

DIKU Kursus Interaktionsdesign – Inter

Rolf Molich

# Opgave 6 og 7 Usabilitytest

## Overblik over videoforelæsninger

- ▶ Opgave 6+7  
Opgave 8
- ▶ Usabilityevaluering  
Planlægning af usabilitytest
- ▶ Udførelse af testsessioner
- ▶ Formidling af resultater  
Modenhed
- ▶ Kvantitativ usabilitytest

## Opgave 6+7 - Hovedformål

- ▶ Vise, at selv en primitiv test med begrænsede ressourcer kan finde interessante usabilityproblemer i en dialog
- ▶ Teste et produkt.  
Finde problemer i en brugergrænseflade
- ▶ Teste processen.  
Udføre en test under realistiske betingelser
- ▶ Forberede jer bedst muligt på eksamen.

## Årets testobjekter

<i>Kunde</i>	<i>Produkttype</i>	<i>Hold</i>
▶ GoDream.dk	Websted	2, 8
▶ Mouseflow.com	Websted	6, 7, 9
▶ Mobilbank, Danske Bank	App	3, 5
▶ Danish Muslim Aid	Websted	1, 4

## Oversigt over opgave 6

- ▶ Fastlæg testopgaver uden at se på webstedet hhv. produktet
- ▶ Løs jeres egne testopgaver på webstedet hhv. produktet
- ▶ Ret testopgaver ud fra jeres egne erfaringer
- ▶ Upload jeres drejebog til Absalon til kommentering

## Oversigt over opgave 7

- ▶ Udfør mindst fem testsessioner med typiske brugere på grundlag af jeres godkendte drejebog
- ▶ En bruger pr. testsession
  - En session med to brugere tæller som én testsession
  - I skal selv skaffe testdeltagere
  - I må gerne bruge de samme personer som i interviewopgaven
- ▶ Skriv rapport og aflever den på Absalon

## "Kunden"s rolle

### Kunden

- ▶ Har brug for jeres konkrete resultater
- ▶ Gør testen mere realistisk
- ▶ Besvarer spørgsmål om webstedet og testens formål.  
Kontaktoplysninger står i opgavehæftet. Cc til mig, tak.
- ▶ Stiller gaver til testdeltagere til rådighed
  - Jeg refunderer gaver til testdeltagere for op til i alt 300 kr.
  - Aflever originalkvitteringer påført gruppenummer til mig i pausen ved en forelæsning.
  - Angiv nummer til MobilePay.  
Hvis ingen gruppemedlemmer har MobilePay,  
så angiv reg.nr. og kontonummer.

## Beslægtede opgaver

- ▶ Opgave 9  
Sammenligning af usabilitytestresultater,  
mandag 20. marts
- ▶ Opgave 11  
Diskussion af usabilitytestresultater med kunde  
mandag 27. marts

Opgave 9 og 11 er obligatoriske  
og tilstedeværelse under øvelserne er obligatorisk for alle gruppemedlemmer

## I får udleveret ...

- ▶ Beskrivelse af opgave 6+7 i kursushæftet
- ▶ Navn på websted eller produkt
- ▶ Gaver  
Jeg refunderer gaver til testdeltagere for op til i alt 300 kr.  
De præcise regler står i kursushæftet
- ▶ Eksempel på testrapport (Kolding Kommune)  
Findes på Absalon > Files > Opgaver og miniopgaver
- ▶ Yderligere eksempler på testrapporter (Irma og Ryanair)  
Findes på Absalon > Files > Eksamen og reeksamen
- ▶ Eksempel på brev til testdeltager  
Findes på Absalon > Files > Opgaver og miniopgaver

## Jeg forventer fra jer ...

- ▶ Drejebog
  - Startbetingelser
  - Tjekliste til briefing
  - Tjekliste til interview før testsession
  - Testopgaver  
Selve opgaven, formål, startbetingelser, forventet svar
  - Tjekliste til interview efter testsession
- ▶ Testrapport

## Kommentarer til drejebog

- ▶ For at forebygge konsekvenser af alvorlige problemer i drejebogen, bør I aflevere jeres drejebog til Absalon som opgave 6
- ▶ Opgave 6 er individuel for at forberede jer bedst muligt på eksamen.  
Opgave 6 er ikke obligatorisk
- ▶ I kan forvente en tilbagemelding fra jeres instruktør inden for 48 timer, så I hurtigt kan komme videre
- ▶ Aflever drejebogen senest mandag den 6. marts kl. 23.55.
- ▶ Drejebøger afleveret efter dette tidspunkt kan ikke forventes kommenteret af instruktør

## Brugsvenlig testrapport

### Krav til

- ▶ Indhold
- ▶ Format
- ▶ Længde og størrelse - 35 sider
- ▶ Antal kommentarer: Mindst 20, højst 40
- ▶ Indhold af kommentarer: Nyttige, ingen gentagelser, forståelige

Jeg anbefaler, at I bruger samme testrapport format som i eksemplet på Absalon til denne opgave, men det er ikke noget krav.

## Sammenfatning

- ▶ Opgave 6+7: Usabilitytest med mindst 5 testdeltagere
- ▶ Skriv rapport, som er brugsvenlig og nyttig
- ▶ Drejebog, herunder testopgaver, afleveres senest mandag 6. marts.
- ▶ Afsluttende rapport afleveres senest fredag 17. marts
- ▶ Sammenligning af usabilitytestresultater: Mandag 20. marts i de sædvanlige øvelser. Obligatorisk fremmøde.
- ▶ Diskussion af testresultater med kunden: Mandag 27. marts, én time i tidsrummet mellem kl. 13 og kl. 17. Obligatorisk fremmøde.

DIKU Kursus Interaktionsdesign – Inter

Rolf Molich

## Opgave 8

### Studie af videnskabelig artikel

#### Opgave 8 - Hovedformål

- ▶ Formålet med denne opgave er at vise, hvordan en videnskabelig artikel om usability kan være opbygget.
- ▶ Formålet er også at give jer et indblik i en væsentlig problemstilling omkring usability.
- ▶ I skal læse artiklen og derefter besvare en række spørgsmål om den.
- ▶ Besvarelsen afleveres senest fredag 17. marts



Opgave 7 og 8  
har samme  
afleveringsdato:  
17. marts  
Rettidig omhu!

17

## Artiklen

Artiklen er

- ▶ Kami Vaniea, Yasmeen Rashidi  
*Tales of Software Updates: The process of updating software*  
Proceedings of CHI'16, May 07 - 12, 2016, San Jose, CA, USA

I kan finde artiklen i Absalon > Files > Opgaver og miniopgaver.

## Artiklens indhold

Forfatterne har bedt 307 personer fortælle om deres erfaringer med softwareopdateringer.

Forfatterne har bearbejdet disse erfaringer og opstillet en model for softwareopdateringer.

Jeg har kort præsenteret et af resultaterne af undersøgelsen som et godt eksempel på en model ved forelæsningen den 8. februar.

## Artiklens opbygning

- ▶ Abstract (sammenfatning af indhold)
- ▶ Introduction (tidligere resultater, som artiklen bygger på)
- ▶ Background (problemstillingen)
- ▶ Method
- ▶ Findings
- ▶ Discussion (selvkritisk analyse)
- ▶ Conclusion
- ▶ Limitations
- ▶ Acknowledgments
- ▶ References

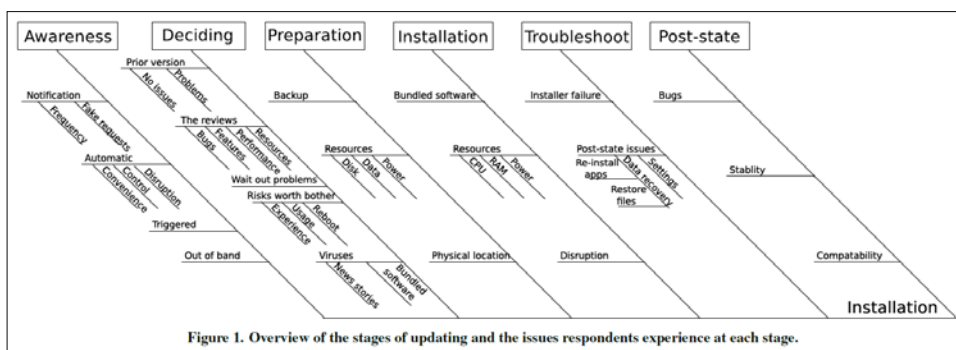
## Læsevejledning

- ▶ Start med at danne jer et overblik over artiklen ved at læse afsnittet *Abstract*.
- ▶ Dan jer derefter et overblik over artiklen ved at se den løseligt igennem – bemærk især overskrifter, tabeller og figurer.
- ▶ Læs *Introduction* og derefter *Discussion and Conclusions*.
- ▶ Læs derefter hele artiklen og besvar 6 delopgaver.

Artiklen er på 10 tætbeskrevne sider plus 2 sider litteraturhenvisninger

## Delopgave 1-3

- ▶ Artiklen bygger på en spørgeskemaundersøgelse på Amazon Mechanical Turk. Beskriv med jeres egne ord, hvad Amazon Mechanical Turk er.
- ▶ Skriv et resumé på højst 200 ord, hvor I sammenfatter problemstillingen, fremgangsmåden og de vigtigste resultater.
- ▶ Hvordan når forfatterne frem til figur 1?



## Delopgaver

- ▶ Artiklen bygger på en spørgeskemaundersøgelse på Amazon Mechanical Turk. Beskriv med jeres egne ord, hvad Amazon Mechanical Turk er.
- ▶ Skriv et resumé på højst 200 ord, hvor I sammenfatter problemstillingen, fremgangsmåden og de vigtigste resultater.
- ▶ Hvordan når forfatterne frem til figur 1?
- ▶ Sammenlign de tre konsekvenser ("implications") i afsnittet *Discussion and Conclusions* med de 10 heuristikker, som er gennemgået ved forelæsningen den 8. februar.
- ▶ Beskriv mindst to yderligere konsekvenser ("implications") på grundlag af resultaterne under *Findings*.
- ▶ Lad os antage, at artiklen er en foreløbig udgave, som forfatterne har bedt jer reviewe. Angiv jeres 3 væsentligste forbedringsforslag.

