

Estimering av IT-prosjekter: Hvorfor bommer vi og hvordan kan vi bli bedre?



Knowit-seminar, Feb 28, 2013



Magne Jørgensen
Simula Research Laboratory
www.simula.no



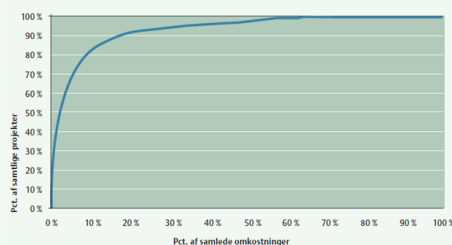
Hvor gode er vi til å estimere?

- Stort sett så er vi ikke så aller verst.
- IKKE tro på rapporter fra Standish Group og andre som ikke har et bevisst forhold til seleksjon av prosjekter.
- Trolig er gjennomsnittlig overskridelse på 20-30%.
- Men, det er noen få store prosjekter som trekker veldig opp dersom vi måler i absolutt overskridelse (se figur).
- Det er forskjell på de "skumle" og de "normale" prosjektene og vi er ikke veldig gode i å anslå hvor skummelt et prosjekt er.



Lille gruppe prosjekter driver store dele af samlede omkostninger i staten

Figur 2.3 Sammenheng mellom de kumulerede omkostninger og antallet af it-prosjekter, n=197



Indikatorer på skumle prosjekter

- Iboende faktorer:
 - Vi gjør ting som er grunnleggende forskjellig fra det vi har gjort før
 - Mange grensesnitt til andre systemer og/eller mange interesserter
 - Involverer stor grad en prosessendring
 - Problemene som skal løses er komplekse
 - Uflaks (mange små eller få store "uflakser")
- Faktorer vi kan gjøre noe med
 - Realisme i ambisjoner
 - Situasjoner og motivasjoner som gir over-optimisme
 - Kompetanse til kunde og leverandør (store forskjeller i kompetanse som ikke fås frem gjennom CV-er og referanser)
 - Oppfølging, støtte, og ledelsesforankring
 - Kommunikasjon, inkludert økt oppmerksomhet rundt kulturelle forhold
 - Anbudsprosesser (unngå "winner's curse" og "adverse selection") – kundefok
 - Egnede utviklingsmodeller (særlig relatert til oppdeling av leveranse)

[**simula** . research laboratory]

Noen funn i ESSU (UK):

- Many projects are over-ambitious.
- Long-term or delayed projects are often overtaken by new technology, changes in legislation and public policy.
- The private sector often overstates its ability to deliver.
- Clients are often under-resourced and/or do not have the required skills.
- The procurement process is a high-risk strategy, heavily influenced by market forces in respect to who bids, the level of competition and private sector strategies to increase market share.
- Some projects are driven by the application of the latest information and communications technology to meet 'customer demand' for seamless one stop contact centres combined with pressure to achieve substantial savings. However, a more incremental approach may be more desirable, effective and economical.
- Off-the-shelf is no guarantee for success. Many projects were based on off-the-shelf products, but nevertheless failed.

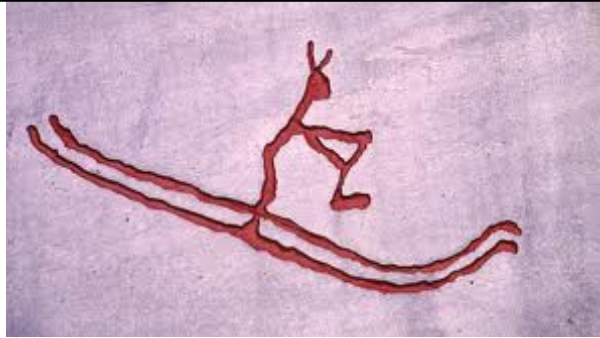


**European Services
Strategy Unit**

(Continuing the work of the Centre for Public Services)

[**simula** . research laboratory]

<http://www.european-services-strategy.org.uk/news/2007/ict-contract-chaos/105-ict-contracts.pdf>



