inte ha funnit den alls: G2:4 skriver "Man blir verkligen varse hur ett dåligt gränssnitt/ett gränssnitt utan några hjälpfunktioner förvirrar och gör en maktlös." G2:5 bekräftar: "Jag insåg nu att jag inte har använt hjälpfunktionen alls!/--/När jag nu tänker efter så minns jag inte att jag såg någon hjälpfunktion." Och G2:2 bekräftar ytterligare: "jag tyckte att det var lätt att missa den hjälp knappen som man hade möjlighet att använda. Upptäckte den först i slutet av sökningen." G2:1 kanske kan komma med förklaringen: "Jag missade helt och hållet hjälp-funktionen. Trodde inte den var klickbar". Utbytet av färger från det traditionella blå till okonventionellt rosa och grått och borttagandet av understrykningar verkar alltså ha lika stor betydelse som Nielsen skriver.

¹⁴⁵ Lif, Magnus, 1998. Adding Usability: Methods for Modelling, User Interface Design and Evaluation. s. 19

¹⁴⁶ Min gissning är att testpersonen avser ordet *sökfält* snarare än *söksträng* här.

För gränssnitt 3 var det flera användare som kommenterade att det var svårt att inte få göra om sökningen. G3:3 skriver: ”Jag brukar inte tänka igenom ordentligt vad jag söker på första gången jag söker, utan mer fundera utefter de träffar jag fått när jag gör nya sökningar.” G3:4 har en liknande kommentar: ” Lite svårt eftersom man bara fick göra en sökning, man fattade ofta vad som varit problemet men man fick ju inte göra om.” Helt enligt Shneiderman är det tydligen vanligt att göra upprepade sökningar för att nå sitt mål. Det intressantaste är kanske ändå att kommentarerna för gränssnitt 3 består mer av reflexioner kring undersökningen än gränssnittets brister. Ett undantag är dock G3:5: ”vet inte om det dåliga resultatet berodde på min bristande förmåga att hitta rätt sökord eller om det hade behövts utförligare förklaringar vid sökfälten.” vilket bekräftar min misstanke om att systemet med flera sökfält komplicerade sökningarna för testpersonerna.

Kommentarerna har alltså bekräftat flera teorier och förklarat vissa resultat. Det mest intressanta är kanske att kommentarerna kring gränssnitt 1 och gränssnitt 2 rör sig runt klagomål på gränssnitten medan kommentarerna kring gränssnitt 3 har fokuserat mer på själva experimenten, som om gränssnittet inte påverkat testpersonerna så negativt att det behöver kommenteras.

Något som förvånat mig är att gränssnitt 3 har skilt sig så mycket från de andra två, då det gäller tid, söksträngslängd och tillfredsställelse. Vid designen trodde jag att gränssnitt 2 låg närmare gränssnitt 3 än gränssnitt 1 i stil och kvalitet. Det är möjligt att den lilla förändringen med länkutseendet i gränssnitt 2 gav ett helt annorlunda resultat än vad som hade varit fallet om hjälplänken hade varit blå och understruken.

7.5 SLUTSATSER

Avslutningsvis vill jag besvara de frågor som ställdes i kapitel 2.3. Eftersom mitt urval av BHS-studenter, för att inte tala om alla användare av sökmotorer på webben, inte är så stort kan jag inte generalisera över mina resultat som om det skulle gälla alla tänkbara användare men det finns helt klart tendenser som tyder på att det finns samband mellan gränssnitt och sökbeteende.

När det gäller frågan om precision är dock svaret negativt. Det verkar inte finnas några samband mellan gränssnitt och precision.

När det gäller frågan om tid finns ett klart samband mellan gränssnitt och tid, där ett mer avancerat gränssnitt med fler valmöjligheter tar längre tid att söka i, än ett enkelt där användaren bara har ett fält och inte lyckas hitta hjälpfunktionerna.

När det gäller frågan om söksträngslängd finns också ett klart samband. Söksträngslängden påverkas av sökfältets/sökfältens utformning. Användare skriver längre söksträngar i ett gränssnitt med flera och/eller större sökfält.

När det gäller frågan om tillfredsställelse finns också ett tydligt samband mellan gränssnittets utformning och den personliga tillfredsställelsen. Det gränssnitt som strävar efter att följa samtliga konventioner och rekommendationer har givit betydligt högre tillfredsställelse än de två som medvetet formgivits ”dåligt” i olika grad.

De skillnader som beskrivits ovan är signifikanta.

Skillnaderna som uppstod var till stor del i enlighet med förväntningarna och relevant litteratur. Det är något förvånande att precisionen inte blev bättre i ett ”bättre” och mer avancerat gränssnitt, förklaringen kan dock vara att ett gränssnitt inte är avgörande för sökresultat, utan att användarnas skilda förutsättningar och systemet väger in lika mycket eller mer. Frågeställningen om tid var inte riktad, i den meningen att det inte fanns en färdig hypotes om hur tiderna skulle fördela sig utan enbart en fråga om samband. Det kan dock ändå tyckas att det ”bästa gränssnittet” också borde ha givit den snabbaste tiden. Samtidigt är det rimligt att ett mer avancerat gränssnitt engagerar användaren mer och motiverar till en mer tidskrävande söksträngsformulering. Söksträngslängden påverkades helt enligt förväntningarna. Det var möjligtvis något förvånande att skillnaden mellan gränssnitt 1 och gränssnitt 2 inte blev större. Samma sak gäller för tillfredsställelsen. Gränssnitt 3 har givit överlägset störst tillfredsställelse, medan gränssnitt 1 och gränssnitt 2 ligger betydligt närmare varandra. En slutsats av detta kan vara att ett gränssnitts utformning inte får kompromissas, för en god design får inga undantag göras från rekommendationerna.

Metoden jag har använt har fungerat bra. Ett större och fullständigt slumpmässigt urval testpersoner hade varit spännande att använda till en större studie. Det kan ha varit ett misstag att välja den avancerade varianten som förval i gränssnitt 3 och kanske hade jag kunnat vara tydligare med instruktionerna att bara göra en enda sökning. Sammantaget har dock metoden och de nyskapande frågeställningarna varit spännande att arbeta med och de har förhoppningsvis också lagt grunden för framtida uppsatser.

Resultaten är både väntade och oväntade och i mitt tycke också mycket intressanta.

7.6 FÖRSLAG PÅ VIDARE FORSKNING

Under arbetet med denna uppsats har många nya frågeställningar vuxit fram, och jag tror att det med den här typen av metod går att utforska en mängd olika områden.

En idé är att göra ett liknande experiment men variera gränssnitten och testgruppen. Det är möjligt att större grupper med fler testpersoner hade gett ett annorlunda resultat. Det hade också varit intressant att göra fler gränssnitt med mindre skillnader, till exempel fem identiska gränssnitt där enbart hjälpfunktionerna skiljer sig åt. Som har framgått här kan något så enkelt som färgerna på länkarna ha en avgörande betydelse. För en student som inte själv kan konstruera webbsidor med HTML skulle det kunna vara intressant att utvärdera redan existerande sök-gränssnitt. Det är möjligt att koppla befintliga sökmotorers gränssnitt till en och samma databas och undersöka resultatet.

En studie av inlärning skulle vara spännande. Antingen med en helt ny funktion, som en ny slags sökfunktion för bibliotekskataloger, eller med extremt oerfarna användare som får lära sig att söka på webben eller annan databas. I en sådan studie skulle ett mer långsiktigt projekt vara lämpligt, där användarna får söka vid flera olika tillfällen och där utvecklingen av färdigheter och skillnader mellan gränssnitten sedan utvärderas.