DIKU Kursus Interaktionsdesign – Inter

# Usabilitytest

Rolf Molich

## Overblik over videoforelæsninger

- ▶ Opgave 6+7  
Opgave 8
- ▶ Usabilityevaluering  
Planlægning af usabilitytest
- ▶ Udførelse af testsessioner
- ▶ Formidling af resultater  
Modenhed
- ▶ Kvantitativ usabilitytest

## Litteratur

Evaluering af brugergrænseflader, oversigt	HP-12.5.2: The Think-Aloud Technique, s. 440-444
Usabilitytest – Forberedelse	HP-14: Rigorous Empirical Evaluation: Preparation, s. 503-536
Usabilitytest – Gennemførelse	HP-15: Rigorous Empirical Evaluation: Running the Session, s. 537-553
Usabilitytest – Rapportering	Molich, Hornbæk, Krug, Scott, Johnson: Recommendations on Recommendations
	Alternativ fremstilling – ikke pensum Molich: Vurderingsteknikker, kapitel 11

## Usabilityevaluering

En proces, som indsamler oplysninger om usability af et produkt med henblik på at forbedre produktet (formativ usabilityevaluering) eller sammenligne produktets usability med fastlagte kriterier (summativ usabilityevaluering)

## Usabilityevaluering

### Hovedformer for usabilityevaluering

- ▶ Med brugere
  - Usabilitytest
  - Spørgeskema
- ▶ Uden brugere
  - Inspektion

## Usabilitytest

En usabilityevaluering, hvor typiske brugere én for én løser konkrete opgaver med et produkt med henblik på at finde usabilityproblemer eller måle usability, dvs. gennemførlighed, effektivitet og tilfredshed.

En usabilitytest er opdelt i en række usabilitytestsessioner. I hver usabilitytestsession løser én bruger usabilitytestopgaver.

## Usabilitytest: Tænke-højt metoden

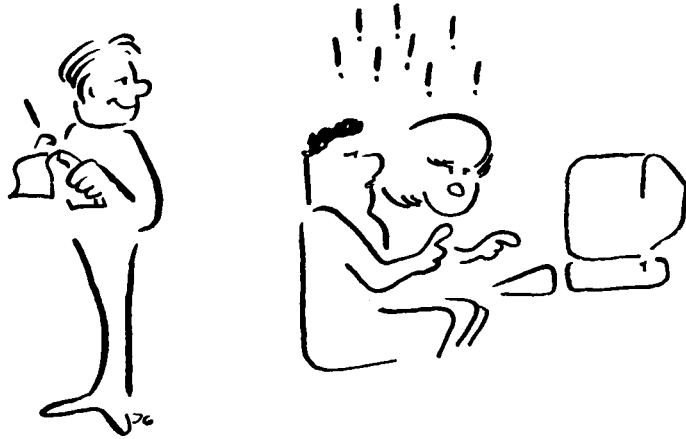


## Usabilitytest: Tænke-højt metoden





## Tænke-højt metoden – variant



## "Tænke Højt" test af en papirprototype



## Vigtige varianter af usabilitytest

	Klassisk Ansigt-til-ansigt	Fjerntest	Umodereret
Hvor befinder moderatoren sig?	Ved siden af testdeltageren	Langt fra testdeltageren	Ingen moderator
Kommunikation testdeltager-moderator	Direkte, ansigt-til-ansigt	Internet, telefon	Per e-mail efter usabilitytest-sessionen

## Usabilitytest

Man tager:

- ▶ 3-10 repræsentative brugere ("testdeltagere").
- ▶ 5-25 repræsentative opgaver, som er relevante for brugere.

Gør følgende:

- ▶ Lad hver bruger forsøge at løse opgaverne.
- ▶ Notér evt. problemer og successer.
- ▶ Hvis to eller flere brugere løber ind i det samme problem, så har produktet sandsynligvis et usabilityproblem.

## Usabilitytest: The Opinion Killer



## Testdeltagernes meninger er uvæsentlige

I en usabilitytest interesserer vi os for, om testdeltagerne kan løse de stillede testopgaver med produktet.

Testdeltagernes meninger om produktet er temmelig ligegyldige.

## Hovedformål med usabilitytest

- ▶ Tjek om produktet overholder brugerkravene
- ▶ Tjek om produktet virker fornuftigt
- ▶ Vis kolleger og interessenter, at der også er usabilityproblemer i deres produkt, og at der findes metoder til at finde disse problemer og eliminere dem

DIKU Kursus Interaktionsdesign – Inter

# Usabilitytest Planlægning

Rolf Molich

## Forbered testen

- ▶ Udnævn en neutral testleder ("moderator", "facilitator")
- ▶ "Don't test your own baby"



A good test facilitator is like a duck - serene on the surface, but paddling like heck underneath (Carolyn Snyder)



En god moderator er nysgerrig – præcis som en god interviewer Picture credit: broterham -

## Forberedelser

- ▶ Beskriv formål med usabilitytest  
Udarbejd kravspecifikation
- ▶ Diskuter testen med udviklingsteamet
- ▶ Fastlæg tid og sted
- ▶ Rekrutter testdeltagere
- ▶ Informér testdeltagere
- ▶ Udarbejd drejebog  
= Interview tjeklister + testopgaver
- ▶ Vælg dataopsamlingsmetode

## Kravspecifikation til usabilitytest

### Testplan:

- ▶ Formålet med testen
- ▶ Ansvarlig person: Testleder, moderator
- ▶ Testobjekt
- ▶ Målgruppe
- ▶ Antal testdeltagere
- ▶ Længde af hver testsession
- ▶ Tid og sted for testen
- ▶ Tidsplan
- ▶ Overslag over omkostninger

## Drejebog

### Test script

- ▶ Startbetingelser for test
- ▶ Introduktion ("briefing")
- ▶ Interview før test ("pre-test interview")
- ▶ Moderation (løs testopgaver med tænke-højt)
- ▶ Interview efter test ("debriefing")

## Fastlæg tid og sted

- ▶ En testsession tager normalt 10 – 100 minutter
- ▶ Afsæt mindst 15 minutter mellem testsessioner
- ▶ En trænet moderator kan overkomme 3-6 testsessioner pr. dag
- ▶ Nogle relevante testdeltagere kan kun deltage før eller efter arbejdstid
- ▶ Interessenter skal have mulighed for at overvære testsessionerne



## Antal testdeltagere

- ▶ Fem testdeltagere er ofte et passende kompromis mellem udbytte og omkostninger
- ▶ En populær myte er "Fem brugere er nok til at finde 85% af problemerne i et produkt". Det er ikke korrekt.

useit.com → Alertbox → Mar. 2000 Test With 5 Users

Jakob Nielsen's Alertbox, March 19, 2000:

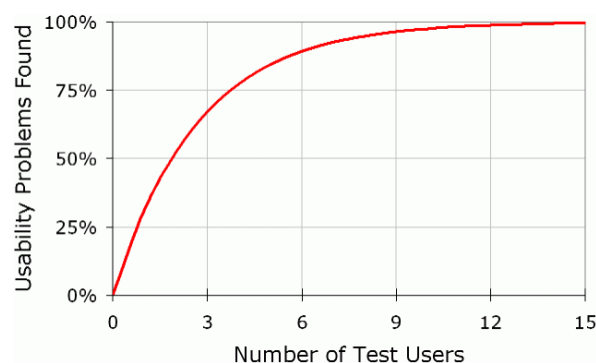
### Why You Only Need to Test with 5 Users

Some people think that usability is very costly and complex and that user tests should be reserved for the rare web design project with a huge budget and a lavish time schedule. Not true. Elaborate usability tests are a waste of resources. The best results come from testing no more than 5 users and running as many small tests as you can afford.

In earlier research, Tom Landauer and I showed that the number of usability problems found in a usability test with  $n$  users is:

$$N(1-(1-L)^n)$$

where  $N$  is the total number of usability problems in the design and  $L$  is the proportion of usability problems discovered while testing a single user. The typical value of  $L$  is 31%, averaged across a large number of projects we studied. Plotting the curve for  $L=31\%$  gives the following result:



The most striking truth of the curve is that **zero users give zero insights**.

Flot formel og kurve, men forkert resultat. For en korrekt analyse se artikler om CUE studies.

Fem brugere  
er nok til en  
nyttig iterativ  
cyklus

51

## Rekrutter testdeltagere

- ▶ Definér profilen for testdeltagerne i samarbejde med udviklingsteamet
- ▶ Testdeltagere skal være typiske hvad angår
  - viden om IT
  - viden om produktområdet
  - viden om produktet
- ▶ Kontrollér viden ved at stille specifikke spørgsmål, ikke ved at bede testdeltagerne vurdere sig selv
- ▶ Hverken for lidt eller for megen viden

## Informér testdeltagerne

- ▶ Bekræft aftalen skriftligt
- ▶ Afmystificer:  
Forklar, hvordan testen foregår
- ▶ Se eksemplet på et brugsvenligt informationsbrev i
  - Absalon > Files > Materiale som supplerer lærebogen > Molich – vurderingsteknikker
  - Absalon > Files > Opgaver og miniopgaver > Opgave 7 – Eksempel på brev til testdeltager

## Skriv testopgaver

- ▶ Ikke kun opgaver – scenarier.
- ▶ Typiske – ikke nødvendigvis simple.
- ▶ Bed undervisere, support, brugere om at kommentere testopgaverne
- ▶ Ret testopgaverne efter en pilottest i overensstemmelse med testdeltagernes kommentarer og forventninger.  
Med andre ord: Usabilitytest dine testopgaver.