TO SMÅ AVSTIKKERE

AVSTIKKER 1:

ESTIMERING AV KOSTNADER VINTER-OL, OSLO 2022

[**simula** . research laboratory]

Forskning viser:

**Å se forover gjør en optimistisk
Å se tilbake skaper mer realisme**

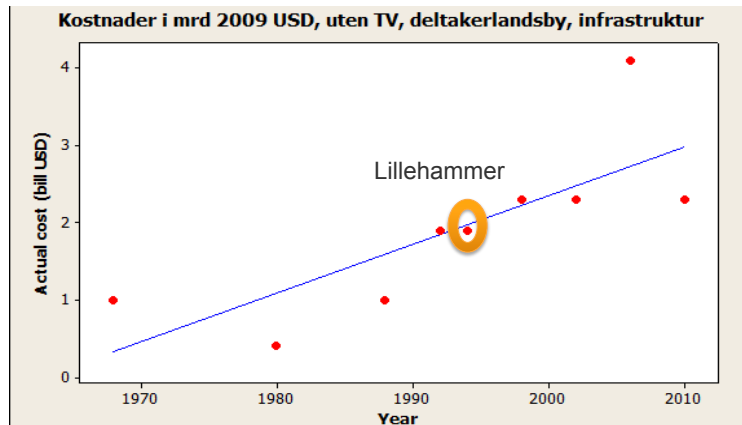
•Vi ser derfor litt tilbake på tidligere vinter-OL

- Utgangspunkt i "bid" og "operational costs incurred by the Organising Committee for the purpose of staging the Game". Kostnader til deltagerlandsby, fjernsyn/presse-tjenester, veier, hoteller m.m. IKKE inkludert.
- Omregnet til 2009 valuta-kurs i OL-landet, så til 2009-USD.
- Mange vinter-OL mangler pålitelige data
- Basert på data fra Flyvbjerg & Stewart, University of Oxford, 2013

[**simula** . research laboratory]

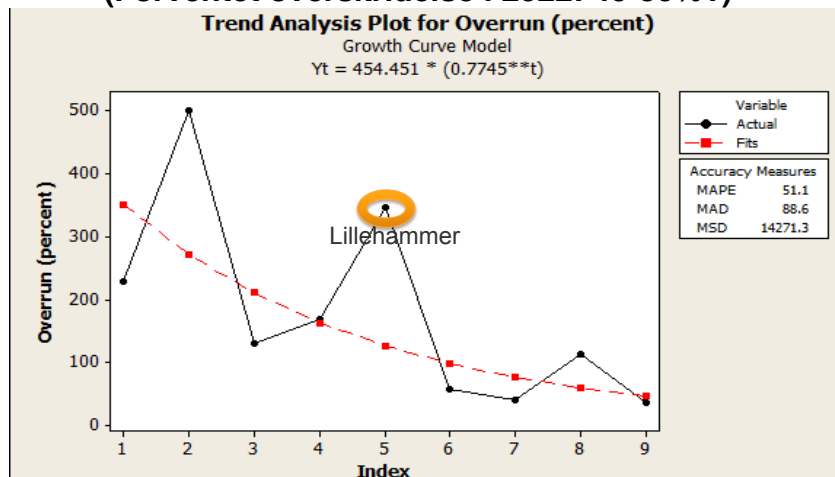
- by thinking constantly about it

**Antatt linear utvikling i kostnader så vil OL i 2022 koste ca. 20 mrd 2009-NOK (30 mrd 2022-kr?)
95% "sikkert" at det vil koste mindre enn 35 mrd. 2009-kr**



Eksempel på analogi-basert estimering

Overskridelser i olympiske vinterleker (1968-2010) (Forventet overskridelse i 2022: 40-50%?)



