

5. Brugbarhedsvurdering 3



- Empirisk arbejde og forskning
 - Generalisering og konklusioner
 - Gentagelse af forsøg
- Saml og opsummer data
- Fortolkning
 - Identificering af problemerne
 - Kategorisering af problemerne

Hvordan kan man bevise noget?



- Argumentation
 - Standardiserede principper for argumentation
 - For eksempel: $(-1)(-1)=1$ kan bevises med "Tertium non datur" (en tredje mulighed findes ikke) og "reductio ad absurdum" (reduktion til absurd konklusion)
- Deduktion (typisk formelt)
 - $x \in \mathbb{N} \wedge x > 7 \wedge x < 9 \Rightarrow x = 8$
- Empirisk arbejde

Observationer → konklusion

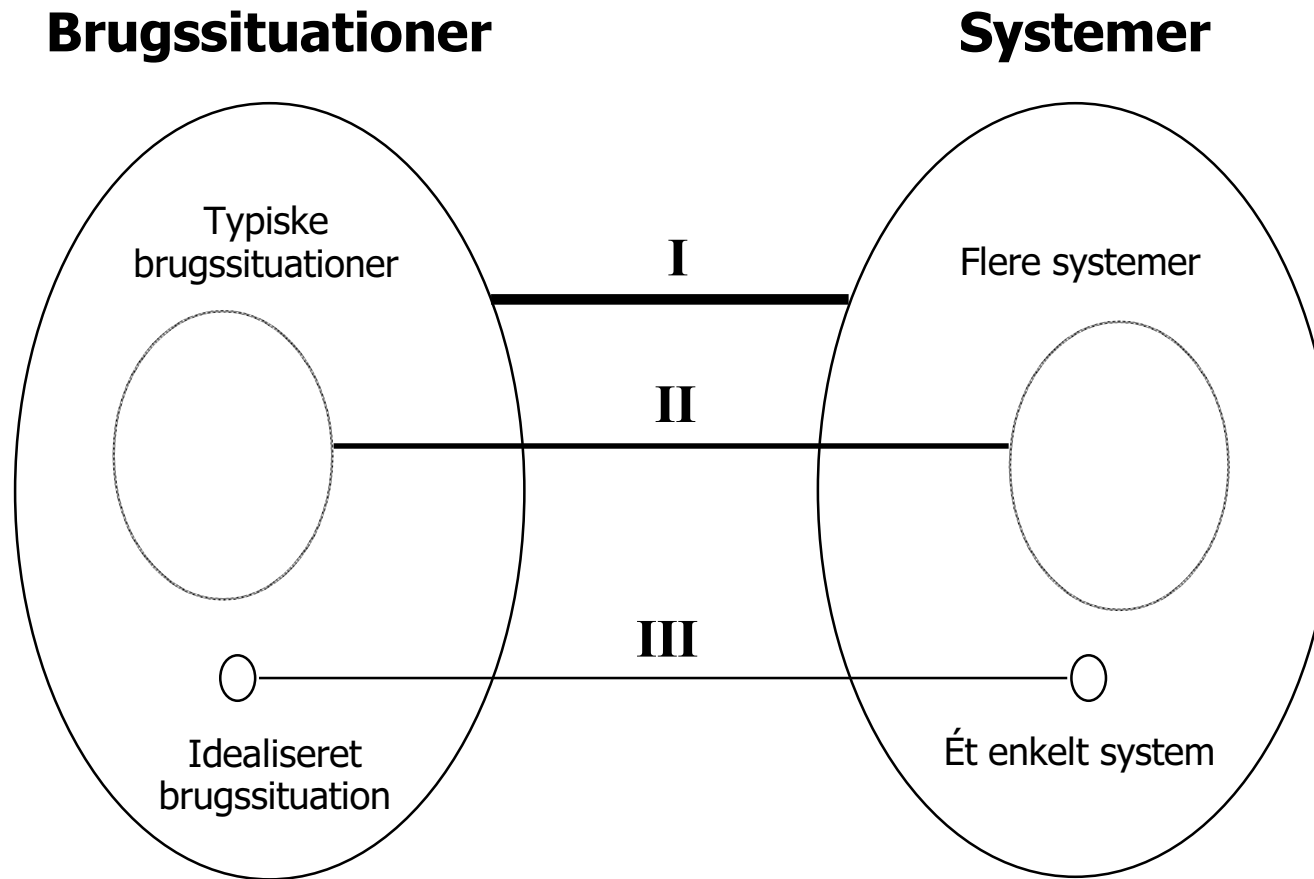
Hvor mange observationer skal vi have?