## Skriv testopgaver

- ▶ En test på 30-100 minutter kan ofte kun dække produktet delvist, så du skal prioritere
- ▶ Fokus på
  - Hyppigt forekommende opgaver
  - Komplicerede opgaver
  - Opgaver, som forekommer sjældent, men som er kritiske, når de forekommer.

The screenshot shows the Avis website interface. At the top is the red Avis logo and navigation links: Om Avis, Fakturakopi, FAQ, Kontakt, Hjælp. Below this is a menu bar with links: Tilbud, Biler, Produkter og services, Elbil, Avis MiniLease, Erhverv, Kontorer, Avis Preferred, Billejekort, Partnere, Udenlandsdanskere, Bestil biljeje. The main content area is titled 'Biludlejning fra Avis - Book biljeje her' and includes links for 'Se eller ændre en reservation' and 'Hent fakturakopi'. The booking form contains fields for 'Startdato/-tid' (14 september 2014, 09:00), 'Slutdato/-tid' (16 september 2014, 09:00), 'Antal lejedage' (2), 'Afhentningsland' (Danmark), 'Udlejningsstation', and 'Aflleveres på' (Samme som udlejningskontoret). There is also a 'Log ind på din brugerprofil' section with a 'Vælg kontotype' dropdown. A red button at the bottom says 'Få tilbud på leje af Avis bil'. On the right, a large red banner advertises the 'Nissan LEAF-elbil - alt inklusiv: 3495,- - Pr. måned inkl. moms' with a 'Bestil' button. Below the main form, there are three promotional tiles: 'Minileasing af elbil' (Nissan LEAF), 'Danskere i udlandet' (Som udenlandsdanskere har du adgang til særligt fordelagtige taxfree priser på både kort- og langtidsleje), and 'Spar op til 10%' (Lej bil i Norge og Sverige, og få alt med i prisen. Ubegrænset kilometerforbrug, forsikringer med 0 kr. i selvrisiko - og endda op til 10% rabat).

Du har kun mulighed for at stille én testopgave til dette websted. Hvad ville du vælge?

## Traditionel testopgave – Avis biludlejning

- ▶ Reserver en bil i Roms lufthavn
- ▶ Afhentning fredag den 7. april omkring kl. 16
- ▶ Tilbagelevering samme sted en uge senere, dvs. 14. april, inden kl. 7
- ▶ Folkevogn Passat eller tilsvarende størrelse
- ▶ Vælg CDW-forsikring (Collision Damage Waiver)

## Åben testopgave

- ▶ Planlægger du en ferie- eller forretningsrejse, eller har du for nylig været ude at rejse?
- ▶ Hvis ja, fortæl mig om hvor du skal hen, eller hvor du har været.
- ▶ Lej en bil hos Avis, som svarer til dit behov og din økonomi, på dit rejsemål og på en tid, som passer dig.

Fordel ved åben testopgave: Mere motiverende for testdeltager.

## Vigtige tips om testopgaver

- ▶ Stil væsentlige, typiske opgaver
- ▶ Mindst én åben opgave, gerne flere
- ▶ Undgå skjult hjælp  
Undgå f.eks. fagudtryk, som optræder i brugergrænsefladen
- ▶ Undgå systemorienterede opgaver, som f.eks. login, registrér dig

## Vigtige tips om testopgaver

- ▶ Undgå "morsomme" opgaver  
Undgå specielt navne, som distraherer testdeltagere
- ▶ Undgå opgaver, hvor testdeltageren skal forestille sig, at han eller hun er en anden end de rent faktisk er ("pretender task")
- ▶ Den første opgave bør være let

## Testopgaver: Tjekliste for opgave 6

- ▶ Mindst 10 forskellige, gyldige opgaver.  
Tip: Lav 12+ opgaver
- ▶ Start med at bede testdeltageren finde webstedet
- ▶ Angiv for hver opgave
  - Formålet med opgaven
  - Hvad det forventede svar er

## Test User Experience

Brugeroplevelse (UX, User eXperience)  
er ofte mere end selve brugergrænsefladen:

- ▶ Påvirkning fra venner, annoncer, presseomtalte
- ▶ Introduktion, "Kom godt i gang"
- ▶ E-mails mellem kunde og leverandør
- ▶ Oplysninger om ordrestatus
- ▶ Modtagelse af varer
- ▶ Varernes kvalitet
- ▶ Kundeservice (tilgængelighed, kompetence)

## Sammenfatning - Forberedelse

- ▶ En vellykket usabilitytest skal forberedes ordentligt.  
Start tidligt
- ▶ Rekruttering er ikke trivial: Tjek deltagere, påmindelser
- ▶ Kvaliteten af testopgaver er vigtig:
  - Brug væsentlige, typiske opgaver
  - Mindst én åben opgave
  - Undgå skjult hjælp
  - Skriv scenarier – ikke bare testopgaver
  - Undgå "morsomme" testopgaver
  - Testdeltageren må ikke skulle foregive noget han ikke er
  - Gør den første testopgave enkel



DIKU Kursus Interaktionsdesign – Inter

Rolf Molich

# Usabilitytest

## Udførelse af testsessioner

### Overblik over videoforelæsninger

- ▶ Opgave 6+7  
Opgave 8
- ▶ Usabilityevaluering  
Planlægning af usabilitytest
- ▶ Udførelse af testsessioner
- ▶ Formidling af resultater  
Modenhed
- ▶ Kvantitativ usabilitytest

## Gennemførelse af testsession

- ▶ Klargøring inden testdeltageren ankommer
- ▶ Hils på testdeltageren
- ▶ Briefing, introduktion – Oplysninger til testdeltageren om testen
- ▶ Interview før test
- ▶ Testdeltager løser testopgaver, mens han/hun tænker højt
- ▶ Interview efter test
- ▶ Overræk gave.  
Tak testdeltageren for indsatsen.  
Følg testdeltageren ud.

## Drejebog

### Test script

- ▶ Startbetingelser for test
- ▶ Briefing, introduktion – Oplysninger til testdeltageren om testen
- ▶ Interview før test (“pre-test interview”)
- ▶ Moderation (løs testopgaver med tænke-højt)
- ▶ Interview efter test (“debriefing”)

## Briefing, Introduktion

- ▶ "Må jeg starte videooptagelsen?"
- ▶ Forklar kort, hvordan testen forløber
- ▶ "Vi tester ikke dig, men webstedet"
- ▶ "Du kan stoppe når som helst"
- ▶ "Jeg beder dig tænke højt under testen, så jeg kan følge med i de overvejelser, du gør dig"
- ▶ "Har du nogen spørgsmål?"

## Interview før test    Pre-test interview

- ▶ Om testdeltageren: Beskæftigelse
- ▶ "Testen handler om bibliotek.dk. Kender du dette websted?"
  - Hvis ja,
    - "Hvad kan dette websted?"
    - "Har du nogensinde besøgt webstedet?"
    - "Hvis ja, hvad var dit ærinde?"
    - "Hvordan gik det?"
      - "Vis mig hvad du gjorde"
      - "Hvad synes du om webstedet?"
- ▶ "Har du nogensinde besøgt websteder for andre biblioteker?"
  - "Hvis ja, hvad var dit ærinde?" Osv.

## Interview efter test    Debriefing

- ▶ "Hvilke 2-3 ting fungerer bedst på webstedet?"
- ▶ "Hvilke 2-3 ting trænger mest til at blive forbedret?"

## Interview efter test    Debriefing

- ▶ "Hvilke 2-3 ting fungerer bedst på webstedet?"
  - ▶ "Hvilke 2-3 ting trænger mest til at blive forbedret?"
- 
- ▶ "Kender du andre, tilsvarende websteder?  
Hvis ja, hvilke og hvordan er de i sammenligning med dette websted?"
  - ▶ "Vil du anbefale dette websted til andre?"
  - ▶ "Har du nogen kommentarer til de opgaver, som jeg har stillet?"

## Moderation: Vigtigste udfordringer

- ▶ Moderatoren hjælper for hurtigt.  
Testsessionen udarter til en hyggelig demonstration af produktet i stedet for en test.
- ▶ Moderatoren snakker for meget.  
Testsessionen udarter til hyggesnak eller et interview.

En usabilitytest er  
fundamentalt  
forskellig  
fra et interview

## Moderation: Vigtigste udfordringer

- ▶ Moderatoren holder sig strengt til drejebogen og følger ikke nysgerrigt op på uventede begivenheder.
- ▶ Moderatoren giver skjult hjælp, f.eks. "tampen brænder"
- ▶ Ingen interview efter test. Testsessionen slutter, når testdeltageren har løst den sidste opgave.
- ▶ Interview efter test er en demonstration af alle de vidunderlige ting, som produktet kan, men som testdeltageren desværre ikke kunne finde ud af.