För denna uppsats sparades söksträngarna i sin helhet och det var intressant att se konstruktionen av söksträngarna. Vissa söksträngar var felstavade eller hade felaktiga operatörer, något som gett nollresultat. Andra var långa och korrekta men hade ändå inte genererat någon perfekt träfflista. Sambandet söksträngslängd och precision vore intressant att utvärdera. Materialet från en liknande undersökning skulle kunna

användas till en kvalitativ undersökning om sökbeteende och söksträngskonstruktion, men också av databasens prestation och brister.

Sökbeteende vore överlag ett spännande och tacksamt uppsatsämne. Med mitt material hade jag kunnat undersöka samband mellan tid och precision, samband mellan precision och tillfredsställelse och samband mellan upplevelse och verklighet. Dessa samband vore förmodligen än mer intressanta på individnivå. Detta är dock något som inte behandlats i denna uppsats eftersom det enbart var gränssnittets påverkan som skulle undersökas.

Min förhoppning är att denna uppsats kommer att inspirera till fler uppsatser med liknande metoder och att några av de aspekter som inte kunnat behandlas här kommer att bli utredda av andra.

8. SAMMANFATTNING

Denna uppsats syfte var att ta reda på om rådande teorier om användbarhet (ordet användbarhet/usability har ersatt den gamla termen användarvänlighet) applicerade i tre olika gränssnitt påverkar verkliga sökningar på ett förväntat sätt. Precision, tidsåtgång för sökning, söksträngslängd och subjektiv tillfredsställelse var de fyra områden där jag förväntade mig att se skillnader i resultat. Frågeställningarna var:

- Påverkas sökningars precision av gränssnittets utformning?
- Är tiden som går mellan läsning av sökuppdrag och formulering av färdig söksträng beroende av gränssnittets utformning?
- Påverkas söksträngens längd av storleken på sökfältet?
- Påverkas användarnas subjektiva tillfredsställelse och upplevelse av sökning av gränssnittets utformning?
- Kan de variationer som uppstår, då det gäller de fyra faktorerna: precision, tid, söksträngslängd och tillfredsställelse, förklaras med slumpen eller är skillnaderna signifikanta?
- Påverkas resultatet enligt förväntningarna och relevant litteratur?

En teorigenomgång presenterade olika faktorer som kan skilja sig åt i gränssnitt och hur de bäst ska utformas för att vara användbara:

- **Sökfunktioner** kan skilja sig åt, det kan finnas olika antal sätt att söka på när det gäller booleska operatörer, trunkeringstecken och liknande. Sökning kan också utföras med hjälp av olika slags menyer. Det bör vara tydligt för användaren vad som kan göras och vad som orsakat en bra eller dålig träfflista. Sökfältet bör vara stort för att uppmuntra till långa söksträngar som ger högre precision.
- **Psykologiska faktorer** kan förklara orsakerna till varför ett gränssnitt bör designas på ett visst sätt. Den mänskliga hjärnan är konstruerad så att korttidsminnet belastas för mycket om inte all relevant information befinner sig på samma skärmbild. Det är också vanligt att användare relaterar ett nytt program eller funktion till liknande program och funktioner hon/han redan känner till. Därför bör gränssnitt designas så att användaren kan dra nytta av dessa ”förkunskaper” och inte bli lurad av dem.
- **Hjälpfunktioner** bör placeras i nära anslutning till sökfunktionen så att användaren kan använda sig av informationen där och ändå kunna fokusera på formuleringen av söksträngar. De bör också vara tydliga och fokusera på användarens uppgifter i stället för programmets funktioner.
- **Grafik och tydlighet** är också avgörande för ett gränssnitts användbarhet. I detta avsnitt beskrivs att länkar bör vara blå och understruken och all navigering ska vara tydlig för användaren med beskrivande länkar. Bilder bör användas sparsamt och majoriteten av skärmutrymmet ska utnyttjas till huvudinnehållet.

Utifrån dessa premisser designade jag tre gränssnitt.

- **Gränssnitt 1.** Ett ”dåligt” med stora mängder oväsentlig information, dåliga hjälpfunktioner som låg tre klick bort, länkar i fel färger och utan understrykningar. Sökfältet var mycket litet.
- **Gränssnitt 2.** Ett ”mellan” med relativt avskalad design men med irrelevanta länkar, och fel länkfärger. Hjälpen låg två klick bort men var placerad intill ”sökknappen” för att göra den lätt att hitta. Sökfältet var av mellanstorlek.
- **Gränssnitt 3.** Ett ”bra” med avskalad design, många olika sökmöjligheter, och med hjälpen placerad i ”pop-up-rutor” som visades på samma skärmbild som sökningen. Länkarna var blå och understruken. Det fanns också möjlighet att välja mellan avancerat och enkelt gränssnitt. I båda varianterna av gränssnittet var sökfälten stora, i den avancerade varianten var de dessutom fem till antalet.

Gränssnitten kopplades till Altavistas databas. 20 sökuppdrag med allmänna ämnen konstruerades. 15 studenter från BHS delades in i tre grupper som tilldelades varsitt gränssnitt som de fick utföra sökuppdragen i individuellt. Testpersonerna fick relevansbedöma sina träfflistor och svara på en mindre enkät. Precision och tid för formulering av sökfråga mättes och söksträngarna sparades i en databas för ändamålet.

Resultatet testades för statistisk signifikans med hjälp av Mann-Whitney testet. Skillnaderna i precision mellan gränssnitten blev ej statistiskt säkerställda och slutsatsen blev att det förmodligen var slumpen som avgjorde skillnaderna.

När det gäller tiden går det dock att med mycket hög säkerhet slå fast att det tar längre tid att formulera en sökfråga i gränssnitt 3 än i de båda övriga. Detta kan bero på stressande faktorer som animeringar i gränssnitt 1 och en låg känsla av personlig kontroll i de gränssnitt där användarna inte hittade hjälpfunktioner, medan det i gränssnitt 3 kunde ta längre tid att välja korrekta fält och operatörer.

Söksträngslängden gav i likhet med tiden en stark övervikt för gränssnitt 3 som signifikant översteg söksträngslängden i de båda övriga gränssnitten.

Tillfredsställelsen var liksom tiden och söksträngslängden signifikant större i gränssnitt 3.

För att besvara mina frågeställningar ovan går det alltså att svara nej på frågan om precision, men ja på frågorna om tid, söksträngslängd och tillfredsställelse. Skillnaderna är signifikanta. När det gäller förväntningar och litteratur är resultatet både väntat och oväntat. Precisionen påverkades inte enligt förväntningarna medan söksträngslängd och tillfredsställelse gjorde det. I avseende på tid fanns inga förväntningar men resultatet kan tyckas negativt, eftersom det ”bästa gränssnittet” också borde ha givit den snabbaste tiden. Samtidigt är det rimligt att ett mer avancerat gränssnitt engagerar användaren mer och motiverar till en mer tidskrävande söksträngsformulering.

KÄLLFÖRTECKNING

TRYCKTA KÄLLOR

Ahrgren, Caroline, (2000). *Elektroniska tidskriftsdatabaser: En studie om webb-presentation och gränssnittsproblematik*. Borås: högskolan i Borås, (magisteruppsats biblioteks- och informationsvetenskap 2000:39).

Allwood, Carl Martin (1991). *Människa-datorinteraktion: ett psykologiskt perspektiv*. Lund: Studentlitteratur.

