

Linköping Studies in Arts and Science

DEPARTMENTSERIES No. XXX

Dissertations. No. THESISNUMBER

En explorativ designstudie av läroverktyg för surfplattor

av

Johan Rittfeldt



Linköping University

Institutionen för datavetenskap
Linköpings universitet
581 83 Linköping

Linköping 2015

Copyright © 2015

ISBN ISBN
ISSN 0282-9800
Tryckt av LiU Tryck 2015

URL: <http://XXX>

Innehåll

1	Introduktion	2
1.1	Syfte	3
1.2	Frågeställningar	3
1.3	Avgränsningar	3
2	Bakgrund	4
3	Teori och empiri	6
3.1	Korttidsminnet	6
3.2	Skissning	7
3.3	Skärmläsning	8
3.4	Skärmtypografi	8
3.5	Färgseende	9
4	Metod	10
4.1	Design	10
4.2	Teknisk implementation	11
4.3	Datainsamling	13
4.3.1	Pilot	14
4.3.2	Huvudstudie	14
4.3.3	Analys	15
5	Resultat	16
5.1	Observationer	16
5.2	Intervjuer	17
5.3	Testpoäng och preferenser	18
6	Diskussion	19
6.1	Resultatdiskussion	19
6.2	Metodval och genomförande	20
6.3	Validitet och generaliserbarhet	20
6.4	Framtida studier	22
7	Slutsats	23

I samarbete mellan flera svenska universitet framarbetas ett verktyg för grundskoleelever som bland annat avser att automatiskt förenkla, sammanfatta eller finna befintliga texter på lämplig nivå för individen. Det läsförståelsetest verktyget baseras på har tidigare delvis skett på en datorskärm. Uppsatsen har undersökt preferensen av två olika designer som tagits fram i syfte att digitalisera testet till fullo. Skillnaden de båda alternativen diskuteras i rapporten liksom dess implikation för resultatet. En design där text presenteras parallellt med frågor föredras av en överväldigande majoritet av eleverna över alternativet att skifta mellan visningen av de båda. Inga tydliga signaler ges på att det skulle vara vare sig en nackdel eller fördel över alternativ lösning.

Kapitel 1

Introduktion

Skolan måste ständigt reformeras för att hänga med i den tekniska utvecklingen i samhället. Kursinnehåll, pedagogik, studiemiljö och utrustning är några av alla de faktorer som kan påverka utfallet av undervisningen. På många platser erbjuds eleverna egna bärbara datorer eller surfplattor som stöd för undervisningen med alla fördelar som följer tekniken. Flera rapporter pekar dock på nedsättningar förknippade med att läsa från skärm kontra papper. Exempelvis menar Mayes et al. (2001) att läsning från skärm är avsevärt långsammare och Wästlund et al. (2005) fann i sin studie att papper var överlägsen gällande förståelse. Många rapporter visar dock att skillnaderna inte är signifikanta de båda medium emellan (Bodmann and Robinson, 2004; Garland and Noyes, 2004). Problematiken med att nå åtminstone en jämbördig nivå aktualiseras för läsförståelsetestet T-MASTER som är den tilltänkta produkten av ett samarbete mellan flera svenska universitet. Det avser att kunna lyfta fram för grundskoleelever lämpliga texter och automatisk sammanfatta eller omvandla andra texter till lättläst svenska (Mühlenbock et al., 2014). T-MASTER avses att kunna genomföras på ett spektrum av plattformar där surfplattor antas tillhöra de minsta enheterna. För uppsatsen faller fokus på mindre skärmar. Testet har under tidigare skede av utvecklingen genomförts på papper. Det kräver manuell hantering vid rättning vilket tar tid och öppnar för potentiella misstag under processen. Utmaningen är att på den begränsade skärmytan inkludera testets två huvudkomponenter, text samt tillhörande frågor, på ett sätt som inte missgynnar eleven. För att efterlikna pappersalternativet i största möjliga mån krävs att texten occuperar hela bredden av skärmen. Men kan text och frågor dela på skärmytan utan att det går ut över läsbarheten? Två olika designer har framarbetats i syfte att stödja digitala plattformar med mindre skärmar. Dessa testas för att undersöka hur väl de mottas av eleverna och huruvida de inverkar på prestationen.

1.1 Syfte

Uppsatsen är en del av ett större forskningsprojekt där ett läsförståelsetest bestående av text och frågor ska presenteras på en begränsad yta på lämpligt vis. Studiens syfte blir således att identifiera en passande lösning för presentation och interaktion av en tvådelad informationsmängd.

1.2 Frågeställningar

- Med begränsad skärmyta för ett läsförståelsetest, föredras text och frågor separerat i fullskärm eller parallellt med varandra?
- Inverkar utformningen på läsförståelsen?

1.3 Avgränsningar

Olika skolor har sina respektive IT-plattformar. Vissa har enbart datorsalar, andra har till elever individuellt tilldelade datorer eller surfplattor av varierande typ. För att täcka uppsatsens syfte inom de ramar som föreligger har en specifik hårdvara valts som utgångspunkt. iPad är den modell som i arbetet givits fokus på önskemål av projektledaren. Det går även i linje med utvecklingsstrategin “mobile first” (Johnson, 2013). Detta koncept innebär att i stället för att börja anpassa sin produkt till desktop direkt angriper de minsta plattformarna först. Detta grundas i stadigt ökande mobilanvändande men också för att det ofta i praktiken är enklare att skala upp sitt gränssnitt än att skala ner det. Surfplattor av diverse fabrikat antas tillhöra de minsta skärmar som elever använder i samband med undervisning. Den mindre skärmen utgör större utmaningar i form av att nyttja tillgänglig yta på bästa vis.

Utöver fysiska aspekter knutet till hårdvaran kan också fysiska varianser hos de tilltänkta användarna medföra upplevelsen blir olika. Idag beräknas ungefär 8% av alla män i Nordamerika och Europa ha någon form av defekt färgsinne medan kvinnor är mer förskonade med från nedsättningen med 1% (MacDonald, 1999). Hänsyn tas i uppsatsen till nedsättningar i färgseendet, dock ej till dyslexi och andra funktionsnedsättningar som kan inverka på testet.

Läsförståelsetestet riktar sig till elever mellan 10-15 år, dvs årskurs fyra till nio. I uppsatsen genomförs tester på årskurs fyra samt åtta. De som faller utanför målgruppen kan ändå resoneras kring i avsnittet för diskussion.

Kapitel 2

Bakgrund

Uppsatsen är en del av ett gemensamt forskningsprojekt mellan universiteten i Linköping, Göteborg och Uppsala som involverar forskare från flera olika discipliner. Projektet, som går under namnet MASTER, som syftar till utveckla tekniker och verktyg för att lyfta fram lämpliga texter, automatisk sammanfatta texter samt omvandla texter till lättläst svenska för elever mellan tio och femton år (Mühlenbock et al., 2014).