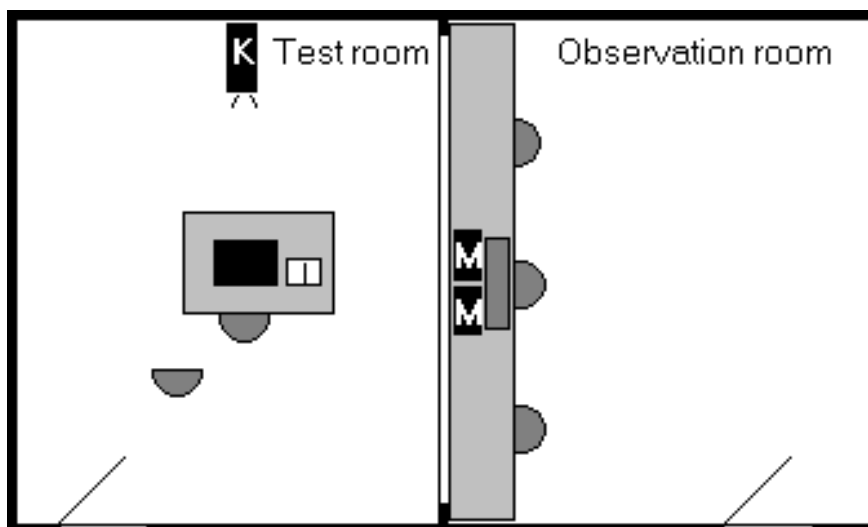
## Dataopsamling



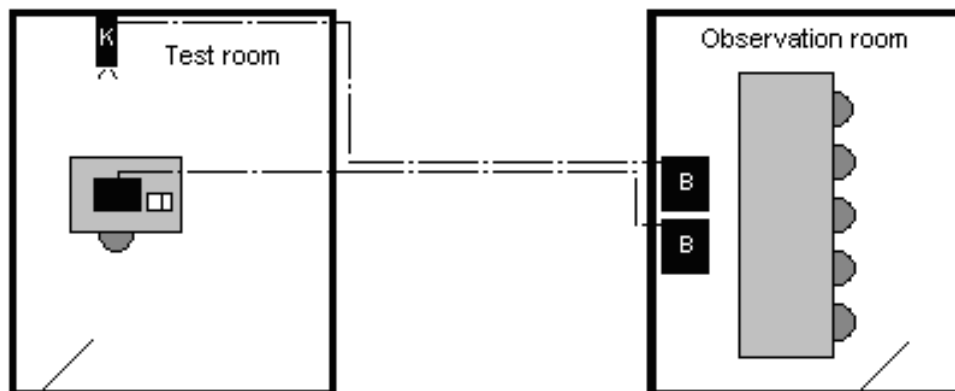
## Dataopsamling

- ▶ Blyant og papir
  - Øv dig på forhånd
  - Brug forkortelser
  - Skærbilleder til noter.
- ▶ Lydoptagelse
- ▶ Video af skærm, f.eks. Camtasia fra Techsmith
- ▶ Video af skærm og testdeltager
  - Pas på: Nogle testdeltagere kan ikke lide det. Fortæl dem om evt. video på forhånd
  - Overvej på forhånd hvad du vil bruge video til, f.eks. dokumentation, opsummering

## Usability Lab

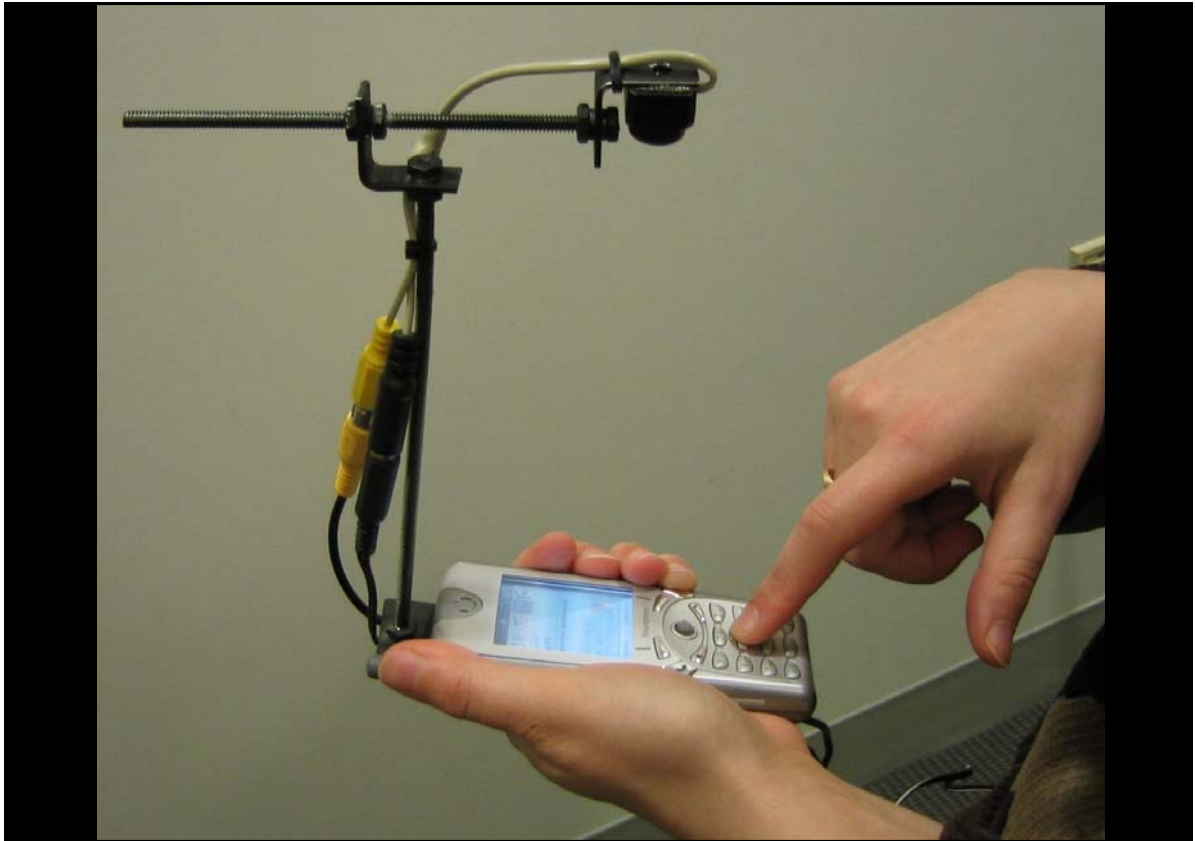


## Usability Lab









## Sammenfatning - Testsession

- ▶ Fire faser:
  - Briefing
  - Interview før test,
  - Løs testopgaver,
  - Interview efter test (debriefing)
- ▶ Moderation er vanskeligt:
  - Ti stille,
  - Undgå skjult hjælp,
  - Lad være med at hjælpe for tidligt
- ▶ Vær åben over for det uventede, uforudsigelige
- ▶ Behandl testdeltagere som du selv ønsker at blive behandlet. Prøv bare én gang selv at være testdeltager.

DIKU Kursus Interaktionsdesign – Inter

Rolf Molich

# Usabilitytest

## Formidling af resultater

### Overblik over videoforelæsninger

- ▶ Opgave 6+7  
Opgave 8
- ▶ Usabilityevaluering  
Planlægning af usabilitytest
- ▶ Udførelse af testsessioner
- ▶ Formidling af resultater  
Modenhed
- ▶ Kvantitativ usabilitytest

## Formidling af testresultater

- ▶ Workshop
- ▶ Testrapport
- ▶ Præsentation

En perfekt  
testrapport,  
som ingen læser,  
løser ingen  
problemer

## Formidling af testresultater

En perfekt testrapport, som ingen læser, løser ingen problemer

Altså:

- ▶ Du skal *sælge* din testrapport
- ▶ Du skal *engagere* brugerne af din testrapport

## Hovedformål med usabilitytest

- ▶ Hovedformålet med en usabilitytest er at få forbedret brugergrænsefladen.
- ▶ Hovedformålet er IKKE at skrive en god testrapport.

## Skab enighed (1/3)

KJ-metoden (KJ = Jiro Kawakita):

- ▶ Bed udviklere og andre interessenter om at overvære en eller flere testsessioner og tage notater.
- ▶ Saml alle observatører umiddelbart efter den afsluttende testsession.
- ▶ For at få lov til at deltage i denne workshop skal man have observeret mindst én testsession "live". Ellers bygger ens indlæg og forslag på meninger og ikke på observationer af brugere.

## Skab enighed (2/3)

- ▶ Bed hver deltager om at notere hvert større iagttaget problem på en Post-it seddel
- ▶ Placér alle sedlerne på en stor tavle. Ingen diskussioner!

## Skab enighed (3/3)

- ▶ Læs de andre sedler i tavshed.  
Tilføj evt. yderligere problemer
- ▶ Sortér problemerne efter hovedområder.  
Kombinér identiske problemer – hvis enighed
- ▶ Giv hver gruppe et navn
- ▶ Udpeg de vigtigste problemer ved afstemning  
Hver mødedeltager har 10 stemmer (x'er)

## Kommunikation

Testresultater skal:

- ▶ Leveres hurtigt
- ▶ Være præcise og letforståelige
- ▶ Være brugbare

Giv udviklerne en chance for at kommentere testrapporten,  
før du viser den til andre.

## Usabilitytestrapport: Grundprincip

- ▶ Din usabilitytestrapport skal være *brugsvenlig*

## Usabilitytestrapport – Indhold

- ▶ Sammenfatning ("Executive summary")
- ▶ Testresultater
  
- ▶ Drejebog
- ▶ Metode, udstyr
- ▶ Beskrivelse af testdeltagere



## Brugsvenlig usabilitytestrapport

- ▶ Højst 25 sider - plus appendikser
- ▶ Sammenfatning står først
- ▶ Sammenfatning fylder højst én side
- ▶ Testresultater kommer umiddelbart efter sammenfatning
- ▶ Testresultater er illustrerede
- ▶ 15-40 testresultater på 5-20 sider

## Sammenfatning Executive summary

- ▶ Målgruppe: Ledelsen
- ▶ Én side, hverken mere eller mindre
- ▶ Sammenfatningen skal være en reel sammenfatning – ikke en måde at luske yderligere problemer ind i rapporten.  
Altså: Alle problemer i sammenfatningen skal også være omtalt i selve rapporten

## Sammenfatning Executive summary

Indhold:

- ▶ Formål med testen
- ▶ Dato, sted, antal testdeltagere
- ▶ De tre vigtigste positive kommentarer
- ▶ De tre vigtigste problemer

## Usabilitytestrapport: Testresultater

For hvert resultat:

- ▶ Overskrift
- ▶ Beskrivelse
- ▶ Klassifikation – Hvor alvorligt er problemet?
- ▶ Løsningsforslag

## Usabilitytestrapport: Testresultater

### Klassifikation:

- ▶ Positivt resultat.
- ▶ Eksistenstruende problem
- ▶ Kritisk problem
- ▶ Alvorligt problem
- ▶ Mindre problem
- ▶ Forbedringsforslag
- ▶ Programfejl ("bug").