Baeza-Yates, Ricardo & Ribeiro-Neto, Berthier (1999). *Modern Information Retrieval*. Harlow Addison-Wesley.

Dumas, Joseph S & Redish, Janice C (1994). *A practical guide to usability testing*. 2 uppl. Norwood, N J: Ablex pub. Corp cop.

Galitz, Wilbert O (1997). *The essential guide to user interface design: an introduction to GUI design principles and techniques*. New York: Wiley.

Hansagi, Helen & Allebeck, Peter (1994). *Enkät och intervju inom hälso- och sjukvård: handbok för forskning och utvecklingsarbete*. Lund: Studentlitteratur.

Head, Alison J (2000). *Design wise: A guide for evaluating the interface design of information resources*. 2 uppl, Medford N J: Information Today cop.

Hearst, Marti A (1999). User Interfaces and visualization. Ingår i Baeza-Yates, Ricardo & Ribeiro-Neto, Berthier. *Modern Information Retrieval*. Harlow Addison-Wesley. S. 257-323.

Johansson, Ann-Catrin & Kangas, Rita (2001). *Folkbibliotekens webbplatsers grafiska användargränssnitt: kvalitativa intervjuer med webbansvariga*. Borås: högskolan i Borås, (magisteruppsats biblioteks- och informationsvetenskap 2001:66).

Körner, Svante & Wahlgren, Lars (1998). *Statistiska metoder*. Lund: Studentlitteratur.

Large, Andrew, Tedd, Lucy A & Hartley, Richard J (2001). *Information Seeking in the Online Age: Principles and Practice*. München: Saur.

Lif, Magnus (1998). *Adding Usability: Methods for Modelling, User Interface Design and Evaluation*. Uppsala: Acta Universitatis Upsaliensis. (Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 359.) Diss. Uppsala universitet.

Lindgren, Karin (2000). *Användarvänlig design av bibliotekskataloger på webben*. Borås: högskolan i Borås, (magisteruppsats biblioteks- och informationsvetenskap 2000:78).

Löwgren, Jonas (1993). *Human-Computer Interaction: What every system developer should know*. Lund: Studentlitteratur.

Marchionini, Gary (1997). *Information Seeking in Electronic Environments*. Cambridge: Cambridge Univ. Press. (Cambridge Series on Human-Computer Interaction 9)

Mizzaro, Stefano (1997). Relevance: The Whole History. *Journal of the American Society for Information Science (JASIST)*, vol. 48, nr. 9, s. 810-832.

Nielsen, Jakob (2001). *Användbar webbdesign*. Stockholm: Liber.

Pagano, Robert R (1998). *Understanding Statistics in the Behavioral sciences*. 5 uppl, New York: Brooks-Cole.

Ramstedt, Martina & Ågren, Erika (2001). *Fritt val! Om individanpassning av sökmotorers gränssnitt*. Borås: högskolan i Borås, (magisteruppsats biblioteks- och informationsvetenskap 2001:48).

Rudisill, Marianne... [et al.], ed (1996). *Human-Computer interface design: success stories, emerging methods, and real world context*. San Fransisco, Calif. Morgan Kaufmann cop.

Stukát, Staffan (1993). *Statistikens grunder*. Lund: Studentlitteratur.

Yu, Byeong-Min & Roh, Seak-Zoon (2002). The Effects of Menu Design on Information-Seeking Performance and User's Attitude on the World Wide Web. *Journal of the American Society for Information Science and Technology*, vol. 53, nr. 11, s. 923-933.

ELEKTRONISKA KÄLLOR

AltaVista

<http://www.altavista.com/> [2003-01-29 till 2003-04-29]

Dialog Web Overview.

<http://www.dialog.com/products/dialogweb/> [2003-03-06]

Google

<http://www.google.com/> [2003-01-29 till 2003-04-29]

Gulz, Agneta (1999). "Does one size fit all?" Kognitiv variation och IT. *Human IT* 2/1999.

<http://www.hb.se/bhs/ith/2-99/ag.htm> [2003-03-03]

Lowry, Richard (1998-2003). Concepts and Applications of Inferential Statistics.

<http://faculty.vassar.edu/lowry/webtext.html> [2003-03-10]

Nielsen, Jakob (2001). Search: Visible and Simple (Alertbox May 2001). *The Alertbox: Current Issues in Web Usability*, May.

<http://www.useit.com/alertbox/20010513.html> [2003-03-03]

- Nielsen, Jakob (1996). Top Ten Mistakes in Web Design (Alertbox May 1996). *The Alertbox: Current Issues in Web Usability*, May.
<http://www.useit.com/alertbox/9605.html> [2003-03-03]
- Nielsen, Jakob (1999). "Top Ten Mistakes" Revisited (Alertbox May 1999). *The Alertbox: Current Issues in Web Usability*, May 2.
<http://www.useit.com/alertbox/990502.html> [2003-03-03]
- Nielsen, Jakob (2000). Test With 5 Users (Alertbox Mar. 2000). *The Alertbox: Current Issues in Web Usability*, March 19.
<http://www.useit.com/alertbox/20000319.html> [2003-02-16]
- Nielsen, Jakob (2001). Usability Metrics (Alertbox Jan. 2001). *The Alertbox: Current Issues in Web Usability*, Jan 21.
<http://www.useit.com/alertbox/20010121.html> [2003-03-03]
- Nationalencyklopedin på Internet, NE.se.
<http://www.ne.se/> [2003-02-16]
- Sandin, Åsa (2000). *Stilguide för webb*.
http://www.orebroll.se/webbnetet/stilguide_webb/html/typsnett_stil.htm [2003-05-29]
- Shneiderman, Ben, (1997). Clarifying Search: A User-Interface Framework for Text Searches. *D-Lib Magazine*, januari 1997.
<http://www.dlib.org/dlib/january97/retrieval/01shneiderman.html> [2003-02-26]
- Svenska datatermgruppen (2001). *Sv. datatermgruppen: Ordlista (v20)*. Ordlista, version 20, 27 november 2001 (rev. 8 maj 2002).
<http://www.nada.kth.se/dataterm/> [2003-02-26]
- User Interface Engineering (2001). *User Interface Engineering – “People Search Once, Maybe Twice” Article*. Från UIE-tips e-mail newsletter november 20 2001.
http://www.uie.com/Articles/search_once.htm [2003-03-02]

BILAGOR

BILAGA 1. STENCILEN SOM DELADES UT VID EXPERIMENTET

Hej och tack för att du ville komma och hjälpa till med detta experiment!

Nu kommer 20 sökuppdrag som du ska försöka utföra så naturligt du kan. Föreställ dig alltså att du har ett verkligt uppdrag eller intresse av att få svar på dessa frågor och bete dig därefter. Det är ingen tävling och mitt syfte är främst att undersöka det "naturliga" resultatet. Ansträng dig varken mer eller mindre än vanligt att hitta rätt sökord och/eller teknik.

Förutsätt att du inte kan mer än den information som ges i sökuppdraget. Det vill säga om frågan är: "vem skrev muminböckerna?" och du vet att svaret är Tove Jansson ska du ändå inte söka på Tove utan försöka få fram svaret ändå.