*Estimerer fra forrige runde for tre norske byer var mellom 12 og 16 mrd (?).
Tar vi 12 mrd (Oslo?), ganger med forventet overskridelse på ca. 50% + litt prisstigning,
får vi ca. 20 2009-kroner mrd ... (Eksempel på fornuftig justering av ekspertestimat)*

AVSTIKKER 2:

HVA SKAPER PROSJEKTPROBLEMER?

[**simula** . research laboratory]

Table 7: Results from the logistic regression model

Predictor variable	Coefficient	p-value	Odds ratio	95% confidence interval	
				Lower	Upper
Constant	-2.90	0.00			
SatisfactionScoreProvider=Low	0.35	0.00	1.42	1.39	1.45
SatisfactionScoreProvider=Unknown	0.91	0.00	2.49	2.33	2.67
FailureRateProvider=Low	-0.66	0.00	0.52	0.51	0.53
FailureRateProvider=Unknown	-0.34	0.00	0.71	0.67	0.76
SkillTestPassRateProvider=Low	0.07	0.00	1.07	1.02	1.12
SkillTestPassRateProvider=Unknown	0.58	0.00	1.79	1.74	1.85
SatisfactionScoreClient=Low	0.18	0.00	1.20	1.17	1.23
SatisfactionScoreClient=Unknown	0.25	0.00	1.28	1.23	1.33
FailureRateClient=Low	-0.64	0.00	0.53	0.52	0.54
FailureRateClient=Unknown	-0.63	0.00	0.53	0.51	0.56
PreviousCollaboration=Yes	-1.74	0.00	0.17	0.17	0.18
FocusLowPrice=Low	-0.19	0.00	0.83	0.81	0.85
FocusLowPrice=Medium	-0.08	0.00	0.92	0.89	0.95
FailureRateProviderRegion="High"	0.27	0.00	1.31	1.28	1.33
FailureRateClientRegion="High"	0.42	0.00	1.53	1.48	1.58
GeographicalDistance="neighbor"	-0.07	0.02	0.93	0.90	0.97
GeographicalDistance="offshore"	0.02	0.10	1.02	1.00	1.05
logSize	0.71	0.00	2.03	1.99	2.06

[**simu**

Hva kommer overoptimismen av?
Hva bør vi gjøre for å treffe bedre med
estimatene?

Overvurdering av egne evner er normalt
hos mentalt friske mennesker!



Undervurdering av risiko er normalt

(Risiko for egen skilsmisse er for eksempel sterkt undervurdert selv om de fleste vet at nær 50% av ekteskap sprekker)



[**simula** . research laboratory]

13



Biologi



- Evolusjonen synes å ha belønnet selvsikkerhet og risikoundervurdering.
- IT-ledere belønner selvsikkerhet og bruker den som indikator på dyktighet.
- En del indikasjoner på sammenheng mellom optimisme og biologi. F eks, så hjelper trolig optimisme på en del sykdomsforløp (styrking av immunforsvar), og håndtering av problematiske situasjoner ("coping").
- Simulering indikerer at overoptimisme i noen sammenhenger gir systematisk bedre utfall enn realisme. Dette gjelder særlig når man vet lite om sannsynligheten for ulike utfall, men mer om konsekvensene av utfallene.

[**simula** . research laboratory]

14

Grupper og optimisme

- Mange studier viser en økt risikovilje og over-sikkerhet (over-confidence), kanskje pga individenes reduserte ansvar for beslutningene.
- Våre studier på IT-prosjekter indikerer imidlertid mer realisme ved bruk av grupper til å estimere!
 - Flere hoder husker på mer (union av aktiviteter)
 - Krav om begrunnelse kan gi økt realisme
 - Noen ganger forekommer imidlertid "group think" og dominerende individer
 - F eks at **alle** i gruppen går ut av en estimeringsdiskusjon med opplevelsen av at estimatet ble for lavt.
- Effekten av grupper på realisme trolig svært kontekstavhengig, dvs vanskelig å finne generelle mønstre.
- Bruk av Delphi-baserte metoder (strukturerte grupper med uavhengige vurderinger) gir stort sett svært bra innvirkning på realisme.
 - Som for eksempel "Planning Poker" eller Wideband Delphi