## Tips til din rapport

Kritisk problem: Gav anledning til hyppige katastrofer. En katastrofe er en situation, hvor webstedet "vandt" over testdeltagerne, dvs. en situation som forhindrede testdeltagerne i at løse en rimelig arbejdsopgave på webstedet, eller som irriterede testdeltagerne voldsomt.

- ▶ Hvis kun én testdeltager oplever et "kritisk" problem, er det måske ikke så "kritisk" endda.

## Tips til din rapport

Tal pænt til brugerne af din rapport.

Kommentarer som

- ▶ *Forsiden er uprofessionel*
  - ▶ *Lad være med at bruge ordet "ulovlig"!!!*
- skaber fjender, ikke et mere brugbart produkt

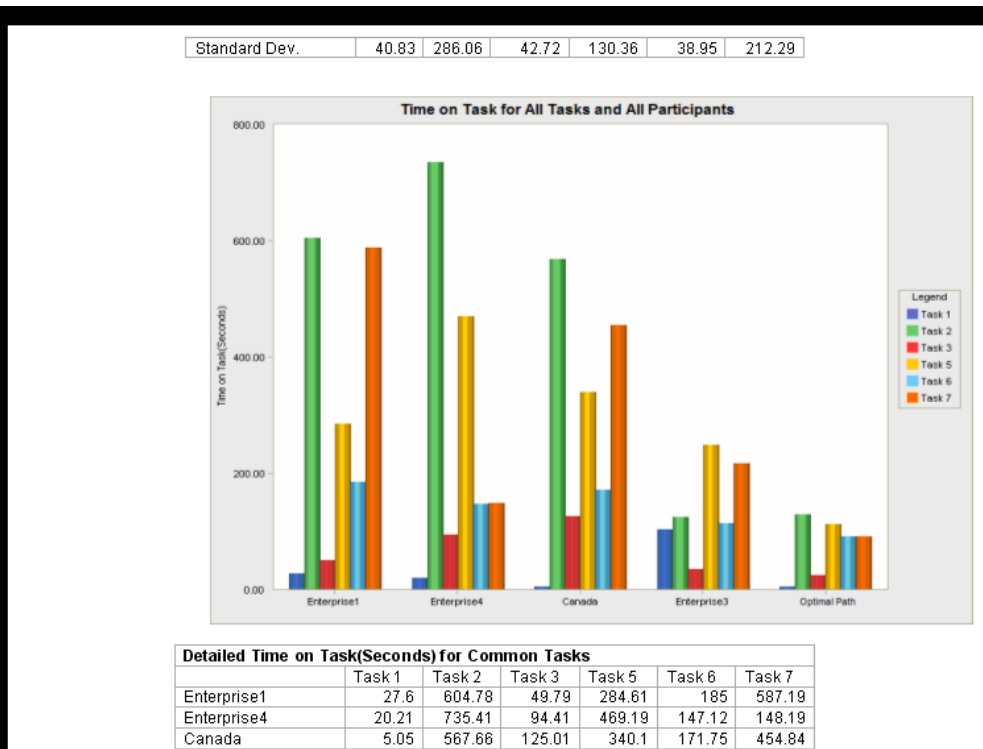
## Tips til din rapport

- ▶ Medtag kun forbedringsforslag, hvis du er nogenlunde sikker på, at de rent faktisk forbedrer brugergrænsefladen
  - ▶ Skeln skarpt mellem
    - Dine egne meninger
    - Testdeltagernes meninger
    - Hvad testdeltagerne ikke kunne finde ud af
- |                   |
|-------------------|
| Uvæsentligt       |
| Mindre væsentligt |
| Væsentligt        |

## Tips til din rapport

- ▶ Brug citater.  
Velvalgte citater "formidler brugerens smerte" og tillader forfatteren at viderebringe budskaber, som han ikke selv kan lægge navn til  
  
Eksempel: *Stavefejlene får det til at se sjusket og uprofessionelt ud. Hvem fanden har læst korrektur på det her?*
- ▶ Ordrette testprotokoller er tidsspilde og ubrugelige

Skriv  
*brugsvenlige*  
testrapporter



Uddrag af testrapport - CUE-6, Team C. Flot diagram, men for mange betydende cifre. Opsætningen antyder en lørdighed i resultaterne, som der ikke er dækning for.

TaskTimer Jan'98 Page 11

Users commented negatively on TaskTimer's use of color. For instance:

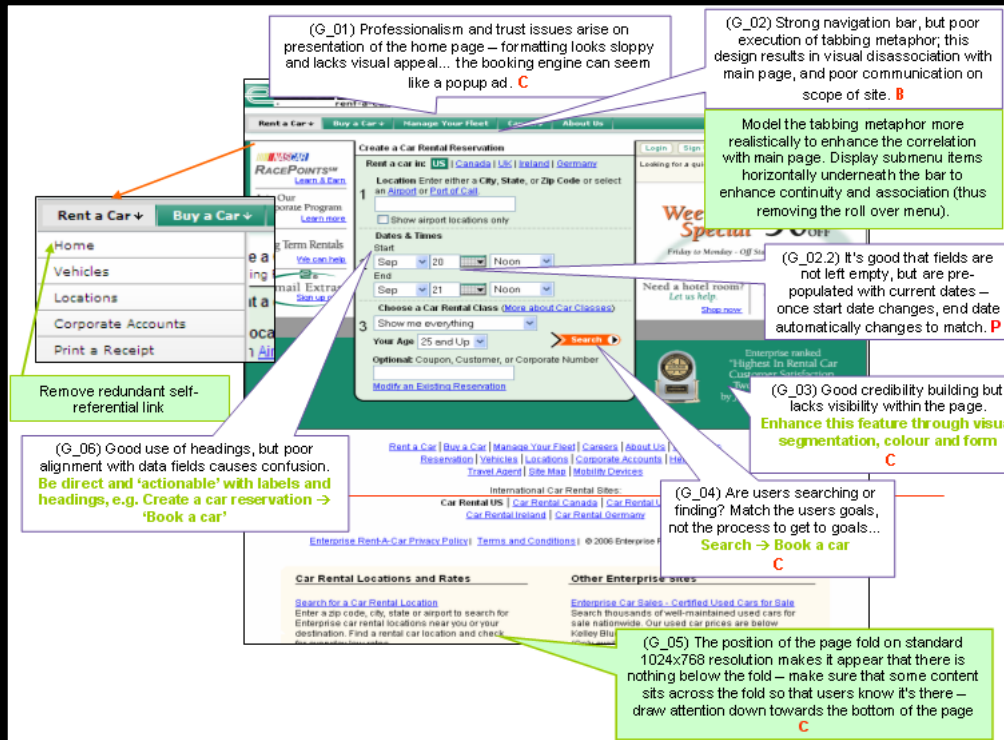
- Users did not like the use of red as a background color in toolbars. One user gave credit for attempting to deviate from standard Windows colors, but thought the attempt failed.

- Users thought the use of two shades of gray in dialogs yielded dialogs that were too busy.

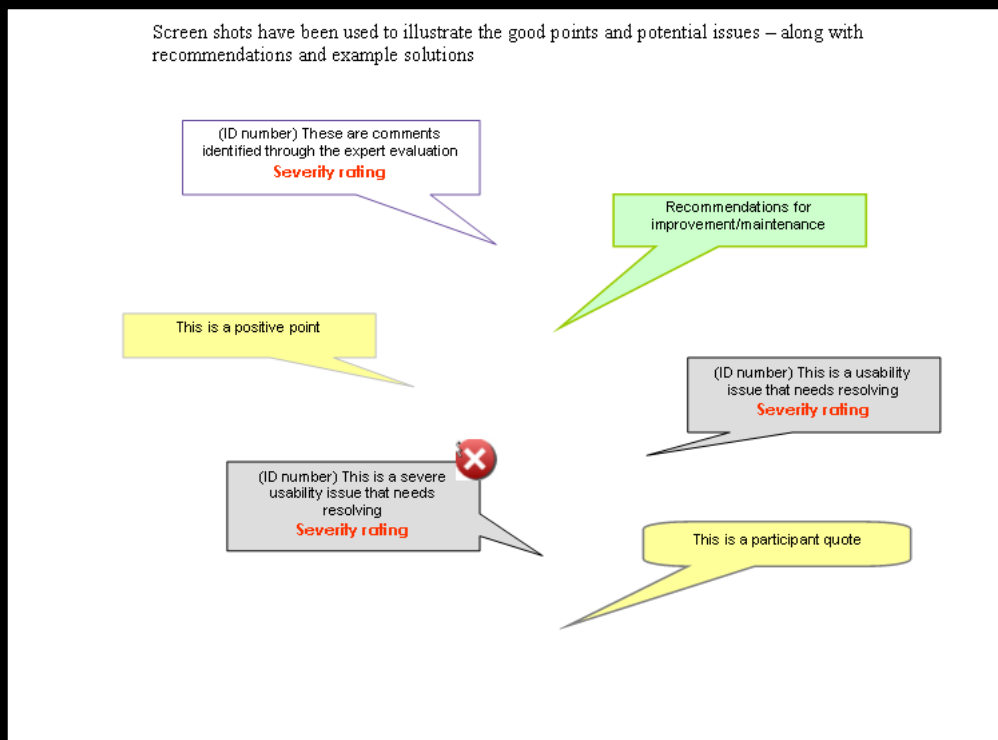
Toolbar icons, in general, were unclear. Users commented upon the use of color and the use of text in these icons as being strange.

All users wanted standard Microsoft tool tips (shown at right) even though they noticed the status bar messages. Users said tool tips were easier to use because they could keep focused on the top of the screen instead of having to shift their gaze to the bottom of the screen.