När du har utfört din sökning ska du undersöka dina 10 första träffar, är de relevanta eller inte? De så kallade sponsrade matchningarna i början räknas inte. Börja titta nedanför dem. Det är viktigt att du gör en enda söksträng! Dvs tänk igenom frågeställningen, fyll i de ord och fält du tror krävs och gör en enda sökning. Gör inte om sökningen om du inte är nöjd med resultatet. Bedöm sedan träffarna efter följande kriterier: Den specifika fråga som finns i sökuppdraget ska kunna besvaras *eller* de kriterier som sökuppdraget ställt upp ska vara uppfyllda. I de sökuppdrag som kan vara svårtolkade står det exakt hur träffarna ska relevansbedömas. Om ett dokument är relevant ska du markera dess position med ett kryss. Tänk på att det är just webbsidan du har fått träff på som ska vara relevant. Du ska inte behöva klicka på ytterligare för att få den information du söker, den ska finnas på sidan du fått träff på. Om du får två träffar från samma webbplats är båda relevanta.

T ex: du har sökt efter information om Britney Spears senaste singel. Du får först upp en rad sponsrade träffar från skivbolag och skivaffärer som du ska ignorera, efter dem kommer en träfflista med tio träffar. Av dem behandlar de första tre träffarna Britney Spears och har uppdaterade diskografier, medan träff 4-9 är webbsidor som handlar om Britney Spears som en kvinnlig förebild och inte alls tar upp hennes musik. Träff 10 handlar bara om själva singeln. Kryssa då ruta 1, 2, 3 och 10.

Du ska också fylla i hur många träffar du fått totalt.

Döda länkar (när du klickar på en träff och sidan är försvunnen) betraktas inte som relevanta träffar. Du måste därför undersöka att en webbsida finns innan du bedömer den som relevant.

Om din söksträng inte genererar några relevanta träffar, eller om du inte får någon träff alls ska du lämna samtliga fält tomma och gå vidare till nästa uppgift. Om din söksträng genererar färre än tio träffar undersöker du bara de träffar du fått. Detta gäller även om du råkat göra fel på något sätt. Om du ändå är tveksam till vad ett relevant dokument är så tänk så här: "Om jag ville veta det som står i uppdraget, skulle jag då vara nöjd med den här träffen?"

Försök att inte snegla på hur de andra testpersonerna gör, och fråga inte varandra om hjälp.

1. Gå in på startsidan. (Adressen stod uppskriven på vita tavlan. Förf. anm.)
2. Klicka på den symbol som du blir tilldelad.
3. Du ser nu första **sökuppdraget**. Vänta med att läsa det.
4. Klicka omedelbart på ”**sök-knappen**”. Ett gränssnitt visar sig i ett nytt fönster.
5. Gå tillbaka till sökuppdragets fönster. **Läs** sökuppdraget och tolka det.
6. Växla till gränssnittet och gör din **sökning** på ditt eget sätt i din egen takt.
7. **Undersök träfflistan**. Ignorera dock ”sponsrade matchningar”. Titta på samtliga träffar för att se att de inte är döda länkar.
8. Bedöm relevansen på de tio icke-sponsrade träffarna. **Markera relevanta** träffars positioner i samma fönster som sökuppdraget.
9. Skriv in **antalet** träffar du fick totalt.
10. Klicka på ”**Nästa sida**”. Nästa sökuppdrag visar sig.
11. Gå tillbaka till punkt **4**.

BILAGA 2. SÖKUPPDRAGEN

1. Hitta information om arbetsmarknaden för bibliotekarier. En relevant sida ska ta upp arbetsmarknaden och kännas relevant för dig som blivande bibliotekarie.
2. Hitta webbsidor som handlar om både Bert Karlsson och Gudrun Schyman.
3. Hitta webbsidor som handlar om bondkatter och/eller golden retriever hundar. En relevant sida ska ta upp minst en av dessa djurtyper.
4. Hitta information om sommarkurser på högskolan i Borås som går nu till sommaren. Ett relevant dokument ska ge dig någon form av information, som skulle kännas användbar om du funderade på sådana sommarkurser.
5. Hitta något om filosofen Nietche. Nietche är dock inte korrekt stavning. En relevant sida ska ge dig korrekt stavning och behandla filosofen på något sätt.
6. Blond on blond, eller var det blonde on blonde? vad är det, vem har gjort den? En relevant sida ska besvara båda dessa frågor.
7. Kommer Håkan Hellström spela på några svenska festivaler i sommar? En relevant sida ska antingen säga att han inte gör det, eller namnge minst en festival där han ska spela.
8. Hitta ett vandrarhem eller hotell i Norrland, en relevant sida ska innehålla namn på vandrarhem/hotell och gärna prisuppgift eller direkt länk till sådana uppgifter. Du som turist ska känna att det är en relevant länk du har fått.
9. Du letar efter en billig flygresa och har hört talas om Ryan Air, hitta företaget eller information om företaget.
10. Du står och väger mellan att köpa minidisc och Mp3-spelare, försök hitta sidor där du kan jämföra fördelar och nackdelar med de två. En relevant sida tar upp båda alternativen och gör någon form av jämförelse.
11. Hitta en bild på en katt, det går bra med en tecknad bild eller symbol också.
12. Har Irak kärnvapen? Hitta information och resonemang kring detta. En relevant sida ska ge dig någon form av inblick i ämnet.
13. Hitta information om Karina/Carina (oklart hur det stavas) Rydbergs bok *Den högsta kasten*.
14. Hitta en sida som har ordet snö i titeln. (Det som syns allra högst upp i webbläsaren Internet Explorer mot blå bakgrund.) Om ordet är i titeln är dokumentet relevant, annars inte.
15. Jag har hört att Malmö är en av Europas mest segregerade städer, med en tydlig uppdelning mellan svenskar och invandrare geografiskt. Ligger det något i detta? En relevant webbsida ska ge dig någon form av inblick i fenomenet.

16. Sök efter information om marknadshyror men undvik det som är skrivet om Stockholm. En relevant sida behandlar marknadshyror men nämner inte Stockholm.
17. *And sometimes when the night is slow*, var kommer detta citat ifrån, vem har skrivit det, var finns det? I ett relevant dokument får du svar på dessa frågor.
18. Det har varit en del debatt om möjligheten till adoption för homosexuella. Hitta information om den svenska debatten. En relevant webbsida ska ge dig någon form av inblick i fenomenet.
19. Du ska flytta från Borås och letar efter länkar till flyttfirmor, städfirmor och biluthyrning som du kan använda dig av. Relevanta dokument är hjälpsamma och ger dig konkreta tips och adresser.
20. Du letar efter information om Sjöbo bibliotek, men inte det Sjöbo som ligger i Skåne.

Webbsidan som innehöll sökuppdraget, relevansbedömningen, angivelsen av totala antalet träffar och de nödvändiga länkarna till gränssnittet och följande sökuppdrag visas här. En bild av det första sökuppdraget:

Sökuppdrag 1		
Hitta information om arbetsmarknaden för bibliotekarier. En relevant sida ska ta upp arbetsmarknaden och kännas relevant för dig som blivande bibliotekarie.	<input type="checkbox"/> Träff 1 är relevant	<input type="checkbox"/> Träff 2 är relevant
	<input type="checkbox"/> Träff 3 är relevant	<input type="checkbox"/> Träff 4 är relevant
	<input type="checkbox"/> Träff 5 är relevant	<input type="checkbox"/> Träff 6 är relevant
	<input type="checkbox"/> Träff 7 är relevant	<input type="checkbox"/> Träff 8 är relevant
	<input type="checkbox"/> Träff 9 är relevant	<input type="checkbox"/> Träff 10 är relevant
<input type="button" value="Sök"/>	Antal återvunna sidor: <input type="text"/>	
		<input type="button" value="nästa sida!"/>

Bilaga 3. Enkäten

Mätningen av personlig tillfredsställelse bestod av nio påståenden som testpersonerna fick instämma med i olika grad. Påståendena var:

1. Jag uppskattade detta gränssnitt och tyckte att det var trevligt att använda.
2. Jag känner mig nöjd med antalet relevanta dokument.
3. Jag använde hjälpen och läste vad som stod där.
4. Jag fann informationen i hjälpen tillräcklig, jag hittade den hjälp jag sökte.
5. Jag tyckte att hjälpen var väl placerad och lätt att komma åt.
6. Jag tyckte att gränssnittet var estetiskt tilltalande.
7. Jag hade tillräckligt med fält och möjligheter för mina sökningar.
8. Jag upplevde att sökuppdragen gick snabbt att utföra.
9. Jag tyckte att sökuppdragen var tydliga och lätta att förstå.