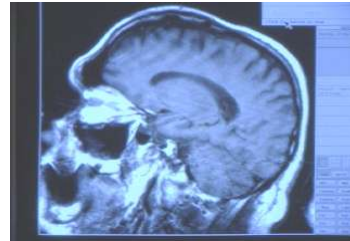
Motivasjon

- Mange resultater viser sterk sammenheng mellom motivasjon for høy ytelse (ønsketenkning) og overoptimisme.
- Optimisme kan påvirke ytelsen positivt, MEN
 - Optimistiske estimater synes typisk å ha kun en kortvarig positiv effekt på gjennomføringen.
 - Hvor stor og hvor positiv effekten av optimisme er på ytelse overvurderes oftest sterkt.



Kognitive prosesser

- Planlegging (visualisering av fremtidsscenarier) gir ofte mindre grad av realisme enn historierefleksjoner (hvordan det har gått)!
 - Analogi-basert estimering gir ofte gode resultater
- Overvurdering av hvor mye vi kan kontrollere/håndtere er en viktig kilde til overoptimisme (the illusion of control)
 - Dyktighet vs uflaks er ofte et spørsmål om man lykkes eller mislykkes



En liten oppgave ...

- Anta at du har estimert et prosjekt til mest sannsynlig å ta 1000 timeverk og skal legge inn et pristilbud til kunde. Det er flere (5-10) andre firma også som vil komme til å legge inn tilbud på samme prosjekt.
- Anta to ulike konkurransegrunnlag-situasjoner:
 - S1: Pris vektlegges 70% og kvalitet 30% (pridfokusert kunde)
 - S2: Pris vektlegges 30% og kvalitet 70% (kvalitetsfokusert kunde)
- Anta at du kun vil ha oppdraget dersom det vil gi et overskudd. Et overskudd krever at prosjektet ikke tar med enn maksimum 800 timeverk, som vi for enkelhets skyld kan si koster dere 800.000 kr.
- Vil du anbefale å gi samme tilbud til kunde uavhengig av S1 eller S2, dvs pris eller kvalitetsfokusert kunde?
- Hvis ikke, når er det riktigst å være mest forsiktig i tilbudet? Hvorfor?

Kunder velger over-optimistiske leverandører

- "Winner's curse"
 - Vektlegging av pris + mange tilbydere: "Vi vinner nesten bare når vi har vært over-optimistiske".
 - Kan gjøre gode leverandører dårlige
- Manipulerende informasjon
 - Budsjettinformasjon (Lånekassens IT-system hevdes å ha blitt estimert av stortingspolitikere, som bestemte budsjettet).
 - Bruk av ladede ord ("lite system")
- "Adverse selection"
 - Jo mindre kompetanse hos leverandør, jo lavere estimat og pris
 - Jo mindre kompetanse hos kunde, jo mer vektlegging av lav pris
 - Mye av prosjektproblemene og overskridelse skyldes prislefokus i valg av leverandør.

[**simula** . research laboratory]

19

Winner's curse formelen ...

$$(6) \text{ rel} = \frac{\text{act} - \text{est}}{\mu} = \frac{\mu \left(1 - \rho_{\text{est}, \text{act}} \frac{\sigma_{\text{act}}}{\sigma_{\text{est}}} \right) (1 - w)}{\mu} = \left(1 - \rho_{\text{est}, \text{act}} \frac{\sigma_{\text{act}}}{\sigma_{\text{est}}} \right) (1 - w).$$

[**simula** . research laboratory]

Desto mer fokus på lav pris, desto høyere forventet overskridelse (ROSING-formelen)



$$RO = S \cdot IN + G$$

RO = Relative cost Overrun

S = Selection bias (= focus on low price)