Beskriv testmetoden



- Objektivitet: vi kan observere virkeligheden uden at påvirke den – resultater er uafhængige af forskeren
Positivism
- Subjektivitet: vi kan ikke observere virkeligheden uden at påvirke den – resultater afhænger af forskeren – virkeligheden eksisterer kun i mit hoved
- Intersubjektivitet: vi påvirker virkeligheden, men vi kan beskrive testsituationen på en måde, så andre kan gentage testen og få lignende resultater
- Derfor: beskriv testsituationen omhyggeligt

Empiri om brugbarhed



- Det ultimative mål er at kunne udtale sig om I.
- Det er vanskeligt at hæve sig op over III.
- Sammenlignende undersøgelser prøver at nærme sig II.
- Reduktionen fra I til II eller III medfører et erkendelsesstab.
- Præsentationer af empiriske resultater må forholde sig til dette erkendelsesstab.
- Vær ydmyge i jeres konklusioner.

Saml data



- Formål: at give os mulighed for at identificere mønstre og identificere problemer med interaktionen
- Opgaver:
 - Skriv notater rent og udbyg med forklaring (mens det huskes)
 - Lav oversigt over forløbet for hver forsøgsperson (regneark) – kaldes også en logfil
- Gør det allerede mens der testes, efter hver dags testning – fordi:
 - Giver os mulighed for at begynde at se mønstre
 - Overblik og check på, at vi ikke har glemt noget
 - Hurtigere i gang med analysearbejdet

Opsummer data

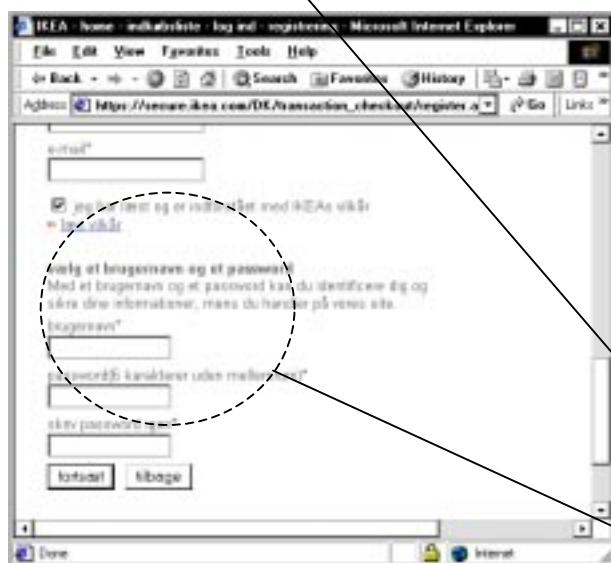
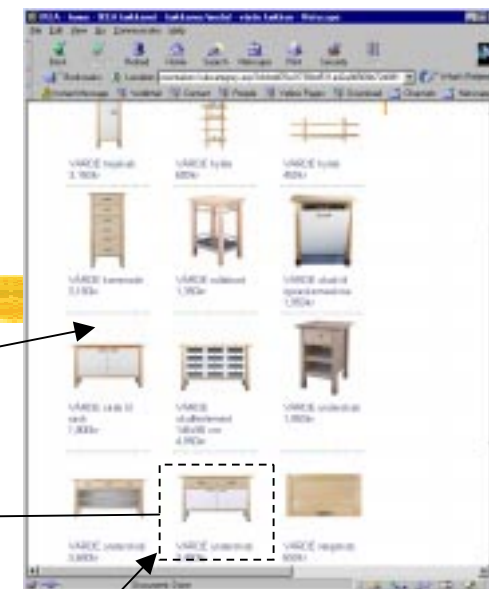
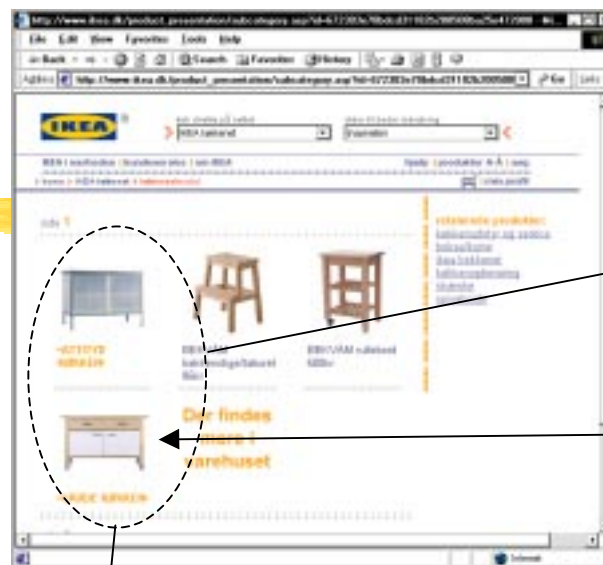
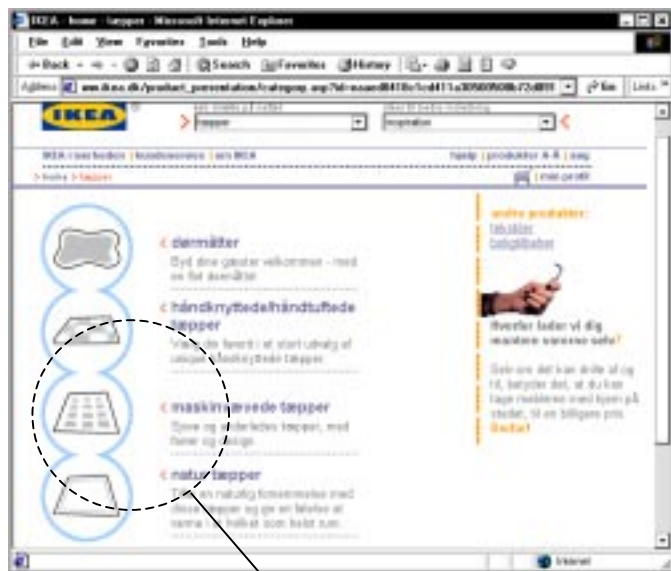