En bild av webbenkäten:

Enkätfrågor - Microsoft Internet Explorer erhållet av Hogskolan i Borås

Adress <http://93.51.70.91/levallnat.asp?interface=2&topic=21>

Nu efter sökuppdraget vill jag att du besvarar frågor om gränssnittet. Tank på att detta är en vetenskaplig undersökning och att det kan vara så att jag medvetet gjort just ditt gränssnitt dåligt. Du behöver alltså inte vara artig i dina bedömningar. Jag vill ha så ärliga svar som möjligt. Frågorna består av påståenden som du ska hålla med om eller inte:

Jag uppskattade detta gränssnitt och tyckte att det var trevligt att använda.	<input type="radio"/> tar helt avstånd <input type="radio"/> tar delvis avstånd <input checked="" type="radio"/> tar varken avstånd eller instämmer <input type="radio"/> instämmer delvis <input type="radio"/> instämmer helt
Jag känner mig nöjd med antalet relevanta dokument.	<input type="radio"/> tar helt avstånd <input type="radio"/> tar delvis avstånd <input checked="" type="radio"/> tar varken avstånd eller instämmer <input type="radio"/> instämmer delvis <input type="radio"/> instämmer helt
Jag använde hjälpen och läste vad som stod där.	<input type="radio"/> tar helt avstånd <input type="radio"/> tar delvis avstånd <input checked="" type="radio"/> tar varken avstånd eller instämmer <input type="radio"/> instämmer delvis <input type="radio"/> instämmer helt
Jag fann informationen i hjälpen tillräcklig, jag hittade den hjälp jag sökte.	<input type="radio"/> tar helt avstånd <input type="radio"/> tar delvis avstånd <input checked="" type="radio"/> tar varken avstånd eller instämmer <input type="radio"/> instämmer delvis <input type="radio"/> instämmer helt
Jag tyckte att hjälpen var väl placerad och lätt att komma åt.	<input type="radio"/> tar helt avstånd <input type="radio"/> tar delvis avstånd <input checked="" type="radio"/> tar varken avstånd eller instämmer <input type="radio"/> instämmer delvis <input type="radio"/> instämmer helt
Jag tyckte att gränssnittet var estetiskt tilltalande.	<input type="radio"/> tar helt avstånd <input type="radio"/> tar delvis avstånd <input checked="" type="radio"/> tar varken avstånd eller instämmer <input type="radio"/> instämmer delvis <input type="radio"/> instämmer helt
Jag hade tillräckligt med fält och möjligheter för mina sökningar.	<input type="radio"/> tar helt avstånd <input type="radio"/> tar delvis avstånd <input checked="" type="radio"/> tar varken avstånd eller instämmer <input type="radio"/> instämmer delvis <input type="radio"/> instämmer helt
Jag tyckte att sökuppdragen var tydliga och lätta att förstå.	<input type="radio"/> tar helt avstånd <input type="radio"/> tar delvis avstånd <input checked="" type="radio"/> tar varken avstånd eller instämmer <input type="radio"/> instämmer delvis <input type="radio"/> instämmer helt
Jag upplevde att sökuppdragen gick snabbt att utföra.	<input type="radio"/> tar helt avstånd <input type="radio"/> tar delvis avstånd <input checked="" type="radio"/> tar varken avstånd eller instämmer <input type="radio"/> instämmer delvis <input type="radio"/> instämmer helt
Har du några kommentarer? Har jag missat något viktigt? Var det något speciellt med gränssnittet du tyckte var bra, underligt, störande? Förklara gärna de punkter där du fyllt i att du tar avstånd från påståendet.	<input type="text" value="Kommentarer: här!"/>

BILAGA 4. GRÄNSSNITTENS UTSEENDE

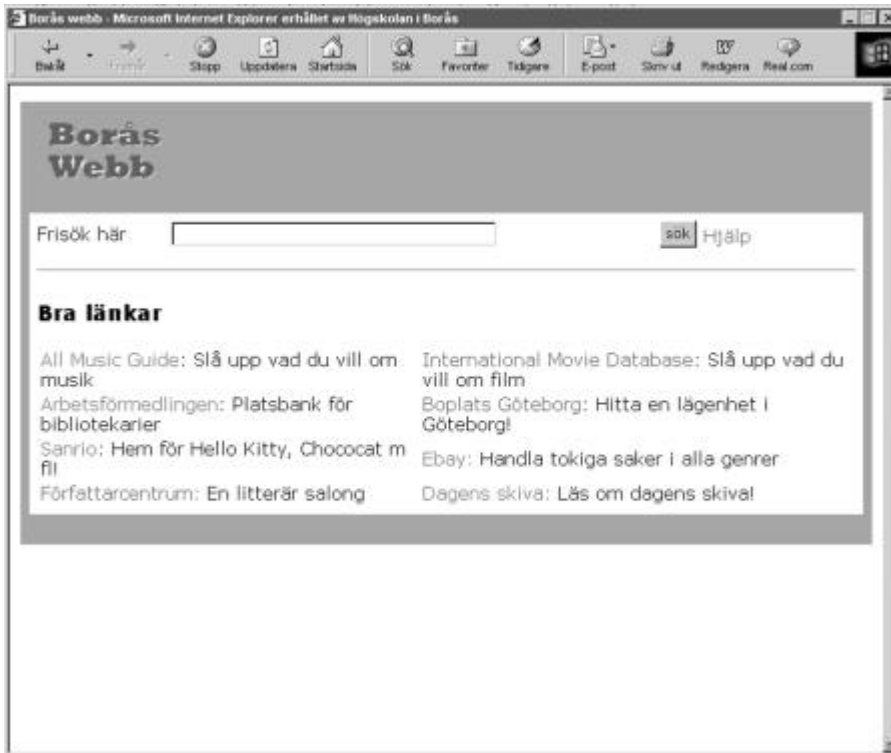
GRÄNSSNITT 1

Eftersom gränssnitt 1 var så långt är bilden ett montage. Detta förklarar den något underliga scroll-listen på bilden. De tre översta reklambilderna var animerade, likaså de två loggorna "Borås webb" som roterade i hög hastighet. Färgerna var starka, Godisreklamen hade klara färger, och färgerna i övrigt bestod av mörkt rött, grönt, blått, ljusgrått, mörkgrått och orange.