IN = INaccuracy

G = General estimation over-optimism

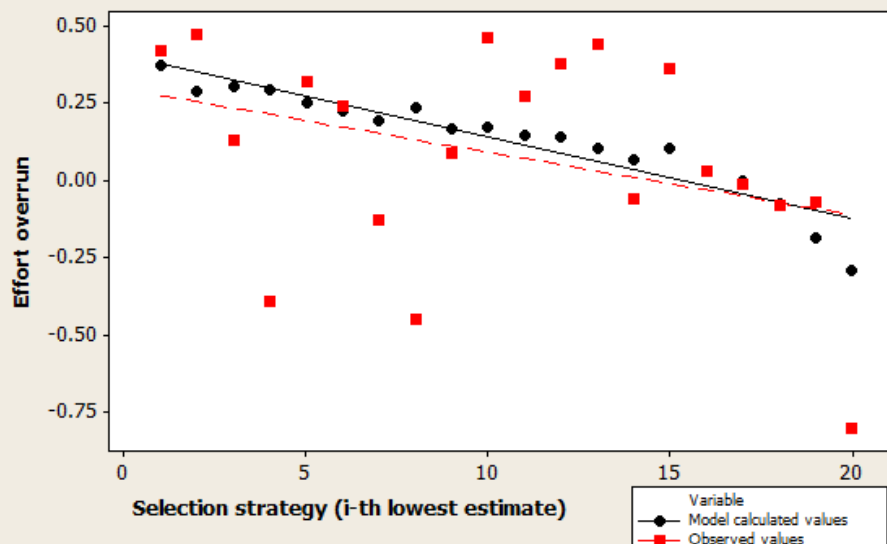


Eksempel:

- Kunde velger et tilbud som er 40% lavere enn det gjennomsnittstilbudet ($S = 0.4$)
- Korrelasjon mellom estimert og faktisk kostnad er 70% ($IN = 100\% - 70\% = 0.3$)
- Leverandørene er i gjennomsnitt 20% over-optimistiske i sine estimater ($G = 0.2$)
- Forventet kostnadsoverskridelse for valgt leverandør er 32% ($0.4 \cdot 0.3 + 0.2$)
- **Sterkere prisfokus fører til høyere S og høyere forventet overskridelse!**

[**simula** . research laboratory]

(Mer om dette på: <http://simula.no/publications/Simula.simula.1478>)



[**simula** . research laboratory]

Ankereffekter

Eksperiment:

- HIGH (LOW) group: *“The customer has indicated that he believes that **1000 (50)** work-hours is a reasonable effort estimate for the specified system. However, the customer knows very little about the implications of his specification on the development effort and you shall not let the customer’s expectations impact your estimate. Your task is to provide a realistic effort estimate of a system that meets the requirements specification and has a sufficient quality.”*
- Deltakere: Erfarne systemutviklere.
- Alle (HIGH, LOW, CONTROL) fikk samme kravspesifikasjon.

Ankereffekter

- Resultater:
 - HIGH group gjennomsnitt: 555 timeverk
 - CONTROL group (uten forventinger) gjennomsnitt: 456 timeverk
 - LOW group gjennomsnitt: 99 timeverk!!!
- Ingen av utviklerne opplevde at de hadde blitt mye påvirket. De fleste mente at de ikke var påvirket i det hele tatt av kundens forventninger.

Feltstudie av det samme

“The preliminary budget of the new system is \$10 000 [corresponding to about 100 work-hours with typical pricing in the country in which it will be built]. The preliminary budget is not built on any knowledge about the actual cost of developing the new system, and will, if needed, be extended to cover the expenses necessary to build a quality system with the desired functionality.”

100 timeverk var en svært lav verdi for dette prosjektet og alle firmaene ble instruert om å IKKE bruke dette som input til estimatet.

Firmaene som deltok visste ikke at de var med i en studie, men skulle gi uavhengige estimater på et arbeid som skulle utføres **av andre (mao ingen direkte grunn til å være optimistiske)**. Firmaene ble betalt for estimeringsarbeidet.