Flera studier (Reichenberg, 2000; Liberg, 2010) har påvisats hur en mindre god läsare kan dra nytta av anpassade texter med alternativa lingvistiska egenskaper. Liberg undersöker i sin studie vilka stödstrukturer som kan hjälpa till i undervisningen då elever besitter olika potential till att läsa och lära. Bland annat dessa insikter har legat till grund för projektet.

För att avgöra en texts läsbarhet används det så kallade SVIT-måttet (Heimann Mühlenbock, 2013). Det kan bedöma mer ytliga lingvistiska egenskaper av en text såväl som på språkligt djupare nivåer. Testen eleverna genomför inkluderar olika texttyper av varierande lingvistisk svårighet fördelat över olika ämnen. I samtliga av texterna testas förmågan att:

1. Finna explicit uttryckt information och
2. göra enkla inferenser och
3. tolka och integrera idéer och information samt
4. reflektera, undersöka och evaluera innehåll, språk och textuella element.

Via ett webbgränssnitt ska lärare ges stöd för att finna texter på lämplig läsnivå för eleverna. Där kan elever genomföra läsförståelsetest för texter på olika svårighetsnivåer och automatiskt få fram en profil för sin läsförmåga för dessa texter. Med hjälp av profilen kan läraren få stöd för att söka fram texter på sådan svårighetsnivå som eleven kan klara av att läsa på ett tillfredsställande sätt. Läraren kommer också kunna förenkla/sammanfatta texter till en sådan svårighetsnivå som lämpar sig för den enskilde eleven.

Efter av eleven fullfört test analyseras resultaten i T-MASTER och mappas mot textuell beskrivning av läsförmågan (Mühlenbock et al., 2014). Beskrivningen görs tillgänglig för ansvarig lärare och presenteras i lättförståeligt format (se appendix för exempel).

Läsförståelsetestet har tidigare skett antingen helt på papper eller delvis där texten är på papper och flervalsfrågorna tillgängliggörs via skärm. Prototypen för uppsatsen har avsett att digitalisera testet fullt ut. Då hela testet kan ske digitalt innebär det inte bara att eleverna alltid har tillgång till senaste versionen i tillfälle av förändringar, men framför allt försvinner behovet för manuell hantering av resultat och återkoppling sker direkt efter testet är slutfört.

Eftersom inga nationella föreskrifter begränsar vilka plattformar svenska skolor ska basera sin IT kring ligger valen fria och öppnar för många variationer när det gäller såväl hårdvara som mjukvara. T-MASTER-applikationen måste därför täcka in ett spektrum av olika typer varför en webblösning är att föredra då den är plattformsoberoende. Det innebär i teorin att det inte spelar någon roll om eleven sitter vid en stationär dator eller surplatta. I praktiken så krävs dock individuella anpassningar och avvägningar. Exempelvis skiljer sig olika webbläsare gällande vilka webbstandarder som följs och i vilken omfattning det sker. Dessutom är de fysiska aspekterna av hårdvaran relevant, i synnerhet gällande proportioner och storlek på skärm. En mindre yta att presentera information på ställer naturligtvis högre krav på utformningen av gränssnittet för att kunna åskådliggöra samma information tydligt och enkelt utan att göra avkall på nödvändig funktionalitet.

Kapitel 3

Teori och empiri

I detta kapitel beskrivs de teorier som legat till grund för de metoder som använts. I tillägg ges en förklaring av begrepp som används i resultatdiskussionen och annan typ av relevant information som påverkat uppsatsen.

3.1 Korttidsminnet

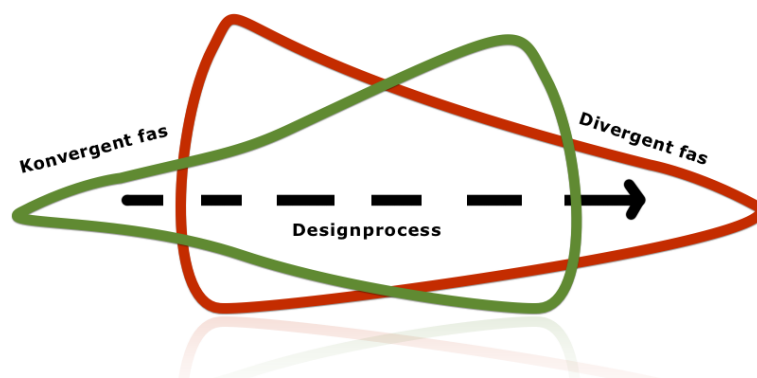
Det finns vissa implikationer för korttidsminne som bör belysas vid design av användargränssnitt. Korttidsminnet är flyktigt och har en snävt begränsad kapacitet. Det är därför viktigt att i den mån som låter sig göras understödja användaren av systemet. Vad som är gjort eller vilken status systemet har ska inte behöva hållas i korttidsminnet när fokus är på ett mål och framskridandet mot det (Johnson, 2010)

På grund av nämnda begränsningar är det ofta rekommenderat att undvika olika lägen (eng. *modes*) i användargränssnitt. Ett exempel på olika typer av lägen är hur kameraknappen kan utföra olika funktionalitet beroende på om kameran är inställd på bild eller video. Även om lägen öppnar för potentiella fel existerar de av en anledning. I exemplet med kameran hade alternativet krävt en ytterligare knapp. Gällande gränssnitt handlar det om att lyfta in ytterligare funktionalitet genom att tillåta användaren att handla i olika lägen. Med det följer risken att hamna i *lägesfel*, vilket innebär att användaren tror den är i ett läge och därmed förväntar sig ett annat beteende än det som sker (Johnson, 2010). Benägenheten att hamna i lägesfel är större med dålig återkoppling om aktuellt läge. Ett kritiskt lägesfel kan uppstå vid exempelvis ett trafikljus och föraren lägger i backen i stället för första växeln. I stället för att accelerera framåt sker det omvända när denne trycker ner gaspedalen. I ett scenario med ett läsförståelsetest riskerar ett lägesfel knappast orsaka en olycka men kan däremot potentiellt skapa förvirring som påverkar resultaten.