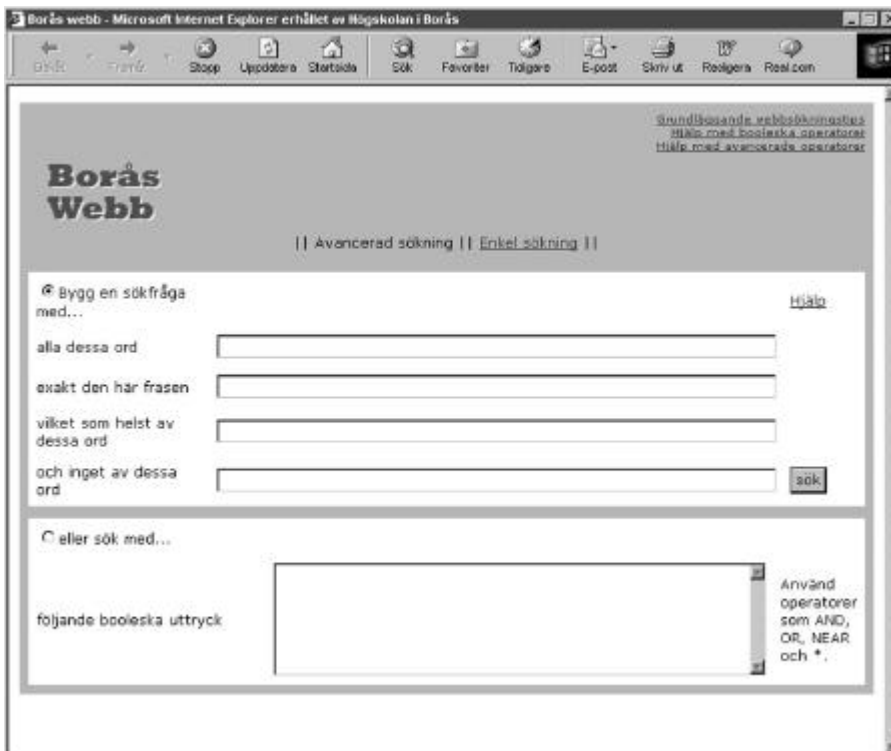
GRÄNSSNITT 2

Här är färgen starkt rosa. Länkarna är rosa och grå.

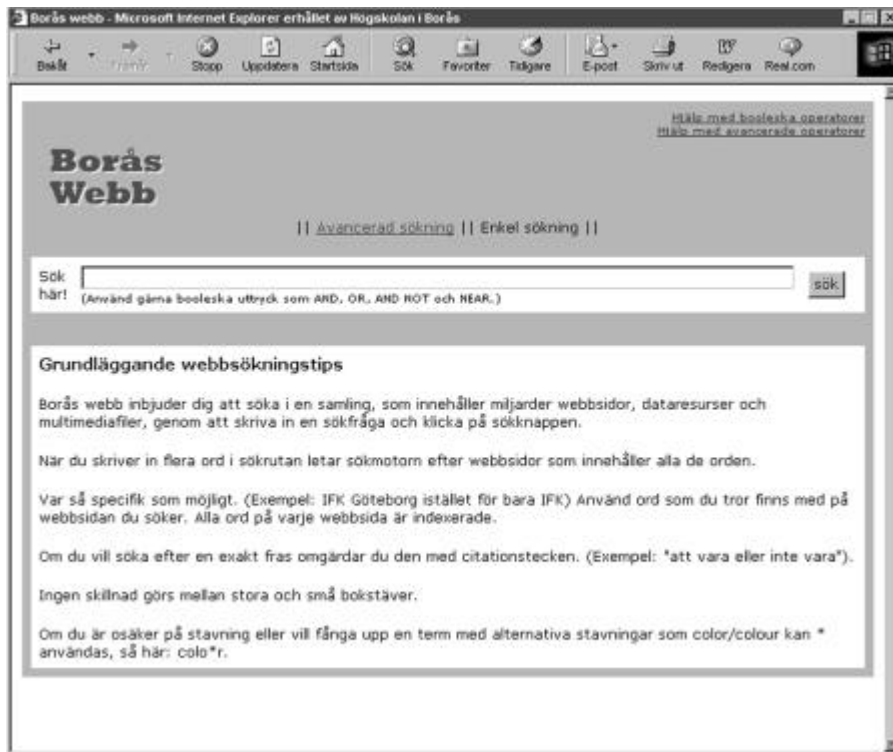


GRÄNSSNITT 3

Här är den avancerade varianten. Färgen är vilsamt rosa, medan logon är blå.



Den enkla varianten:



BILAGA 5. DE ENSKILDA TESTPERSONERNAS SÖKRESULTAT

Här redovisas de enskilda testpersonernas resultat. Även de aritmetiska medelvärdena och medianvärdena redovisas i följande tabeller. Tabellerna redovisas i samma ordning som i övriga uppsatsen, precision, tid, söksträngslängd och sist tillfredsställelse. Inom varje avsnitt kommer gränssnitten att redovisas i nummerordning.

PRECISION

Nedan visas precision för gränssnitt 1, ”dåliga gränssnittet”. G1 är medeltalen för varje enskild sökfråga och G1s totala medelvärde är alltså ett mått på hela gränssnittets medelvärde.

Användare:	G1:1	G1:2	G1:3	G1:4	G1:5	G1
Sökuppdrag 1	0	0	0	0	0	0
Sökuppdrag 2	0,33	0	0	0,3	0,5	0,226
Sökuppdrag 3	0,5	0	0	0,4	0	0,18
Sökuppdrag 4	0,1	0	0	0,1	0	0,04
Sökuppdrag 5	0	0	0,1	0,1	0	0,04
Sökuppdrag 6	0,1	0	0	0,4	0	0,1
Sökuppdrag 7	0	0	0,2	0,09	0,1	0,078
Sökuppdrag 8	0,2	0	0,6	0,3	0	0,22
Sökuppdrag 9	0,3	0	0,3	0	0,5	0,22
Sökuppdrag 10	0	0	0	0,1	0,4	0,1
Sökuppdrag 11	0	0,7	0,6	0,3	0,8	0,48
Sökuppdrag 12	0,1	0,1	0,4	0,7	0,5	0,36
Sökuppdrag 13	0,3	0,5	0,5	0,4	0,9	0,52
Sökuppdrag 14	0,2	0,9	0,4	0,6	0,9	0,6
Sökuppdrag 15	0,1	0,3	0,3	0,3	0	0,2
Sökuppdrag 16	0	0	0,2	0,2	0,1	0,1
Sökuppdrag 17	0,8	0,8	0,7	1	0	0,66
Sökuppdrag 18	0,5	0,1	0,6	1	0,1	0,46
Sökuppdrag 19	0	0	0	0,5	0	0,1
Sökuppdrag 20	0,2	0	0,3	0,3	0	0,16
Medeltal:	0,1865	0,17	0,26	0,3545	0,24	0,2422

Figur 9: Precision för testpersonerna som använde gränssnitt 1

Precision för gränssnitt 2, ”mellangränssnittet” visas nedan. G2:3 var den testperson vars resultat inte blivit korrekt införda och därför redovisas inte dessa resultat här.