Uddrag af testrapport - CUE-1, Team C. Interessant alternativ opsætning af testrapport



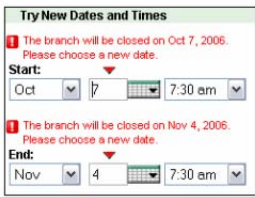
Uddrag af testrapport - CUE-6, Team G. Uoverskuelig og svær at forstå



Uddrag af testrapport - CUE-6, Team G. Signaturforklaring til rapport fra foregående planche

Team K: Show Stoppers \*\* indicates a high-priority issue

**Show Stoppers:** These problems are critical because 1) they encourage users to leave the site, and 2) each was found by several of our small number of users. We recommend that this set of usability issues be addressed first. (Some of the problems that we categorized as bugs in a later table are also show stoppers.)

Problem	Reason it was a problem	Possible solution
<b>**Selection of date and time: Changing Times Without Notification</b> I have selected a pick-up and drop-off date when my pick-up location was closed. When I select the pick up location very clear feedback of my error was shown on the selected date, telling me what the problem was and suggesting what to do to correct it. 	The only problem is that the system has now changed my drop-off and pick up time without advising me of the change. I may change my date without realizing that the time is not the one that I originally entered.	Add a message or a graphical affordance that warns me that the pick-up and drop-off time are changed. Or: Preserve the date I have entered.
<b>Selection of different drop-off location</b> Let's assume that I want to schedule pick up in one location and the drop-off in a different location. How do I do that? The system forces me to have a drop-off schedule that is compatible with the business hours of my pick-up location.	This may force me to reserve the car by phone - in case that my drop-off time is not compatible with the pick-up location, but it is indeed compatible with the actual drop-off location.	Allow users to specify a drop-off location that is different from the pick-up location.

Uddrag af testrapport - CUE-6, Team K. . Interessant alternativ opsætning af testrapport

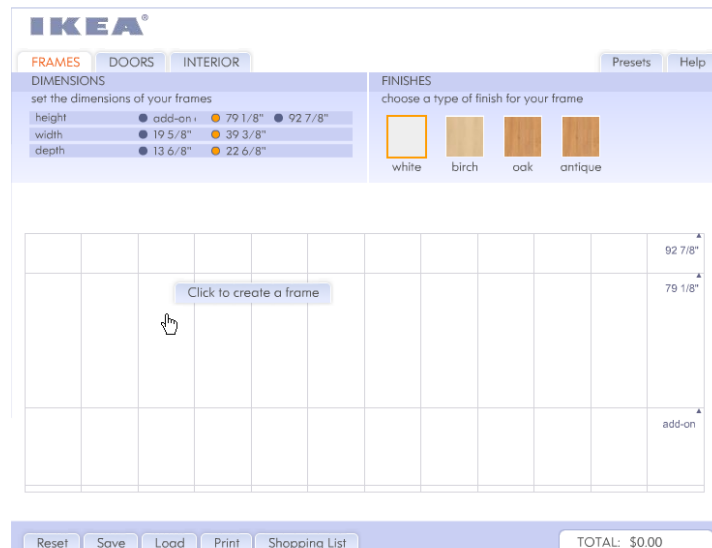
## Et godt forbedringsforslag ...

- ▶ Løser usabilityproblemet
- ▶ Er konstruktivt og direkte
- ▶ Er forståeligt, dvs. taler brugerens sprog
- ▶ Er detaljeret, så kreative misfortolkninger undgås
- ▶ Er illustreret, f.eks. med et redigeret skærbillede
- ▶ Er begrundet
- ▶ Anbefaler den mindst mulige ændring

Læs mere i artiklen *Recommendations on Recommendations* som indgår i pensum

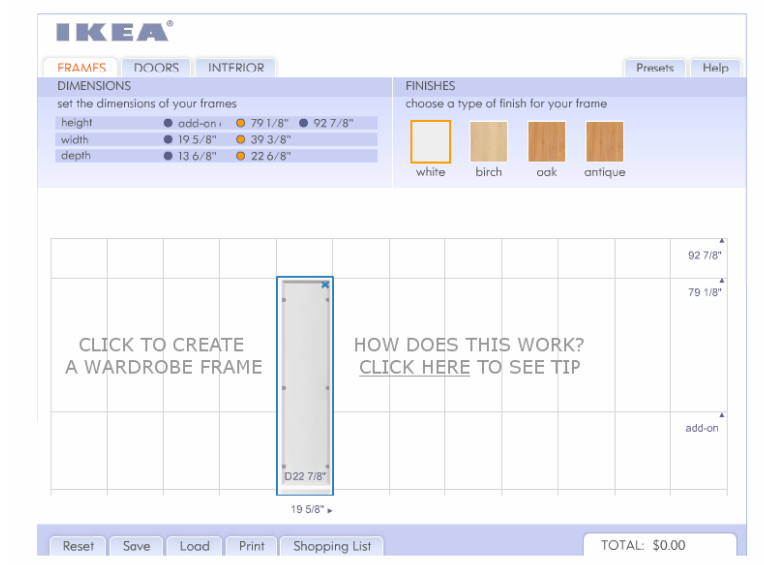


## Eksempel på illustreret forbedringsforslag



Test viste, at brugere var i tvivl om hvad de skulle gøre på dette skærmbillede

## Eksempel på illustreret forbedringsforslag



Illustreret forbedringsforslag

## Videosammendrag Highlights video

Et videosammendrag præsenterer de vigtigste testresultater

- ▶ Målgruppen for videosammendraget
  - Kender normalt produktet i forvejen
  - Har en begrænset viden om usability
  - Har ikke overværet testsessionerne
- ▶ Der findes gratis værktøjer til optagelse og redigering af testvideoer, f.eks.
  - Techsmiths Camtasia fra [www.techsmith.com](http://www.techsmith.com) til Windows og Mac
  - Quicktime til Mac

## Videosammendrag

Brugsvenlige videosammendrag:

- ▶ Mindst 5 minutter og højst ca. 12 minutter.
- ▶ Vis nogle få, alvorlige usabilityproblemer, f.eks. 2 eller 3. Illustrér hvert problem med mindst to klip med forskellige testdeltagere.
- ▶ Vis, hvad der sker på testdeltagerens skærm. Det ikke et krav, at man kan se testdeltageren eller moderatoren på videosammendraget.
- ▶ Man skal klart kunne forstå, hvad testdeltageren siger. Brug evt. undertekster
- ▶ Start med titel: Formål, dato, kontaktpersoner og videoens længde

## Opgave 6+7: Typiske fejl

Den vigtigste årsag til at få en rapport afvist er, at rapporten indeholder væsentlig færre end 20 gyldige kommentarer



## Opgave 6+7: Ugyldige kommentarer

- ▶ Ubearbejdede – Beskriver hvad testdeltageren gjorde uden at forklare, hvad usabilityproblemet er
- ▶ Uforståelige
- ▶ Fejlklassificerede – f.eks. mindre problem <> kritisk problem
- ▶ Vage – f.eks. "Webstedet taler ikke brugernes sprog". Fuldt pointtal kræver mindst to konkrete eksempler.
- ▶ Gentager tidligere omtalte problemer
- ▶ Ingen omtale af testdeltagere, lyder som et inspektionsresultat eller en personlig mening



## Sammenfatning

- ▶ Fokuser på kritiske problemer ("katastrofer")
  - ikke mikro-brugervenlighed
- ▶ Rapportér, hvad testdeltagerne ikke kunne finde ud af
  - ikke testdeltagernes meninger, og slet ikke dine egne meninger
- ▶ Hold øje med det uventede, uforudsigelige
  - "Det er tilladt at undre sig!"
- ▶ Behandl testdeltagere som du selv ønsker at blive behandlet
  - Prøv bare én gang selv at være testdeltager
- ▶ Udvis rettidigt omhu – god forberedelse tager længere tid end du tror
- ▶ Sælg dine resultater godt – de sælger ikke sig selv

DIKU Kursus Interaktionsdesign – Inter

# Usabilitytest Modenhed

Rolf Molich

Cowboy  
programmers  
don't need no  
stinkin' usability

-- Anonym

## Hvorfor er politik vigtigt?

Når ledelsen siger “kvalitet frem for alt” og derefter tilsidesætter kvalitetskontrollen for at gennemtvinge lanceringen af et nyt produkt inden for en bestemt tidsfrist, så forstår vi budskabet. Udviklere er gode til at opfatte sådanne signaler korrekt.

Hvad folk siger, er ord, men hvad folk gør, viser, hvad der virkelig er vigtigt for dem.

-- Phillip G. Armour (oversat af Rolf Molich)

God is on our side  
– but you  
can help him

-- Bruce Tognazzini

## Usability modenhed

- ▶ Den gode usabilityspecialist sælger usability, fjerner forhindringer, motiverer, implementerer og styrker usability
- ▶ Synliggør konsekvenserne af at ignorere usability ved hjælp af usabilitytest
- ▶ Husk: Mange af dine dygtige kolleger har aldrig set en bruger anvende jeres produkt
- ▶ Undgå meninger – Test!

## Usability modenhed – Usability Maturity

- ▶ Skab alliancer
- ▶ Undgå konfrontationer
- ▶ Vælg dine konflikter med omhu;  
du kan ikke vinde dem alle  
– og selv hvis du kunne, ville det være uklogt
- ▶ Positionér dig som en allieret  
– ikke som en fjende eller politifunktion
- ▶ Skab tillid ved at være fuldstændig åben. Afmystificér.  
Usability er ikke raketvidenskab
- ▶ Dokumentér og sælg successer til ledelsen

## Usability modenhed – Primitiv model

- 3. Ingen taler om usability
- 6. Alle taler om usability
- 7. Ingen taler om usability

## Usability modenhed – Forenklet model

- 2. Forudsigelig (bedst)  
Organisationen kan præcist forudsige omkostninger og kvalitet for brugercentreret design.
- 3. Veletableret  
Organisationen har implementeret en brugercentreret designproces med veldefinerede aktiviteter og milepæle. Der foretages regelmæssige målinger af brugsoplevelsen.
- 5. Udføres  
Entusiastiske enkeltpersoner udfører usabilityaktiviteter.
- 6. Ufuldstændig (dårligst)  
Der tales meget om usability, men der anvendes ikke anerkendte metoder. Projektledere siger, at usability er vigtigt for dem, men det omsættes ikke i konkrete beslutninger.