Resultater fra feltstudien

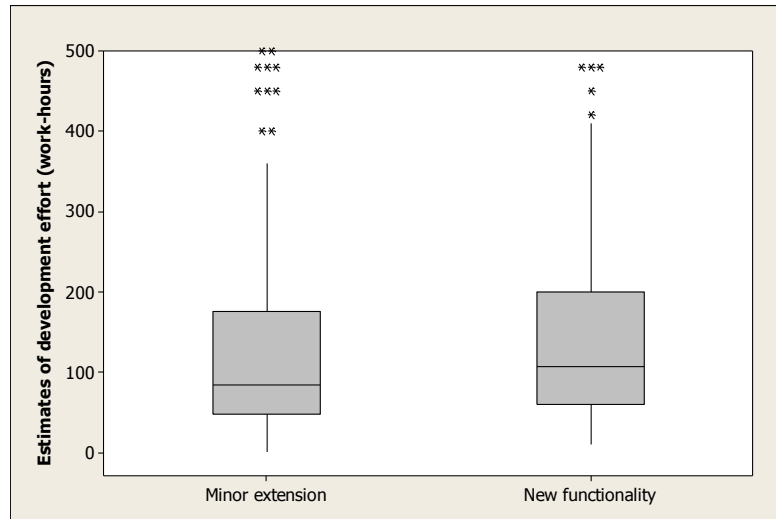
Numerical Anchor

Group	Median estimate
Manipulated (client's expectation, 100 work-hours)	724 work-hours (n=23)
Ordinary	956 work-hours (n=23)

Mindre effekt enn i ”laboratoriet”, men fortsatt en effekt av betydning.

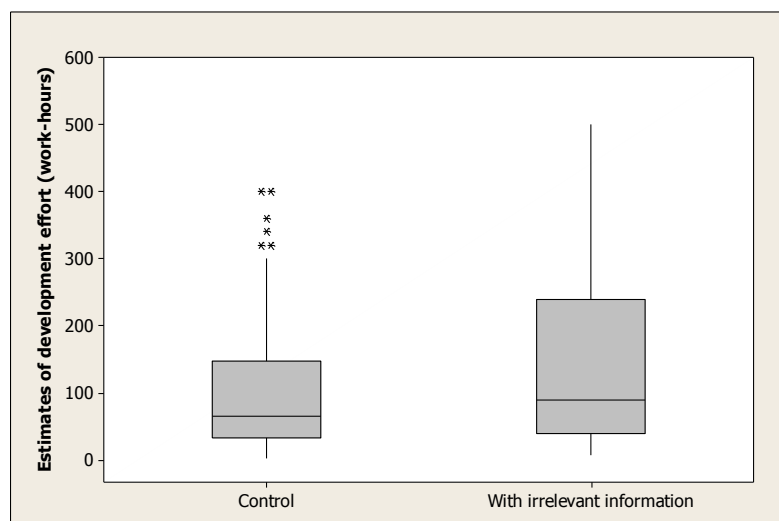
Trolig større effekt dersom man er i tilbuds-modus.

Ladete ord kan også fungere som anker



[**simula** . research laboratory]

Irrelevant informasjon påvirker estimatet



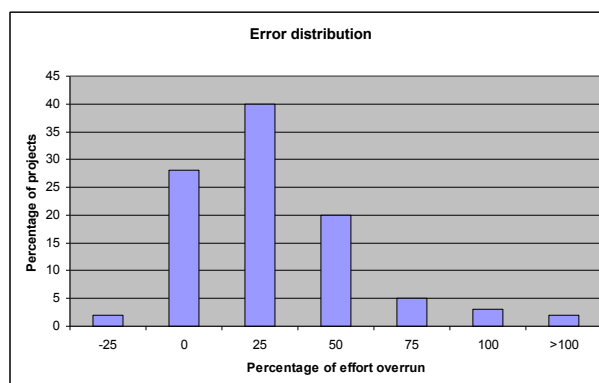
[**simula** . research laboratory]