Användare:	G2:1	G2:2	G2:4	G2:5	G2
Sökuppdrag 1	0	0	0,2	0,6	0,2
Sökuppdrag 2	0	0	0,6	0,3	0,225
Sökuppdrag 3	1	0	0,25	0,6	0,4625
Sökuppdrag 4	0	0	0	0	0
Sökuppdrag 5	0	0	0	0	0
Sökuppdrag 6	0,67	0	0,5	0,8	0,4925
Sökuppdrag 7	0,4	0	0,5	0,2	0,275
Sökuppdrag 8	0	0	0,1	0	0,025
Sökuppdrag 9	0,4	0	0,4	0,4	0,3
Användare:	G2:1	G2:2	G2:4	G2:5	G2

Sökuppdrag 10	0	0	0	0	0
Sökuppdrag 11	0	0,4	0,9	0,3	0,4
Sökuppdrag 12	0,4	0,1	0,5	0,6	0,4
Sökuppdrag 13	1	0,2	0,4	0,6	0,55
Sökuppdrag 14	0,7	0,4	0,4	0,5	0,5
Sökuppdrag 15	0,3	0,2	0	0,4	0,225
Sökuppdrag 16	0	0	0	0	0
Sökuppdrag 17	0,9	0,7	1	0	0,65
Sökuppdrag 18	0,7	0,1	0,6	0,9	0,575
Sökuppdrag 19	0	0	0	0	0
Sökuppdrag 20	0	0	0,2	0,2	0,1
Medeltal:	0,3235	0,105	0,3275	0,32	0,269

Figur 10: Precision för testpersonerna som använde gränssnitt 2

Precision för gränssnitt 3, visas nedan.

Användare:	G3:1	G3:2	G3:3	G3:4	G3:5	G3
Sökuppdrag 1	0	0	0	0,2	0	0,04
Sökuppdrag 2	0	0,2	0,4	0	0,1	0,14
Sökuppdrag 3	0	0	0,5	0,1	0	0,12
Sökuppdrag 4	0	0	0	0	0,1	0,02
Sökuppdrag 5	0	0	0	0,5	0	0,1
Sökuppdrag 6	0	0,1	0,5	0,8	0,1	0,3
Sökuppdrag 7	0	0	0	0,3	0	0,06
Sökuppdrag 8	0	0,1	0,3	0	0,3	0,14
Sökuppdrag 9	0	0,3	0,1	0	0,5	0,18
Sökuppdrag 10	0	0	0,1	0	0	0,02
Sökuppdrag 11	0	0,1	0,3	0	0	0,08
Sökuppdrag 12	0	0,1	0	0,3	0	0,08
Sökuppdrag 13	0	0	0,4	0	0	0,08
Sökuppdrag 14	0	0,3	0,1	0,7	0,3	0,28
Sökuppdrag 15	0	0,2	0,2	0,2	0,1	0,14
Sökuppdrag 16	0	0,2	0,1	0,1	0,3	0,14
Sökuppdrag 17	0,8	0,2	0	0,8	0,9	0,54
Sökuppdrag 18	0,8	0,2	0,8	0	0	0,36
Sökuppdrag 19	0	0	0	0,3	0,8	0,22
Sökuppdrag 20	0,7	0	0,7	0,6	0,8	0,56
Medeltal:	0,115	0,1	0,225	0,245	0,215	0,18

Figur 11: Precision för testpersonerna som använde gränssnitt 3

Tid

Måttenheten här är tusendels sekunder. Observera de stora skillnaderna i tid då det gäller första sökuppdraget.

Gränssnitt 1:

Användare:	G1:1	G1:2	G1:3	G1:4	G1:5	G1
Sökuppdrag 1	280 859	254 406	12 078	22 187	19 187	117 743
Sökuppdrag 2	34 906	42 391	11 969	23 406	39 453	30 425
Sökuppdrag 3	109 187	54 969	26 656	42 562	15 891	49 853
Sökuppdrag 4	78 860	23 672	38 203	43 844	22 328	41 381
Användare:	G1:1	G1:2	G1:3	G1:4	G1:5	G1

Sökuppdrag 5	169 391	38 265	4 969	28 781	11 484	50 578
Sökuppdrag 6	44 750	21 031	22 172	227 141	25 719	68 163
Sökuppdrag 7	211 156	25 109	37 031	109 875	15 062	79 647
Sökuppdrag 8	60 047	44 531	27 907	92 015	25 016	49 903
Sökuppdrag 9	23 828	12 734	6 297	88 109	11 828	28 559
Sökuppdrag 10	131 094	15 500	20 563	37 062	138 734	68 591
Sökuppdrag 11	53 578	13 421	6 593	66 359	12 375	30 465
Sökuppdrag 12	74 875	16 860	8 109	32 234	14 203	29 256
Sökuppdrag 13	31 907	24 281	10 797	53 969	19 672	28 125
Sökuppdrag 14	76 515	6 359	5 312	40 078	4 547	26 562
Sökuppdrag 15	66 422	30 860	20 219	59 328	14 890	38 344
Sökuppdrag 16	39 171	17 094	8 188	55 219	23 750	28 684
Sökuppdrag 17	25 641	16 063	19 188	27 312	27 593	23 159
Sökuppdrag 18	30 187	12 281	9 719	19 078	27 547	19 762
Sökuppdrag 19	75 172	42 375	20 219	20 546	35 688	38 800
Sökuppdrag 20	53 172	23 219	10 735	25 187	25 829	27 628
Medeltal:	83 536	36 771	16 346	55 715	26 540	43 782

Figur 12: Tiden för testpersonerna som använde gränssnitt 1

Testperson G2:3s resultat var felaktigt införda i databasen och därför redovisas inte resultaten här. Gränssnitt 2:

Användare:	G2:1	G2:2	G2:4	G2:5	G2
Sökuppdrag 1	132 938	15 047	13 375	31 750	48 278
Sökuppdrag 2	35 266	13 938	46 937	28 969	31 278
Sökuppdrag 3	62 922	91 219	86 016	32 500	68 164
Sökuppdrag 4	66 563	151 250	49 125	38 109	76 262
Sökuppdrag 5	119 328	48 031	112 500	42 515	80 594
Sökuppdrag 6	24 985	23 515	41 578	32 438	30 629
Sökuppdrag 7	42 375	15 891	64 578	95 203	54 512
Sökuppdrag 8	56 000	77 171	74 297	100 640	77 027
Sökuppdrag 9	33 781	22 891	8 985	9 719	18 844
Sökuppdrag 10	87 797	77 859	61 266	181 859	102 195
Sökuppdrag 11	23 750	13 672	16 297	32 937	21 664
Sökuppdrag 12	72 156	38 734	20 203	24 891	38 996
Sökuppdrag 13	23 828	12 188	24 515	37 562	24 523
Sökuppdrag 14	26 531	14 750	29 437	25 313	24 008
Sökuppdrag 15	31 781	33 422	45 125	38 062	37 098
Sökuppdrag 16	23 187	34 187	84 563	36 422	44 590
Sökuppdrag 17	51 797	3 079	28 766	26 375	27 504
Sökuppdrag 18	27 625	102 172	56 156	42 688	57 160
Sökuppdrag 19	61 828	54 547	62 672	68 172	61 805
Sökuppdrag 20	22 516	64 734	45 875	34 422	41 887
Medeltal:	51 348	45 415	48 613	48 027	48 351

Figur 13: Tiden för testpersonerna som använde gränssnitt 2

Gränssnitt 3:

Användare:	G3:1	G3:2	G3:3	G3:4	G3:5	G3
Sökuppdrag 1	45 672	83 360	5 766	354 703	276 703	153 241
Sökuppdrag 2	45 672	84 437	31 844	121 985	39 422	64 672
Sökuppdrag 3	20 297	40 094	46 734	144 390	115 828	73 469
Användare:	G3:1	G3:2	G3:3	G3:4	G3:5	G3