## Usability modenhed – Fuldstændig model

The maturity (capability) levels of standard process assessment models according to ISO 18529 are:

1. Optimizing – Optimerende  
Organization can reliably tailor the human-centred design process to particular requirements.
2. Predictable – Forudsigelig  
Performance of the human-centred design process is within predicted resource and quality limits.
3. Established – Veletableret  
The organization has implemented a human-centred design process, with multiple activities and milestones.  
The organization tracks user experience quality throughout design projects and across releases.

## Usability modenhed – Fuldstændig model

4. Managed – Styret  
Quality, time, and resource requirements for the human-centred design process are known and controlled.
5. Performed – Udføres  
The human-centred design process achieves its purpose.  
Individuals carry out processes.
6. Incomplete – Ufuldstændig  
The organization is not able to carry out the human-centred design process. Product managers say that they care about usability, but when it comes to spending resources or making inconvenient decisions to achieve usability, nothing happens.  
Usability is fine if it comes for free, but no one is committed to it.

## Usability modenhed – Fuldstændig model

7. Indifference or outright hostility towards Usability  
Developers don't care about users or their needs; their only goal is to build features and make them work on the computer. In this mindset, humans are irrelevant – they're told to use the system, regardless of whether doing so is easy or pleasant. Usability activities are considered counterproductive and suppressed.

## Sammenfatning – Øge modenhed

- ▶ Forklar hvad UX arbejde går ud på  
Det er ikke raketvidenskab
- ▶ Involver interessenter i planlægning af usabilitytest, specielt i fastlæggelse af testopgaver og testdeltagerprofiler
- ▶ Inviter interessenter til at overvære usabilitytest
- ▶ Afhold usabilitytest på tidspunkter, som passer interessenter
- ▶ Involver interessenter i fastlæggelse af resultaterne, f.eks. ved at bruge KJ-metoden

DIKU Kursus Interaktionsdesign – Inter

# Usabilitytest Kvantitativ

Rolf Molich

## Overblik over videoforelæsninger

- ▶ Opgave 6+7  
Opgave 8
- ▶ Usabilityevaluering  
Planlægning af usabilitytest
- ▶ Udførelse af testsessioner
- ▶ Formidling af resultater  
Modenhed
- ▶ Kvantitativ usabilitytest



## Kvantitativ usabilitytest

En usabilitytest, som har til formål at fremskaffe måleværdier for gennemførlighed, effektivitet og tilfredshed for et produkt

Kvantitative usabilitytests bruges til at

- ▶ Sammenholde en prototype eller et produkt med brugerkrav
- ▶ Sammenligne produkter, f.eks. konkurrerende produkter
- ▶ Sammenligne versioner af et interaktivt system

## Måleresultater

Et muligt resultat af en måling er  
"Succes – testdeltageren løste opgaven korrekt på 204 sekunder"

Mulige måleresultater

- ▶ Succes
- ▶ Testdeltageren kunne ikke løse opgaven
- ▶ Katastrofe
- ▶ Ugyldig måling

## Kvantitativ usabilitytest

Parametre, som ofte måles i en usabilitytest:

- |  |                            |
|--|----------------------------|
| ▶ Udførelsestid, tid til at løse opgaven | Et mål for effektivitet    |
| ▶ Succesrate                             | Et mål for gennemførlighed |
| ▶ Fejlrate                               | Et mål for gennemførlighed |
| ▶ Katastroferate                         | Et mål for gennemførlighed |

$\text{Succesrate} + \text{fejlrate} = 100\%$

Fejl inkluderer katastrofer

## Kvantitativ usabilitytest - noter

- ▶ For at få gyldige målinger er det nødvendigt at rekruttere **repræsentative testdeltagere**
- ▶ **Kvalitative resultater**, som opnås i en kvantitativ usabilitytest, bør også rapporteres for at hjælpe interessenter med at finde ud af, hvordan et utilfredsstillende resultat kan forbedres
- ▶ Testdeltagere bør have at vide, at tiden måles, og at de skal give besked, når de mener, at de har løst en opgave eller giver op. Tilføj: "Det er ikke for at stresse dig – vi vil bare gerne vide, hvor lang tid det tager.
- ▶ I kvantitative usabilitytest bør testdeltageren **ikke tænke højt**, da det kan påvirke den tid, det tager at løse opgaven.
- ▶ Tilfredshed måles med et **spørgeskema**, f.eks. SUS, System Usability Scale

## Udførelsestid, tid til at løse opgaven

Sluttid minus starttid for en testdeltagers løsning af en testopgave

- ▶ Starttid: Tidsmålingen starter, når testdeltageren har læst opgaven, og den relevante brugergrænseflade er synlig for testdeltageren. Starttiden er ikke det tidspunkt, hvor testdeltageren modtager testopgaven.
- ▶ Sluttid: Tidsmålingen slutter, når testdeltageren siger, at han eller hun har fundet en korrekt løsning på testopgaven, og løsningen er korrekt. Testdeltageren behøver ikke læse svaret højt, men det er ikke nok, at svaret står på skærmen.
- ▶ Engelsk: Task completion time.

## Udførelsestid – Usikkerhed

Det er ofte svært at fastslå den præcise start- og sluttid.

Usikkerheden er ofte 5 sekunder eller mere.

## Udførelsestid – Bemærkninger

Den rapporterede udførelsestid er det geometiske gennemsnit af udførelsestiderne for alle korrekte løsninger af usabilitytestopgaven.

- ▶ Hvis en testdeltager ikke kan løse én testopgave, er der ingen sluttid, og vi beregner ikke en udførelsestid for den pågældende testdeltager.

I stedet indregnes resultatet i fejlraten.

## Succesrate

Procentdelen af alle forsøg på at løse en usabilitytestopgave, som lykkes.

Succesraten er et mål for gennemførligheden på webstedet.

Et forsøg på at løse en usabilitytestopgave er en succes, hvis testdeltageren kan finde det rigtige svar på egen hånd.

## Fejlrate

Procentdelen af alle forsøg på at løse en usabilitytestopgave, som mislykkes.

Et forsøg på at løse en usabilitytestopgave mislykkes, hvis testdeltageren

- ▶ Giver op
- ▶ Er ude af stand til at finde et korrekt svar inden for en rimelig tid
- ▶ Finder et ukorrekt svar og tror, at det er korrekt  
Sammenhold med katastroferate.

Mislykkede forsøg inkluderer katastrofer.



## Katastroferate

Procentdelen af alle forsøg på at løse en usabilitytestopgave, som resulterer i en katastrofe.

Et forsøg på at løse en usabilitytestopgave er en katastrofe, hvis testdeltageren

- ▶ Finder et forkert svar og tror så meget på det, at han/hun vil fortsætte sit arbejde på grundlag af det forkerte svar, OG
- ▶ Konsekvenserne af at bruge det forkerte svar er kritiske eller en alvorlig trussel mod helbred eller økonomi.

Bemærk den væsentlige forskel mellem fejlrate og katastroferate. Normalt er katastroferaten nul eller meget mindre end fejlraten.

## Katastrofe – Eksempler

- ▶ Problem på webstedet for en kommune: Testdeltageren finder et forældet telefonnummer til borgerservice. Dette er ikke en katastrofe, fordi brugere let opdager at nummeret er forkert, når de ringer op, så konsekvenserne er begrænsede.
- ▶ Katastrofe på webstedet for en kommune: Testdeltageren finder et forældet telefonnummer til nødhjælp.
- ▶ Katastrofe på webstedet for en kommune: Testdeltageren finder en forkert dato for næste afhentning af storskrald.

Den tid, som brugere anvender på et katastrofalt forsøg på at løse en usabilitytestopgave, indgår ikke i beregningen af udførelsestiden.

## Ugyldig måling

En måling af en usabilitytestopgave, som ikke må inkluderes i resultaterne på grund af fejl i udstyr, fejl i software, fejl i fremgangsmåde eller utilsigtet afbrydelse. Denne liste er ikke udtømmende.

Udførelsestiden for en ugyldig måling må ikke medtages i beregningen af den samlede udførelsestid.

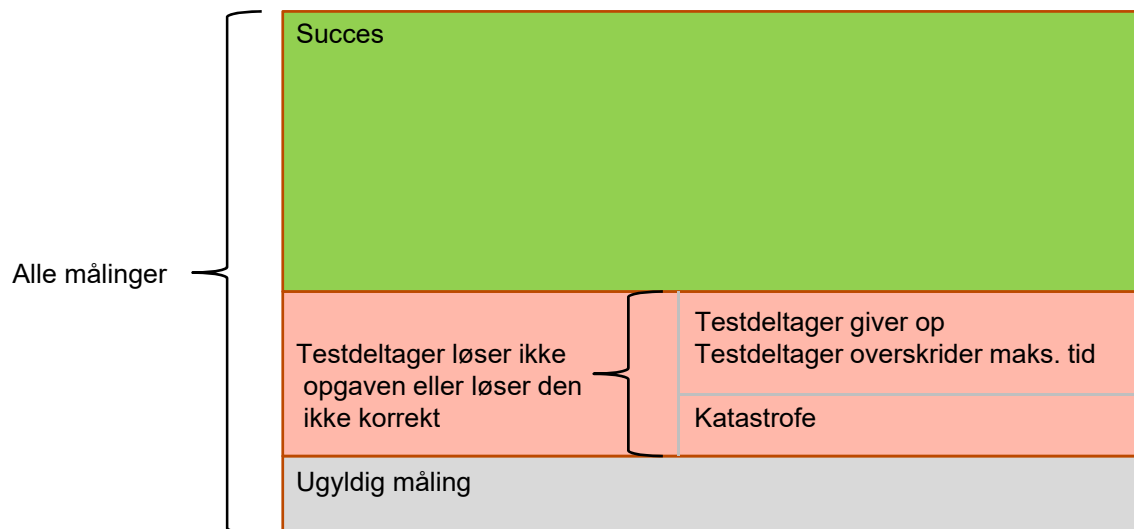
En ugyldig måling må ikke indregnes i fejlraten.