3.2 Skissning

Det finns olika metoder för skissning som lämpar sig olika bra beroende på situationen de ska användas till. Scribble sketching är en metod för att snabbt skissa ner sina idéer och i dessa bara fånga essensen av det och undvika detaljer (Buxton, 2010). Metoden lämpar sig därför att på stående fot inom loppet av några sekunder få ner idéerna då de kommer. Det är också en väl fungerande metod då många idéer ska genereras. Att ta fram en mängd olika skisser med vitt skilda tankar är ofta en bra början för att finna en god utgångspunkt. Som Buxton beskriver kan man illustrera problemet som ett antal kullar där bara en av dem kan ha den högsta toppen/bästa idéen. Inom datorvetenskap finns för detta den engelska termen *local hill climbing*. Designen kan sedan bara vara så god som den aktuella kullen tillåter. Därför är det ofta bra att ta fram vitt skilda idéer innan någon ska utvecklas ytterligare. Brainstorming, diskussioner och observationer av slutanvändaren är några av de vägar som kan utnyttjas för att lyfta fram ytterligare idéer. Laseau (2001) menar för att ta sig förbi lokalt maximum inleds en divergent fas som sedan övergår mot en konvergent fas (se Figur 3.1). Det innebär att man först förfinar en mängd idéer i sin repetoar för att sedan reducera antalet idéer tills den mest lovande är funnen.

Buxton beskriver en ytterligare skissmetod, 10-plus-10-metoden, för att bättre understödja brainstorming. Först definieras problemet innan tio eller fler idéer sedan ska produceras. Målet är att de ska vara så kreativa och varierande som möjligt då det bara är initiala koncept. Att medvetet tvinga upp tempot kan medföra att många spärrar försvinner som annars hade begränsat idérikedomen. När skisserna är klara inspekteras de och de utan potentiell bärkraft läggs åt sidan. För övriga skisserna kan de vidareutvecklas så de kan visas och förstås av andra. De mest lovande kan därefter fortsätta att raffinerars ytterligare.



Figur 3.1: Repertoaren av idéer utvecklas alltjämt som antalet idéer reduceras för att så slutligen finna den mest lovande.

3.3 Skärmläsning

Att vi läser fler digitala dokument idag än för tio år sedan råder ingen tvekan om, men implikationerna för detta är något som (Liu, 2005) tagit fasta på i sin artikel. Essensen av vad som framkommer i studien är att mer tid läggs på bläddrande och scannande, sökning efter nyckelord. Läsningen av en text sker ofta bara en gång och på ett ickelinjärt vis. Läsningen är dessutom mer selektiv. Större informationmängder från fler källor har medfört att vi tvingas bli mer kompetenta på att sälla i materialet. Samtidigt som läsningen blivit mer ytlig har också den ihållande uppmärksamheten blivit lidande. Läsande från papper förknippas oftare med aktiviteten att stryka under och notera än motsvarande i den digitala miljön enligt (Liu, 2005).

Bläddra och skanna ett fysiskt dokument behöver inte vara enbart för att lokalisera information i texten men kan också vara ett sätt att skapa en känsla för texten som helhet. Att bläddra på en datorskärm ger inte stöd för den här typen av läsning och informationsprocessande. Undermedvetet och automatiskt försöker hjärnan skapa ett visuellt minne för var i texten olika objekt finns på en sida i ett dokument. Scrollandet på en dynamisk datorskärm försämrar den relationen (Olsen, 1994).

3.4 Skärmtypografi

Utöver rent lingvistiska aspekter avgörs en texts läsbarhet bland annat av dess typografi. Det inkluderar variabler som typsnitt, typgrad (storlek), radlängd, kolumner, radavstånd med mera. Många av variablerna korrelerar också med varandra och kan således ej isoleras till att alltid ha ett statistiskt värde oberoende av varandra. Exempelvis kräver längre rader också ett större radavstånd för att ögonen korrekt ska kunna lokalisera nästa rad efter slutet på aktuell. Vinkeln bör därav inte vara för liten (Bouma, 1980). Linjelängd kan dessutom mätas på två sätt; antingen i fysisk längd av linjen eller i antal tecken på raden, vilket kan förbrylla ytterligare.

Chaparro et al. (2001) undersökte hur teckenstorlek påverkar läsande från skärm. I experimentet jämfördes texter i punktstorlekarna 10, 12 och 14 med totalt nio olika typsnitt. Testdeltagarna var alla vuxna och majoriteten hade vana att läsa dokument från skärm. Vad man fann vara att text med storlek 10 tog längre tid att läsa än motsvarande med storlek 12. Men det är också en avvägning/kompromiss mellan hastighet och noggrannhet. Teckensnitt som lästes snabbare var också generellt mindre precisa. Upplevd läsbarhet bestämdes inte enskilt utav storlek eller teckensnitt men i kombination de bägge emellan. Bernard et al. (2002) jämförde preferensen mellan tre olika radlängder på 132, 76 respektive 45 tecken per rad. Detta undersöktes för såväl vuxna som för barn utan att konstatera några signifikanta skillnader i vare sig effektivitet eller tid för läsningen. Däremot påvisas att kortare längder generellt föredras av vuxna över fullängds som sträcker sig över hela skärmen. Med det smalaste radspannet upplevdes det

som enklare att upprätthålla koncentrationen men föredrog ändå den medellånga för längdläsning. Barnen däremot uppfattade inga skillnader i egen effektivitet gällande läsning men indikerade likt de vuxna en preferens till smalare rader. Bernard et al. resonerar att det möjligen kan förklaras av det faktum att barn i det undersökta åldersspannet (9-12) inte har fullt utvecklad läsförmåga och därför koncentrerar sig mer på läsningen självt för att linjelängd ska spela någon roll.

3.5 Färgseende

Som noterat i avsnitt 1.3 lider ungefär 8% av alla män i Nordamerika och Europa av någon form av defekt färgsinne. Hos kvinnorna är motsvarande siffra 1% av kvinnorna (MacDonald, 1999). Att någon är fullständigt färgblind är mycket sällsynt. Vanligast är svårigheten att skilja på rött och grönt. Vissa färgscheman upplevs därför mer svårlästa och mindre lämpliga än andra då färgerna kan verka flyta ihop. Kontrasten av mörkare mot ljusare är inget som berörs av någon bland den här typen av nedsättningar varför det också är en bra ledstjärna för val av färger för kritiska områden.

Kapitel 4

Metod

I denna studie kombinerades flera metoder för att utveckla en prototyp. Efter litteraturstudier har egna skisser stått till grund för implementeringen av programvaran. Den har därefter genomgått pilottest i fält under observation och med uppföljande gruppintervjuer. Efter återkoppling från eleverna som genomgått testet har mjukvaran och testförfarandet justerats för att sedan köras skarpt på åtta elever från fjärde klass samt lika många från åttonde klass. Även dessa har observerats under testfasen och sedan deltagit i gruppintervjuer. Metoderna kompletterar varandra väl utan att inkräkta på respektive område. Var metod diskuteras mer ingående nedan.

4.1 Design

Primärt har 10-plus-10-metoden används för att generera en mängd skisser. Vissa har därefter arbetats vidare på och andra har förkastats. Idéerna har innan initieringen av första prototypen diskuterats med några studenter som alla läst minst en kurs interaktionsdesign. De flesta av de skisser som förkastats har gjort det av anledning att de sannolikt blir för omständliga att använda i kontrast till förfarandet på papper. Det är viktigt att betänka att verktyget senare kommer vara något som inte används frekvent av eleverna och bör fungera utan någon inlärningströskel.