- Statistiske, kvantitative data (tænke-højt vs. spørgeskema)
 - Tid for løsning af opgaver: længste, korteste, gennemsnit og standardafvigelse (gør dette i regneark)
 - Antal fejl: antal brugere, som afsluttede opgaven
 - Brug af hjælpesystem, manual og lignende (f.eks. antal gange)
 - Hjælp fra forsøgsleder nødvendigt for at komme videre
- Mere kvalitative data – præferencer:
 - Hvor godt kan du lide ... (skala)
 - hvilken ... kan du bedst lide (rangering)
 - hvad synes du om ... (åbent)
- Grupper data efter kategorier – f.eks. brugerkategori (næste kursusgang)

Fortolkning



- Tænke-højt
 - En opgave tager meget længere tid at løse end andre opgaver
 - En bruger tager meget længere tid end andre brugere for at løse en opgave
 - Kig efter forskelle og ligheder mellem opgaver, brugere, systemer osv.
- Spørgeskema
 - En opgave bliver vurderet meget anderledes end andre opgaver
 - En bruger vurderer en af opgaverne meget anderledes end de andre brugere
 - Kig efter forskelle og ligheder i vurderingerne mellem opgaver, brugere, systemer osv.
- Identificer steder, hvor brugere laver fejl, hvor de skifter angrebsvinkel, osv.
- Find årsagen til, at der bruges længere tid eller laves fejl – vanskeligt og tidskrævende
- *Opsamling af data om alt dette skal overvejes før forsøget*

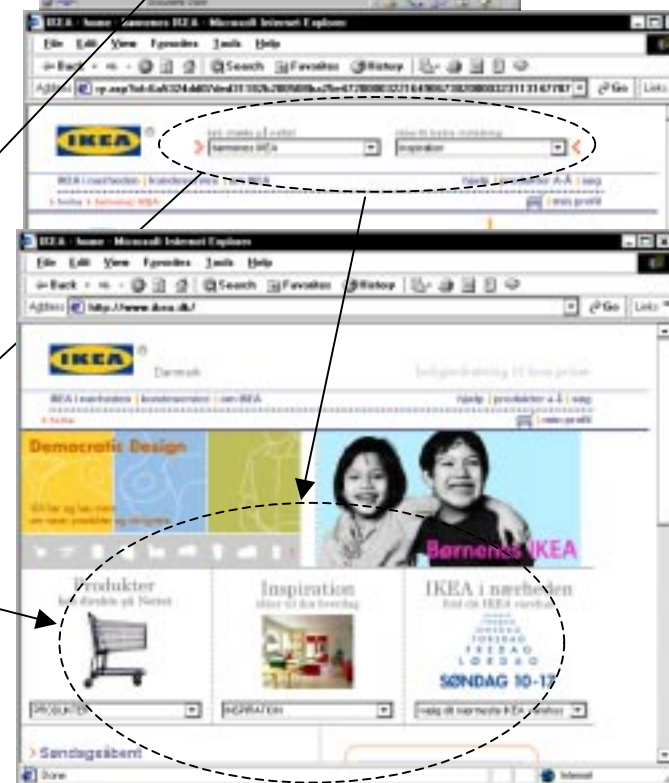


Manglen på forskellige produkter spiller en rolle i opgaveløsningen. Overser muligheden for flere produkter indenfor samme serie. Hvorfor?