Sökuppdrag 4	11 031	50 265	54 329	171 000	35 625	64 450
Sökuppdrag 5	27 828	70 672	108 984	177 875	65 078	90 087
Sökuppdrag 6	45 047	85 438	39 000	258 734	65 422	98 728
Sökuppdrag 7	34 688	32 906	78 203	119 641	114 594	76 006
Sökuppdrag 8	49 860	111 891	97 719	138 359	49 750	89 516
Sökuppdrag 9	5 735	25 547	8 968	116 282	91 781	49 663
Sökuppdrag 10	57 297	62 375	44 172	142 328	118 532	84 941
Sökuppdrag 11	6 297	12 360	21 547	85 359	40 047	33 122
Sökuppdrag 12	32 563	50 641	38 235	104 469	69 954	59 172
Sökuppdrag 13	78 281	47 765	22 359	104 610	61 516	62 906
Sökuppdrag 14	63 906	23 047	7 891	62 391	67 875	45 022
Sökuppdrag 15	20 469	33 750	31 250	176 422	46 687	61 716
Sökuppdrag 16	20 453	28 375	78 641	44 860	61 203	46 706
Sökuppdrag 17	14 000	45 594	27 828	61 203	70 188	43 763
Sökuppdrag 18	23 375	15 031	47 703	116 547	46 500	49 831
Sökuppdrag 19	45 532	63 594	78 922	111 922	62 563	72 507
Sökuppdrag 20	24 500	30 109	33 765	75 266	38 047	40 337
Medeltal:	33 625	49 863	45 193	134 417	76 866	67 993

Figur 14: Tiden för testpersonerna som använde gränssnitt 3

SÖKSTRÄNGENS LÄNGD

Här redovisas resultatet av mätningarna av söksträngarnas längd. Värdena är angivna som antal tecken inklusive mellanslag och operatorer.

Gränssnitt 1:

Användare:	G1:1	G1:2	G1:3	G1:4	G1:5	G1
Sökuppdrag 1	109	53	19	27	34	48,4
Sökuppdrag 2	36	36	21	32	32	31,4
Sökuppdrag 3	31	28	24	28	17	25,6
Sökuppdrag 4	32	25	27	33	32	29,8
Sökuppdrag 5	32	15	7	20	7	16,2
Sökuppdrag 6	18	38	14	39	6	23
Sökuppdrag 7	61	29	38	33	30	38,2
Sökuppdrag 8	33	32	26	42	33	33,2
Sökuppdrag 9	10	28	8	12	8	13,2
Sökuppdrag 10	34	14	21	28	16	22,6
Sökuppdrag 11	28	13	6	11	10	13,6
Sökuppdrag 12	20	10	14	28	16	17,6
Sökuppdrag 13	27	29	21	31	26	26,8
Sökuppdrag 14	3	3	3	9	3	4,2
Sökuppdrag 15	24	29	25	17	16	22,2
Sökuppdrag 16	26	27	13	30	26	24,4
Sökuppdrag 17	38	38	38	38	36	37,6
Sökuppdrag 18	21	6	21	25	28	20,2
Sökuppdrag 19	60	40	41	22	20	36,6
Sökuppdrag 20	22	27	15	31	26	24,2
Medel	33	26	20	27	21,1	25,45

Figur 15: Antal tecken per söksträng som användes i gränssnitt 1

Gränssnitt 2:

Användare:	G2:1	G2:2	G2:3	G2:4	G2:5	G2
Sökuppdrag 1	20	27	28	28	28	26,2
Sökuppdrag 2	36	28	33	36	28	32,2
Sökuppdrag 3	31	30	31	28	28	29,6
Sökuppdrag 4	42	28	33	36	25	32,8
Sökuppdrag 5	20	13	9	21	10	14,6
Sökuppdrag 6	38	16	36	38	34	32,4
Sökuppdrag 7	40	32	31	35	42	36
Sökuppdrag 8	35	25	29	32	33	30,8
Sökuppdrag 9	10	8	10	10	8	9,2
Sökuppdrag 10	31	14	23	24	47	27,8
Sökuppdrag 11	16	11	21	10	9	13,4
Sökuppdrag 12	48	15	14	19	44	28
Sökuppdrag 13	31	19	27	27	29	26,6
Sökuppdrag 14	12	5	8	3	3	6,2
Sökuppdrag 15	21	13	17	32	28	22,2
Sökuppdrag 16	26	26	25	27	35	27,8
Sökuppdrag 17	34	38	38	38	38	37,2
Sökuppdrag 18	25	38	25	36	20	28,8
Sökuppdrag 19	48	53	32	19	39	38,2
Sökuppdrag 20	27	25	21	15	16	20,8
Medel	30	23	25	26	27,2	26,04

Figur 16: Antal tecken per söksträng som användes i gränssnitt 2

Gränssnitt 3:

Användare:	G3:1	G3:2	G3:3	G3:4	G3:5	G3
Sökuppdrag 1	28	27	20	76	54	41
Sökuppdrag 2	38	32	36	28	56	38
Sökuppdrag 3	27	31	32	35	52	35,4
Sökuppdrag 4	30	27	31	32	36	31,2
Sökuppdrag 5	7	8	46	44	30	27
Sökuppdrag 6	38	34	38	20	21	30,2
Sökuppdrag 7	21	32	44	64	35	39,2
Sökuppdrag 8	30	19	35	92	39	43
Sökuppdrag 9	8	8	13	32	29	18
Sökuppdrag 10	25	22	32	43	57	35,8
Sökuppdrag 11	9	13	5	58	29	22,8
Sökuppdrag 12	15	13	19	52	49	29,6
Sökuppdrag 13	25	22	31	39	49	33,2
Sökuppdrag 14	3	13	16	9	6	9,4
Sökuppdrag 15	17	20	17	58	35	29,4
Sökuppdrag 16	34	22	21	22	35	26,8
Sökuppdrag 17	14	36	35	38	58	36,2
Sökuppdrag 18	22	24	50	51	44	38,2
Sökuppdrag 19	60	19	32	42	27	36
Sökuppdrag 20	20	31	33	34	31	29,8

Användare:	G3:1	G3:2	G3:3	G3:4	G3:5	G3
Medel	24	23	29	43	38,6	31,51

Figur 17: Antal tecken per söksträng som användes i gränssnitt 3

TILLFREDSSTÄLLELSE

Påståendena förkortas nedan med "P", P 1 är alltså påstående 1. Eftersom Likertskalan är en ordinalskala har *medianvärdet* angivits för de fiktiva testpersonerna G1, G2 och G3.

	P 1	P 2	P 3	P 4	P 5	P 6	P 7	P 8	Median
G1:1	1	1	1	1	1	1	2	4	1
G1:2	2	1	3	2	1	1	1	4	1,5
G1:3	4	4	1	1	1	5	2	2	2
G1:4	2	1	5	4	1	1	2	2	2
G1:5	1	2	1	1	1	2	1	3	1
G1	2	1	1	1	1	1	2	3	1
G2:1	3	1	1	3	1	2	2	4	2
G2:2	2	1	2	2	2	1	1	4	2
G2:3	3	2	1	1	3	3	3	2	2,5
G2:4	1	2	3	3	3	1	1	2	2
G2:5	4	2	1	1	2	4	2	4	2
G2	3	2	1	2	2	2	2	4	2
G3:1	3	2	4	4	3	4	5	5	4
G3:2	2	2	2	3	3	4	4	2	2,5
G3:3	4	2	2	4	4	4	4	3	4
G3:4	4	2	4	3	4	4	2	3	3,5
G3:5	4	1	4	2	5	3	4	4	4
G3	4	2	4	3	4	4	4	3	4

Figur 18: Ett diagram över genomsnittlig söksträngslängd för de tre gränssnitten