För prototypen har två designförslag tagits fram i enlighet med vad som illustreras med Figur 4.1 och Figur 4.2. De nyttjar den digitala ytan på två skilda sätt men är båda minimalistiska då funktionalitet annat än strikt nödvändigt uteslutits. Högst upp på skärmen finns webbläsarens navigationspanel och under denna finns en förloppsindikator. Denna indikerar framskridandet i testets fyra faser som består av två ordtest och två lästest. Eftersom ordtesten i bägge fallen är oberoende av någon text är dessa identiska för båda lösningarna. Den första (Se Figur 4.1) nyttjar skärmstorleken närmast till max och presenterar texten över större delen av skärmen. Utrymmet längst ner occuperas av en indikator där prickar symboliserar att

fler sidor finns tillgängliga likt en iPads hemskärm. Vid svep till vänster exponeras frågorna för användaren och vid svep tillbaka visas texten åter. I den andra designen (Se Figur 4.2) figurerar text och frågor sida vid sida och delar därmed skärmens yta på mitten. För att komma till nästa fas i testet klickar man på knappen som är i botten av frågeformuläret. Förfarandet är detsamma i de båda lösningarna.

Båda designerna har varit öppna för att användas såväl i stående som liggande läge, vilket elverna också uppmanas till att utforska i samband med inloggningen till testet. Däremot har möjligheten till att zooma låsts i HTML-filen vilket besparar användaren från flera potentiella problem som kan orsaka konflikter i testandet.

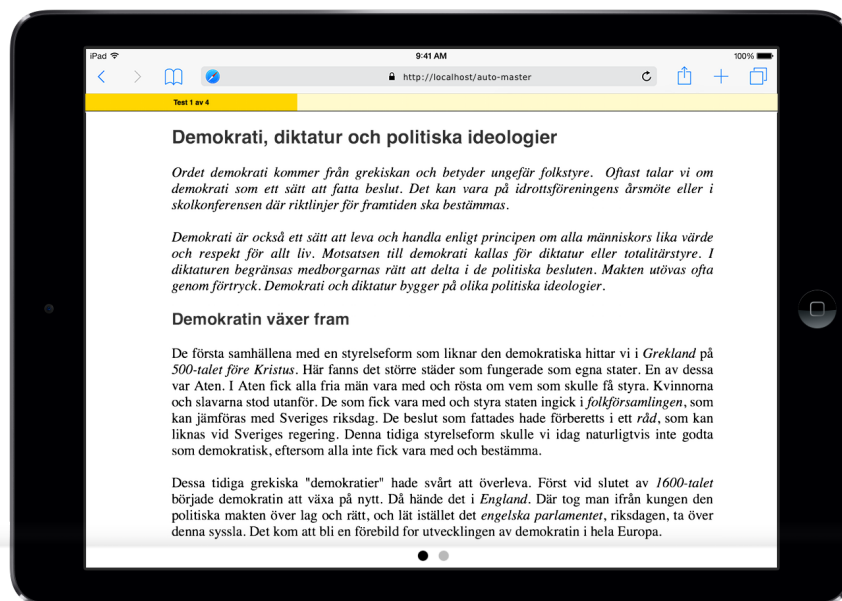
Efter en pilotstudie inkluderades bland annat mjukare scrollning som medför att det upplevs mer naturligt att scrolla igenom texten och färre gester krävdes för att scrolla större mängder text. Det var något som innan noterades vara ett problem även om det aldrig explicit uttalats från eleverna.

För att tillgodose behoven för personer med nedsättningar i färgseendet kontrollerades designen med <http://www.color-blindness.com>. Detta verktyg tillåter olika typer av färgnedsättningar att emuleras för att granskas av personer med normalt seende. Därigenom kan eventuellt bristfälliga designval i avseende färger exponeras.

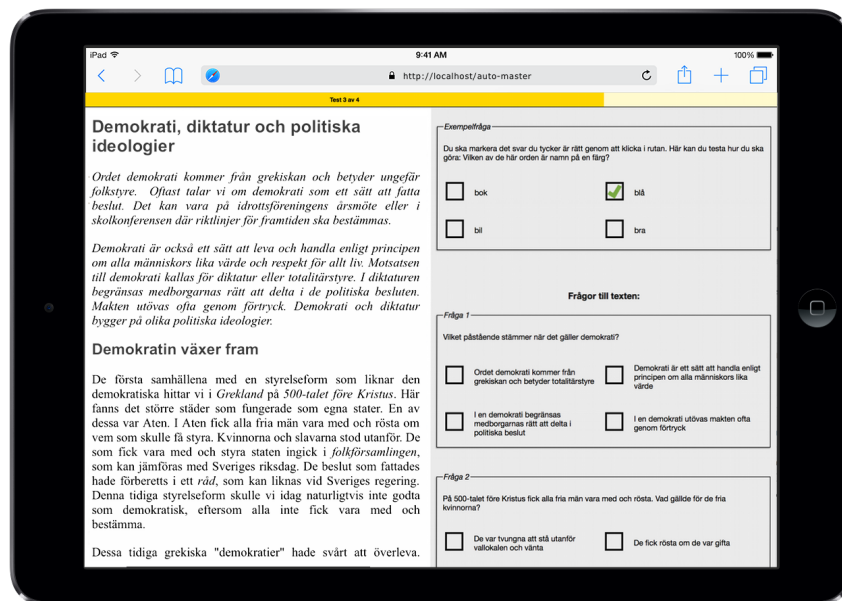
Typografin i textmängderna har på önskemål av involverade forskare i projektet MASTER efterliknat de ursprungliga dokument som tagits fram för testet. Deras avsikt har varit att ligga så nära verkliga originalkällor som möjligt där det i vissa fall kan vara ett ca 30 tecken brett urklipp från en tidningsartikel för i nästa fall vara en fullängds rad med utdrag från en lärobok. Därför har typsnitt, teckenstorlek, radavstånd etc. varierat även i samma texter. Standard för majoriteten av texterna har utgått från 14 punkter med 1,15 radavstånd för årskurs fyra och 12 punkter med 1,0 radavstånd för åttan. Avsteg från att helt efterlikna malldokumenten har varit oundvikligt i betingelsen sida vid sida där texten blivit för smal för att helt efterlikna originalet. Där har därför har bland annat radhöjd anpassats till den kortare radlängden enligt teorier för detta. Radhöjden har varit något högre för texten i liggande läge då raderna blir något längre än vid stående.