Sammenhæng mellem felter. Tror at højre pull-down er undermenu. Løsning?

Mangler faciliteter til brugerens behov, skal finde bestemt størrelse på tæppe

Overser information



Kategorisering af problemer

Brugssituationen

Problemkategorier		Forsinket	Irritation/ Irrationalitet	Forventning vs. faktisk
	Kosmetisk	< 1 minut	Lettere	Mindre forskkel
	Alvorligt	flere minutter	Middel	Betydelig forskkel
	Kritisk	Fuldstændigt (går i stå)	Stærk	Kritisk forskel

Molich, s. 25-26 (brugssituationen), s. 141 (problemkategorier)

IKEA: Kategorisering af problemer

Nedenstående er eksempler på identificerede problemer i brugbarhedstest af Rolf Molich (DialogDesign). Analyser og kategoriser disse tre problemer i forhold til problemkategorier og brugssituationen.

Problem #1: **Kritisk**

To testdeltagere kunne ikke forstå, at deres password ikke virkede. De var ikke opmærksomme på, at web-stedet skelner mellem store og små bogstaver. Den ene testdeltager kunne således ikke logge på med passwordet kaktus fordi han oprindeligt havde indtastet passwordet KAKTUS

Går i stå, kan ikke gennemføre opgaven

Problem #2: **Kosmetisk**

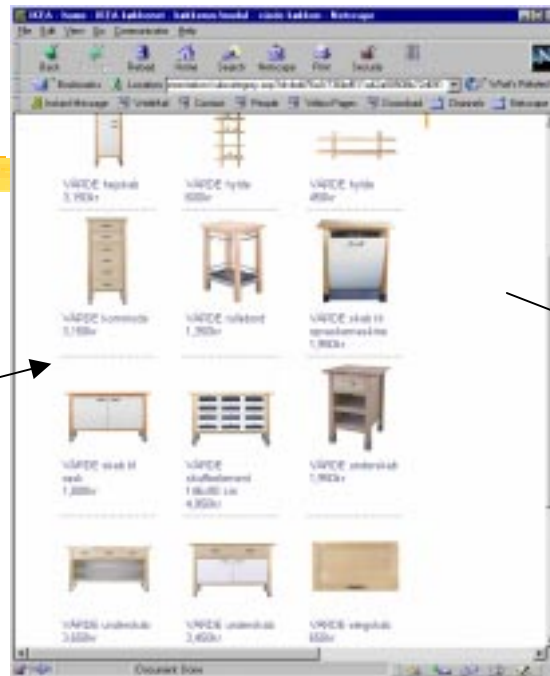
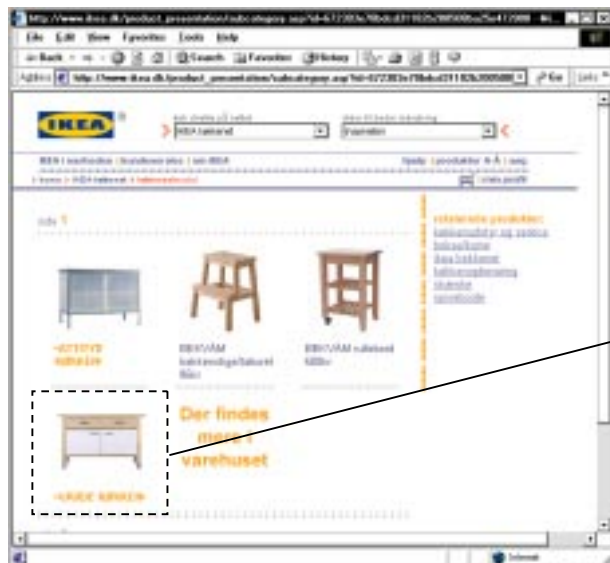
Når testdeltagerne klikkede på Gå til kassen kom de ikke til kassen men til en ny side, som ligeledes bar navnet Indkøbsliste med forsendelsesomkostninger og en knap der hed Gå til kassen.
"Jamen, den knap har jeg jo lige trykket på!"

Bliver sinket i kortere tid, mindre kritisk forskel mellem forventning og handling

Problem #3: **Alvorligt**

Alle testdeltagere havde problemer med at bruge siden Produkter A-Ø, som indeholder en alfabetisk produktoversigt. De fleste testdeltagere opdagede ikke, at de kunne sortere produkterne efter kategori i stedet for produktnavne. Det betød at nogle testdeltagere havde svært ved at finde det de søgte efter.

Sinker brugeren i lidt længere tid men forhindrer ikke gennemførelsen af opgaven



Kritisk, alvorligt eller kosmetisk?



Kritisk, alvorligt eller kosmetisk?