## Ugyldig måling

Eksempler på grunde til at erklære en måling ugyldig:

- ▶ Moderatoren hjælper testdeltageren med opgaven.  
Det er OK hvis testdeltageren får hjælp fra den ordinære support eller online hjælp.
- ▶ Moderatoren griber ind for at løse et teknisk problem
- ▶ Hardware eller software problemer på computeren
- ▶ Nødvendigt software ikke installeret på computeren, f.eks. Adobe Reader
- ▶ Testdeltageren afbrydes i mere end 10 sekunder, f.eks. af en kollega eller en telefonsamtale, medmindre afbrydelsen er en planlagt del af opgaven
- ▶ Testdeltageren diskuterer med moderatoren

## Oversigt over måleresultater



## Statistiske begreber

For at kunne planlægge og gennemføre målinger i en kvantitativ usabilitytest er det nødvendigt at have et vist kendskab til en række statistiske begreber:

- ▶ Gennemsnit
- ▶ Geometrisk gennemsnit
- ▶ Standardafvigelse
- ▶ Konfidensinterval.

## Statistiske begreber

For at kunne planlægge og gennemføre målinger i en kvantitativ usabilitytest er det nødvendigt at have et vist kendskab til en række statistiske begreber:

- ▶ Gennemsnit
- ▶ Geometrisk gennemsnit
- ▶ Standardafvigelse
- ▶ Konfidensinterval.

Hvis du er i tvivl om disse begreber, eller hvis din kvantitative usabilitytest danner grundlag for vigtige beslutninger, så rådfør dig med en sagkyndig på området.

Du kan finde yderligere oplysninger om disse emner sammen med værktøjer, som kan beregne konfidensintervaller mm. på Jeff Sauros website **measuringu.com**

## Gennemsnit

Summen af tiderne fra alle gyldige målinger divideret med antallet af målinger.

Gennemsnit kaldes også aritmetisk gennemsnit (modsat geometrisk gennemsnit, som er omtalt på næste planche)

Det er aldrig tilstrækkeligt at rapportere et gennemsnit. Du skal også rapportere det tilsvarende konfidensinterval.

Eksempler:

- ▶ Gennemsnittet af 3, 5 og 10 er 6, fordi  $(3+5+10)/3 = 6$
- ▶ Gennemsnittet af 5, 6, 7, 7, 7, 8, 9, 9, 9, 23 minutter er 9 minutter

## Geometrisk gennemsnit

En form for gennemsnit, som angiver den væsentlige tendens eller typiske værdi for et antal tal med mindre vægt på ekstreme værdier.

Det geometriske gennemsnit af N tal beregnes som den N'te rod af produktet af tallene.

- ▶ GEOMEAN funktionen i Excel beregner det geometriske gennemsnit.

## Geometrisk gennemsnit – Eksempler

- ▶ Det geometriske gennemsnit af 2 og 8 er 4, fordi kvadratroden af  $2 \cdot 8$  er 4.
- ▶ Det geometriske gennemsnit af 2, 10 og 11 er 6,04, fordi den tredje rod af  $2 \cdot 10 \cdot 11 = 220$  er 6,04
- ▶ Det geometriske gennemsnit af 5, 6, 7, 7, 7, 8, 9, 9, 9, 23 minutter er 8,2 minutter  
Til sammenligning er gennemsnittet af 5, 6, 7, 7, 7, 8, 9, 9, 9, 23 minutter = 9 minutter

## Standardafvigelse

Standardafvigelsen angiver hvor meget målingerne afviger fra gennemsnittet. Jo større standardafvigelse, desto mere er målingerne spredt.

Standardafvigelsen er kvadratroden af variansen.

Variansen er summen af kvadratet på afvigelserne fra gennemsnittet.

- ▶ STDEV funktionen i Excel beregner standardafvigelsen.

## Konfidensinterval

Et skøn over det interval, som den sande værdi befinder sig i.

- ▶ Formålet med konfidensintervallet er at vise det interval, hvor den korrekte værdi ligger med en vis sandsynlighed, f.eks. 95%
- ▶ "Den korrekte værdi" er den værdi, som vi ville få, hvis alle brugere udførte usabilitytestopgaven.
- ▶ Rapporterede måleværdier bør altid ledsages af et konfidensinterval. Med andre ord bør du ikke bare rapportere, at det tog gennemsnitligt 140 sekunder at løse en opgave. I stedet bør du rapportere, at det med 90% sandsynlighed tager gennemsnitligt mellem 90 og 210 sekunder at løse opgaven.

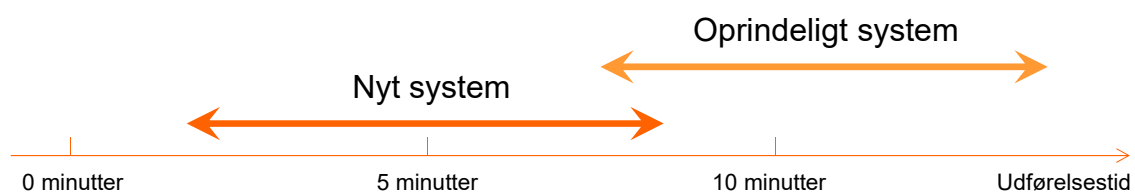
## Konfidensinterval – Bemærkninger

- ▶ Hvis du øger antallet af testdeltagere i en kvantitativ usabilitytest, indsnævres konfidensintervallet.
- ▶ For at opnå et rimeligt smalt konfidensinterval, bør du teste mindst 20 testdeltagere. Det kan være nødvendigt at teste 100 eller flere testdeltagere for at opnå et rimeligt snævert konfidensinterval.
- ▶ En beregner for konfidensintervaller til usabilityformål findes på [measuringu.com](http://measuringu.com)

## Konfidensinterval – Praktisk anvendelse

Figuren viser konfidensintervaller for udførelsestid for kerneopgaven “Lej en bil” på et biludlejningswebsted med 20 testdeltagere.

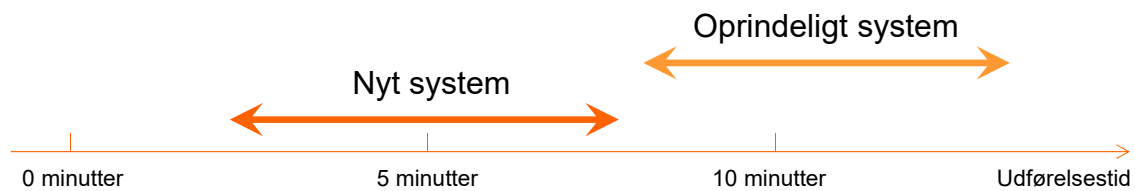
Er det nye system mere effektivt end det oprindelige system?



## Konfidensinterval – Praktisk anvendelse

Figuren viser konfidensintervaller for udførelsestid for kerneopgaven “Lej en bil” på et biludlejningswebsted med 40 testdeltagere.

Er det nye system mere effektivt end det oprindelige system?

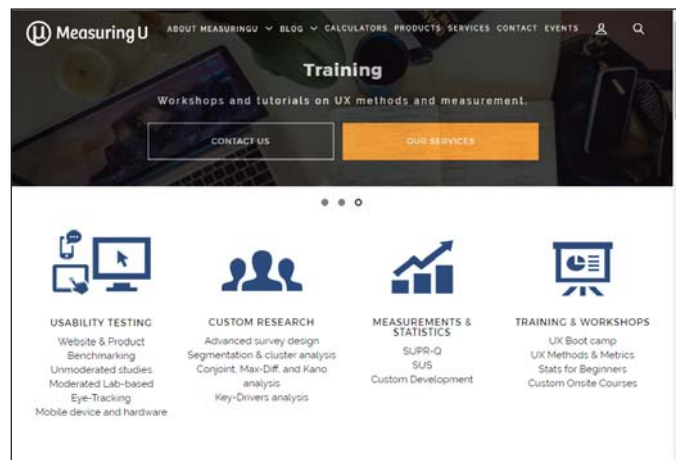


## Statistiske begreber

- Rapportér altid både gennemsnit og konfidensinterval.

Der er stor forskel mellem måleresultaterne 10, 110, 210 sekunder og 100, 110, 120 sekunder, selv om gennemsnittet er det samme

- Oplysninger om statistiske begreber i forbindelse med usabilitytest og nyttige beregningsværktøjer, finder du på Jeff Sauros website [measuringu.com](http://measuringu.com)





## Sammenfatning – Kvantitativ usabilitytest

- ▶ For at opnå pålidelige resultater fra kvantitative usabilitytest er omhu nødvendig. Kvantitativ usabilitytest er langt mindre fejltolerant end kvalitativ usabilitytest.
- ▶ Måleparametre: Udførelsestid, succesrate, fejlrate, katastroferate, tilfredshed
- ▶ Måleresultater: Succes, fejl, katastrofe, ugyldig
- ▶ 4-6 testdeltagere er for det meste tilstrækkelige i en kvalitativ test. Mindst 20 testdeltagere er nødvendige i en kvantitativ test for at opnå et acceptabelt konfidensinterval.
- ▶ Rapportér altid både gennemsnit og konfidensinterval, f.eks.  
“Den gennemsnitlige tid for at bestille en flybillet ligger med 95% sikkerhed mellem 102 og 159 sekunder”