4.2 Teknisk implementation

Från projektet MASTER fanns en befintlig webbaserad miljö tillgänglig för frågeformulären där de elever med tillgång till datorer kunnat logga in och besvara dessa på skärm. Servern som arbetats mot har funnits lokaliserad på Linköpings Universitet. Databasen är av typen MySQL och ligger på en Apache-server. JavaServer Pages genererar dynamiska sidor baserat på HTML, CSS samt JavaScript. Motsvarande komponenter inklusive en kopia av databasen installerades på en separat dator som fick agera server (Se Figur 4.3). Den utgjorde en testmiljö som därför kunde ändras utan



Figur 4.1: Illustration av det koncept som ligger närmast hur originaltexten ser ut på papper.



Figur 4.2: Illustration av det koncept där text och frågor konkurrerar om samma yta.

inverka på befintlig data eller test. Servern togs med till respektive skola för test så att surplattorna skulle kunna koppla upp lokalt mot den. I databasen upprättades nya användarnamn och lösenord för att kunna spara och analysera elevernas resultat.

För svepfunktionaliteten har öppen källkod använts baserad på JavaScript. De figurer i webbsidans nederkant som symboliserar tillgängliga sidor



Figur 4.3: Översiktsbild på komponenter nödvändiga för körning av systemet.

har flyttat på sig beroende på vilket läge surfplattan varit i, vilket därför har korrigerats med CSS så att de alltid återfinns på samma position oavsett läge. För den andra betingelsen har sidan konfigurerats så att vänster respektive höger sida har haft oberoende scroll. Har texten i vänster del scrollats har frågorna till höger ändå bibehållit tidigare position och vice versa. Flera webbt tekniker har nyttjats för att skapa ett så följsamt gränssnitt som möjligt.

4.3 Datainsamling

För att göra test med elever från fjärdeklass kontaktades en utvecklingsstrateg på Norrköpings kommun. E-post skickades på dennes vägnar ut till rektorerna i kommunens respektive mellanstadieskolor, varefter den första som svarade positivt till deltagande valdes. I kontakt med klassföreståndaren förklarades proceduren för testet, först skriftligt per e-post och sedan mer genomgående över telefon. Etiska ställningstagande diskuterades, och huruvida de önskade skicka ut information till föräldrarna om testet. Efter en intern dialog lärarna emellan ansågs det överflödigt av flera orsaker. Dels är testet anonymiserat för utomstående, dvs endast lärarna kan koppla resultaten till respektive individ som endast blivit tilldelad ett numrerat inlogg. Tiden var dessutom knapp från kontakt till avtalat datum för test. Testet ansågs också vara av lärandeform med legitim bakgrund. Enligt utsago var eleverna dessutom mycket vana vid olika typer av externa test och det är också något föräldrarna är medvetna om förekommer.

Då ovan nämnd skola endast sträckte sig upp till femte klass och andra testet skulle utföras på åttondeklass togs kontakt med en skola i Linköping. De mottog på önskemål omfattande information om testet och meddelade att ingen upplysning var nödvändig att förmedla föräldrar då testet är anonymt för oss som sparar datan om eleverna.

Under testen observerades hur interaktionen förlöpte mellan deltagare och surfplattan i syfte att kunna skönja eventuella problem eller lyfta fram övrigt värt att reflektera över. Till stöd användes ett observationsschema (se appendix) som tagits fram på förhand av testet. Det hindrade givetvis inte andra observationer från att noteras. I roll av försöksledare valdes att inta en till synes passiv roll så att eleverna inte skulle känna sig iakttagna och pre-

stera annorlunda av den anledningen. Halvvägs in i testet fick eleverna säga ifrån och vänta in de övriga. Därefter ändras betingelsen till den andra. För att höja validiteten alterneras startbetingelsen mellan varannan testgrupp. De två betingelserna avser alltså de två designförslag som utvärderas mot varandra.

Efter genomförda tester följde semistrukturerade intervjuer i syfte att bland annat finna svar på vilken och varför en design föredrogs över den andra. Intervjuerna skedde i gruppform om fyra elever för att på så vis skapa en trygghet för eleverna att våga uttrycka sig. Underlaget med frågor var främst som stöd och följdes inte nödvändigtvis till fullo. Då frågor bara besvarades med ja eller nej ställdes följdfrågor om *varför* för att få mer uttömmande svar.

Val av metoder föll på observationer samt intervjuer då de båda är väl beprövade och i aktuell studie ansågs kunna komplettera varandra genom att fånga upp olika aspekter av testandet. Observationerna kunde ske ostört och passivt medan intervjuerna aktivt inkluderade eleverna i tänkandet. Videoupptagning övervägdes för mer precis analys men förkastades av etiska och praktiska orsaker.

4.3.1 Pilot

Primära syftet med pilottestet var att undanröja eventuella tveksamheter och buggar i mjukvaran samt få lite vana av att instruera och observera under försöken. Den bestod av fyra elever i årskurs fyra. Den utfördes under en dag på vald mellanstadieskola i Norrköping. Eleverna valdes slumpmässigt av klassföreståndaren på önskemål av testledaren. Två elever åt gången fick sitta avskilt och ostört från övriga klassen i ett grupprum och genomföra testet. Korta instruktioner om vad testet gick ut på förmedlades muntligt varefter eleverna fick lov att börja utföra testet i egen takt. I roll av testledare uppmuntrades även att påpeka tveksamheter men att inga svar kunde delges. Testet skedde vid ett runt bord. Detta gav bättre överblick och medförde enklare observation för testledaren.

Pilotstudien resulterade i mindre korrigeringar av gränssnittet samt rättning av några direkta fel. Bland annat togs överflödiga information i gränssnittets översta del bort för att ge mer utrymme för det väsentliga. Frågor med fler identiska svarsalternativ förekom också. Dessa fel rättades inför den efterföljande huvudstudien.

4.3.2 Huvudstudie

Samma förfarande som för piloten valdes för de skarpa testen. Men i kontrast till pilottestets grupper om två elever skedde det skarpa testet i grupp om fyra elever åt gången. Detta valdes för att på så vis mer effektivt genomföra testen då fler surfplattor fanns tillgängligt samt att observationsfokus inte längre var lika spridd mellan lika många olika punkter som vid det inledande

pilottestet. Eleverna från årskurs fyra satt alla runt det runda bordet. Eleverna från årskurs åtta satt under testet vid var sin bänk vid främsta raderna riktade mot katedern i ett mindre klassrum från var experimentledaren hade översikt.

4.3.3 Analys

Observation- samt intervjumaterial har efter pilotstudien analyserats för att kunna lyfta fram viktiga aspekter för att därefter kunna justera framtagna prototyper. Detta har främst skett genom att klustra samman material och på så vis finna återkommande teman. Enstaka aspekter har även ansetts viktigt nog att överväga även om det bara skett vid enstaka tillfällen.

Kapitel 5

Resultat

I detta kapitel redovisas data från genomförda observationer, intervjuer samt testresultat. Observationerna har medfört en mängd viktiga insikter kring användandet av surfplattan som åskådliggörs i listan nedan. Efter dessa följer data från intervjuerna där essensen av yttranden redovisas tillsammans med mer specifika citat. Sist presenteras elevernas prefererade betingelse samt snittresultat per text och betingelse.

5.1 Observationer

- Samtliga, med undantag av en från årskurs fyra samt en från årskurs åtta använder surfplattan enbart i liggande läge. Ett fåtal av de i fjärde klass testade att vrida skärmen till stående läge för att sedan snabbt vrida tillbaka den igen innan vidare användning.
- Skärmdimmning går igång på en enhet när eleven läser statiskt utan att fysiskt interagera med skärmen under någon minut.
- Samtliga elever tycks förstå hur båda betingelserna navigeras i. En åttåndeklassare insåg det först när all text var läst att det gick att svepa åt sidan för att exponera tillhörande frågor.
- Två elever, en från årskurs fyra och en från årskurs åtta observerades att av misstag svepa från kanten i stället för en bit in och därmed kommit åt webbläsarens, Safari, tillbaka-funktion och hamnat vid login-skärmen. Bägge under betingelsen svep.
- En elev råkar öppna ny flik och undrar hur man kommer tillbaka till testet.
- Flera elever uttryckte förundran efter att ha valt ett alternativ och inte kunde välja bort det igen då de informerats att de inte måste svara om de känner sig osäkra.

- Betingelsen svepning tycks resultera i mycket växlande fram och tillbaka även på samma fråga.
- Ett par elever från fyran försöker börja svara på frågorna direkt utan att ha läst igenom hela texten först.

5.2 Intervjuer

- Alla förstod vad som skulle göras och hur navigeringen fungerade.
- Ingen var helt ny för konceptet surfplattor eller motsvarande teknik.
- Utöver de två observerade eleverna som råkat backa ut från testet var det ytterligare en men som på egen hand förstod att det bara var att klicka fram i webbläsarens igen.
- Ingen av de tillfrågade hade föredragit att göra testet på papper.
- Ingen gav explicit uttryck för behov av paus vid frågan.
- Kryssaboxarna buggade på en av surfplattorna, annars var det inga konstigheter.
- På frågan om vilken betingelse som föredrogs var fyror nästan helt samstämmiga, sånär som på två elever som föredrog svep-betingelsen över den andra utan någon speciell anledning. Första gruppen åttor svarade instämmande att de föredrog svep efter att en elev yttrade följande:

“man läser inte lika grundligt sida vid sida”.

Nästa elev flikte in med:

“Frågar du mig senare så kommer jag nog ihåg mer från första delen [med svepning] eftersom då läste man först allt och sen svarade på frågorna.”

Första eleven reviderade sitt svar och sa:

“Den andra [sida vid sida] var skönare eftersom man kunde läsa frågorna och leta efter svaren.”

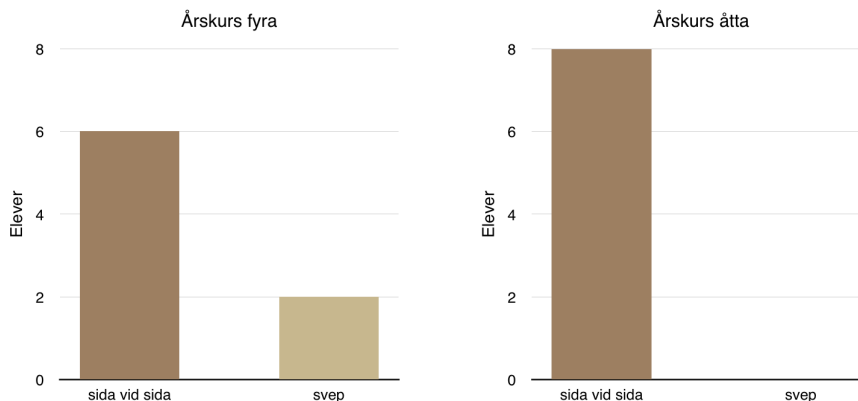
En tredje elev uttryckte ändå att denne tyckte sida vid sida verkade lättare, och fick sedan medhåll av övriga tre i gruppen [underförstått att den ena designen kändes bättre ur lärande synpunkt och den andra för att bara klara besvara frågorna]. Den sista gruppen från årskurs åtta var alla enstämmiga om att de tyckte sida vid sida var enklare och föredrog den. En av eleverna tyckte det blev för mycket svepande fram och tillbaka annars. En av de andra sa:

“Det är enklare att se frågan och kunna leta efter svaret vid sidan”

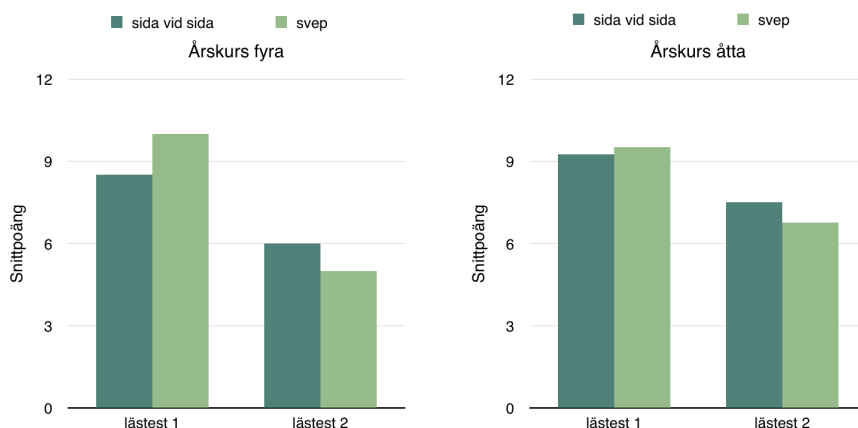
Av de tillfrågade var det ingen som sades sakna möjligheter att anteckna, även om några sa att de inte visste riktigt.

5.3 Testpoäng och preferenser

I Figur 5.1 visas föredragen betingelse från fjärde respektive åttonde klass. Svaren avser vad eleven anser är enklast för att klara uppgiften. I Figur 5.2 visas snittresultaten från fjärde respektive åttonde klass. Varje stapel representerar alltså fyra elever då det är antalet per betingelse.



Figur 5.1: Föredragen betingelse för årskurs fyra till vänster respektive årskurs åtta till höger.



Figur 5.2: Snittresultat för årskurs fyra till vänster respektive årskurs åtta till höger fördelat på de båda betingelserna

Kapitel 6

Diskussion

I detta kapitel diskuteras resultat från studien och relateras till tidigare studier. Dessutom diskuteras metodval, studiens validitet och slutligen om egna reflektioner om studien.

6.1 Resultatdiskussion

Resultatet tyder på att elever föredrar att ha text och tillhörande frågor fördelat på samma yta över alternativet att via en handling växla de båda emellan. Dock har enskilda elever gett uttryck för att text över hela skärmytan kan vara föredraget ur lärandesynpunkt då de upplever sig läsa mindre noggrant och mer aktivt leta efter specifika svar när de har frågorna vid sidan. Då T-MASTER undersöker förmågan att läsa och förstå och inte är fokuserad på lärande i den bemärkelsen anses därför sida vid sida vara en bättre lösning då 18 av 20 elever, pilotstudien inkluderad, föredragit detta. Värt att notera är också att de två eleverna från årskurs fyra föredrog det andra alternativet gjorde det först då frågan ställdes på sin spets och ett val var tvunget att göras. Ursprungligen uppgav de att det inte spelade någon roll. Alternativet med text och frågor sida vid sida är också att föredra ur ett programmeringsmässigt perspektiv då det är mindre kod, lättare att anpassa till olika skärmstorlekar och plattformar samt att problematiken med att av misstag svepa bakåt i webbläsaren därigenom undviks. I resultaten anas en svag tendens till högre prestation med betingelsen svep på de enklare texterna och sida vid sida på de svårare texterna. Detta gäller för såväl årskurs fyra som åtta. Alltså ges inga tydliga indikationer på att någon av utformningarna är överlägsen oberoende textunderlag. Text och frågor på samma sida är rimligtvis mindre kognitivt belastande då det endast krävs en blick för att växla mellan de båda. Då den längre och svårare texten av naturliga skäl kan tänkas kräva mer mentala resurser av läsaren kan det därav vara fördelaktigt att inte behöva lägga svarsalternativen på minnet i samma utsträckning som krävs vid svepningar. Den enklare texten kan eventuellt

tänkas hålla en nivå där i stället översiktligheten blir en fördelaktig faktor för prestationen. Huruvida resultaten håller med större underlag återstår ännu att se.

6.2 Metodval och genomförande

Skissandet för de båda betingelserna har skiljt sig något åt. I den där skärmen i stort sett bara ockuperas av text har möjligheter att låta idéerna sväva ut varit mer begränsade. Den andra har genererat fler idéer för även om valet föll för en av de minst funktionsrika men däremot en synnerligen enkel lösning. Befintliga lösningar för att presentera text i olika form så som iBooks, Kindle och enskilda läroverktyg har givit som inspiration till vissa idéer. Men som Laseau (2001) nämner så finns alltid möjligheten att hamna i en sits där bara lokalt maximum kan nås och inte den globalt optimala lösningen. Det är inte möjligt att med säkerhet konstatera man funnit den mest optimala lösningen då antalet varianter i det närmsta är obegränsade i antal. Som tidigare nämnts har de flesta skisser förkastats då det har varit ett starkt fokus att hålla gränssnittet enkelt ur användarsynpunkt.

Inför testen togs ett observationsschema fram för att kunna lägga särskilt fokus kring vissa på förhand intressanta punkter. En fördel med observationer är att det tillåter att upptäcka sådant som sträcker sig bortom bara förväntade skeenden. Flera intressanta aspekter under testet har framkommit med denna metod. Hade eleverna lämnats ensamma under testets gång hade risken varit stor att även de egna spontana kommentarerna och frågorna de haft inte framkommit i efterhand utan fallit i glömska. Inte allt för oväntat har det varit utmanande att i intervjuerna lyfta fram vad eleverna själva önskar utöver vad de ser och själva testat. Det är svårt att resonera om sådant man inte vet kan finnas. Generellt var intervjuformen mest givande för de äldre eleverna som med större lätthet kunde reflektera över frågorna som ställdes.

Stora delar av den kod som kommunicerar med databasen och producerar frågorna fanns tillgänglig för uppsatsen. Att inkorporera ny kod med befintlig har i viss mån begränsat designen något och även orsakat kodmässiga krockar som sannolikt undvikits om koden skapats helt från början. Den bygger dock på de webbtekniker som oavsett hade valts för implementering av skisserna i uppsatsen. Detta lämpar sig väl då läsförståelseverktyget med nyttjade webbspråk kan göras plattformsoberoende.

6.3 Validitet och generaliserbarhet

Uppsatsen kan i sitt omfång peka på tendenser men kräver för att statistiskt förankra resultaten ytterligare data. Demografiska skillnader är också något att betrakta. Att stora olikheter existerar mellan skolor råder ingen tvekan om. Hur de uttrycks och huruvida de också återspeglas i testet är

svårt att sja om. Utöver resursmässiga skäl har också testen begränsats till elever från endast två skolor. Det har varit ett önskemål att i minsta möjliga mån påverka regionens alla skolor som ligger på tur för skarpt test av T-MASTER. Att elever valdes från åttonde i stället för nionde årskurs berodde på att åttorna inte ligger primärt på tur att köra skarpt. Annars hade elever från nionde årskursen som ligger högst i åldersspannet för T-MASTER föredragits. Detta då elevernas lässkicklighet ständigt utvecklas och eventuella skillnader därmed rimligen är störst när åldersspannet sträcker sig från yngst till äldst.

Under testet med en av grupperna i fjärde klass var en av eleverna väldigt stökig och störde de övriga eleverna när dennes fokus sjönk. Har inte kunnat se något tydligt tecken i resultaten att det skulle ha inverkat på de övriga individerna i gruppen. Däremot antas eleverna kunna påverka varandra under intervjufasen. De har varit väldigt samstämmiga vid olika yttranden trots att de mellan olika grupper kunnat spreta något. Spontant kan enskilda intervjuer tyckas som en given lösning. Men ska många elever testas så kan det av praktiska skäl vara fördelaktigt att i stället nyttja ett digitalt formulär som eleven själv kan fylla i. Förslagsvis antingen i form av en SUS-enkät (System Usability Scale) som består av tio frågor med fem möjliga svar eller ett formulär med öppna svarsalternativ om man önskar ett mer kvalitativt angreppssätt.

En av de två elever som föredrog betingning svep var en av de som också använt surfplattan stående. Det kan tyckas förstaeligt då det i stående läge blir mycket begränsat utrymme om text och frågor ska konkurrera om utrymmet. Den stora majoriteten av eleverna gjorde inga försök att använda surfplattan på annat vis än i det läge den var i när testet inleddes, med andra ord liggande läge. De flesta av surfplattorna hade även fodral av olika slag som uppmuntrade detta läge. Det är svårt att sja om utfallet sett annorlunda ut om testet preparerats med surfplattorna i stående läge och utan fodral som inbjuder till liggande användning.

Många av de teorier om typologi som initialt tänkts att inkluderas i koden fick strykas för att matcha önskemål från projektgruppen att efterlikna verkliga texter. Oavsett skäl skiljer sig texterna bara i måttlig omfattning typologiskt i de olika betingelserna. Men i enlighet med bland annat Bernard et al. (2002) så bör effektiviteten oavsett inte påverkas nämnvärt av radlängden. Tvärtom föredrar många en kortare radlängd och upplever att det kan bidra till att fokus kan upprätthållas längre. Noteras bör dock att gällande läsbarhet av texter finns det många studier som tycks peka i något olika riktning. En möjlig förklaring kan som nämndes i teoriavsnittet var att det kan bero på att olika kombinationer av variabler kan orsaka olika effekter. Många av de artiklar som initialt har hittats på ämnet specifikt om läsning från skärm har förkastats då de i många fall ansetts varit utdaterade då tekniken gått framåt på och ändrat förutsättningarna på många vis. Skärmarna har blivit mer högupplösta, fått bättre färger, betraktningssvinklar etc.

6.4 Framtida studier

Vilket designval som föredras av eleverna tycks i det närmsta vara uppenbart. Däremot behövs bättre underlag för att kunna konstatera att det inte säkert inverkar på resultatet i läsförståelsetestet. Rimligen begränsas då också fler potentiellt påverkande variabler genom att exempelvis utesluta stående läge. Vidare vore det intressant att sätta designerna i relation till elevers prestation på papper. Ur flera perspektiv är det föredraget att ha en genomgående konsekvent utformning oavsett plattform. Men kan det få några implikationer för testresultatet?

Några av de yngre barnen visade tendenser till ovilja att läsa texterna, i synnerhet den andra texten. Kan ett mer lekfullt gränssnitt uppmana till bättre fokus och större engagemang? Eftersom testerna i T-MASTER finns för flera olika kärnämnen är det av stort intresse att veta hur man kan förmå eleverna att vilja ta sig igenom samtliga test.

Kapitel 7

Slutsats

En klar majoritet med 18 av 20 elever föredrar en design på läsförståelsetestet där text och frågor figurerar på samma yta av skärmen i kontrast till de 2 elever som i studien föredrar alternativet att växla mellan visning av bara text eller bara frågor. Observationerna visar att mycket få är benägna att använda surfplattan i stående läge under testet. Med den insikten kan anpassning för detta läget övervägas att uteslutas och därigenom nå ett mer konsekvent gränssnitt. Några elever resonerar att när texten ensamt fyller skärmen är det fördelaktigt ur lärandesynpunkt då den i det läget läses mer noggrant. Inga tydliga skillnader kan skönjas i resultaten de båda designerna emellan. Detta bör dock styrkas kvantitativt med ett större urval.

Litteraturförteckning

- Bernard, M., Fernandez, M., and Hull, S. (2002). The effects of line length on children and adults' online reading performance. *Usability News*, 4(2).
- Bodmann, S. M. and Robinson, D. H. (2004). Speed and performance differences among computer-based and paper-pencil tests. *Journal of Educational Computing Research*, 31(1):51–60.
- Bouma, H. (1980). Visual reading processes and the quality of text displays. *IPO Annual Progress Report*, 15:83–90.
- Buxton, B. (2010). *Sketching user experiences: getting the design right and the right design: getting the design right and the right design*. Morgan Kaufmann.
- Chaparro, B., Bernard, M., Mills, M., Peterson, M., and Storrer, K. (2001). A comparison of popular online fonts: Which is best and when. *Usability News, Software Usability Research Laboratory*, 3(2):50–60.
- Garland, K. J. and Noyes, J. M. (2004). Crt monitors: Do they interfere with learning? *Behaviour & Information Technology*, 23(1):43–52.
- Heimann Mühlenbock, K. (2013). *I see what you mean. Assessing readability for specific target groups*. Dissertation, Språkbanken, Dept of Swedish, University of Gothenburg.
- Johnson, J. (2010). *Designing with the mind in mind: Simple guide to understanding user interface design rules*. Morgan Kaufmann.
- Johnson, J. (2013). Mobile first design: Why it's great and why it sucks.
- Laseau, P. (2001). *Graphic thinking for architects & designers*. John Wiley & Sons.
- Liberg, C. (2010). Texters, textuppgifternas och undervisningens betydelse för elevers läsförståelse: Fördjupad analys av PIRLS 2006.
- Liu, Z. (2005). Reading behavior in the digital environment: Changes in reading behavior over the past ten years. *Journal of documentation*, 61(6):700–712.

- MacDonald, L. W. (1999). Using color effectively in computer graphics. *Computer Graphics and Applications, IEEE*, 19(4):20–35.
- Mayes, D. K., Sims, V. K., and Koonce, J. M. (2001). Comprehension and workload differences for vdt and paper-based reading. *International Journal of Industrial Ergonomics*, 28(6):367–378.
- Mühlenbock, K. H., Kanebrant, E., Kokkinakis, S. J., Arne Jönsson, C. L., af Geijerstam, A. s., Falkenjack, J., and Folkeryd, J. W. (2014). Studies on automatic assessment of students’ reading ability. In *Proceedings of the Fifth Swedish Language Technology Conference (SLTC-14)*, Uppsala, Sweden.
- Olsen, J. (1994). *Electronic journal literature: implications for scholars*. Meckler Corporation.
- Reichenberg, M. (2000). *Röst och kausalitet i lärobokstexter: en studie av elevers förståelse av olika textversioner*. PhD thesis, Acta Universitatis Gothoburgensis, Diss. Göteborg : Univ.,Göteborg.
- Wästlund, E., Reinikka, H., Norlander, T., and Archer, T. (2005). Effects of vdt and paper presentation on consumption and production of information: Psychological and physiological factors. *Computers in human behavior*, 21(2):377–394.

Appendix

Observationsschema

- Tycks de förstå vad de ska göra och hur de ska navigera?
- Används surfplattan liggande och/eller stående?
- Används surfplattan på något specifikt avstånd?
- Kommer de av misstag eller flit åt andra funktioner på surfplattan?
- Syns några tendenser till att fastna vid några frågor?
- Någon som ändrar ljusinställningar och annat?
- Försvinner fokus iväg på annat än testet?
- Orsakas några problem med scrollandet vid betingelsen svep?

Intervjuunderlag

- Förstods vad som skulle göras?
- Hur var det att läsa på surfplatta?
- Finns vana sedan innan av surfplattor?
- Fungerade det bra att scrolla?
- Hade anteckningsmöjligheter använts om de fanns?
- Några påstötta problem?
- Paus nödvändigt?
- Hur fungerade det att använda kryssalternativen?
- Tydligt hur långt man kommit i testet?
- Vilket föredras; sida vid sida eller svep?
- Hellre föredragit test på papper?