Usikkerhetsvurderinger

- Minimum-maksimum intervall: “Nesten helt sikker” = 60-70% sikkert.
- Konfidensnivå betyr minimalt for bredden på intervallene.
 - Eksperiment med fire grupper av utviklere (A: 50%, B: 75%, C: 90, D: 99% konfidensnivå)
 - Nesten samme minimum-maksimum intervaller ble angitt.
- Meningsløst å be om minimum-maksimum intervaller.
- Be i stedet om:
 - Angi en fordeling av estimeringsfeil for lignende prosjekter.
 - Bruk denne til å anslå usikkerhet.
 - Eksempel: Dersom 80% av tidligere prosjekter har overskredet estimatet med mindre enn 50%, så kan du være 80% sikker på at et budsjett som er lik 1,5*estimatet ikke vil bli overskredet.
 - Hovedfordel er fornuftig bruk av historiske data og unngåelse av ønsketenkning. Evaluert i reelle prosjekter med god virkning.

[**simula** . research laboratory]

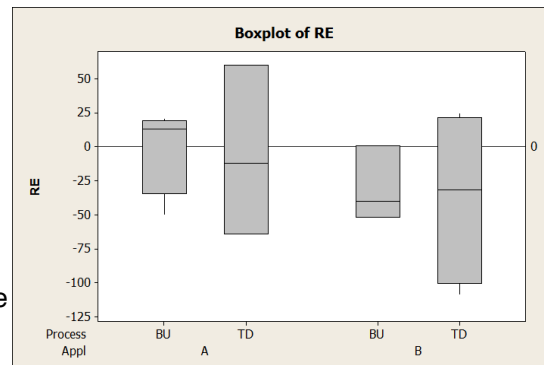
En enkel metode for å bli bedre på usikkerhetsvurderinger: Hvor mye bommet du tidligere på lignende oppgaver?



[**simula** . research laboratory]

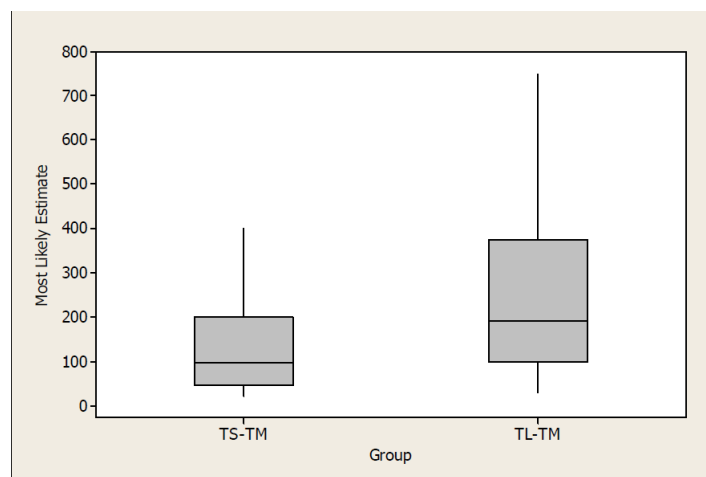
Top-down eller bottom-up

- Ja takk, begge deler.
- Studie viste at dersom gode analogier ble funnet, så var top-down den beste.
- I gjennomsnitt var imidlertid bottom-up best.
- Overraskende funn: Estimeringsteamene var relativt dårlige i å finne og bruke lignende prosjekter.
- Kan være stort potensiale for opplæring og bruk av relativ estimering på prosjektnivå – a la relativ estimering på user story nivå i agile.



[**simula** . research laboratory]

Sekvensseffekt



TS-TM: Estimerte først et lite system (ca. 25 tv), så det middels store.

TL-TM: Estimerte først et relativt sett stort system (ca. 500 tv) så det middels store.

[**simula** . research laboratory]

Kan vi prediktere hvem som vil være mest realistisk?

- Få robuste resultater.
- Eneste faktor som er noenlunde bra (men heller ikke den veldig bra) er tidligere estimeringsnøyaktighet
- Uklart hva estimeringsekspertise egentlig er.

[**simula** . research laboratory]

ØVELSESOPPGAVER

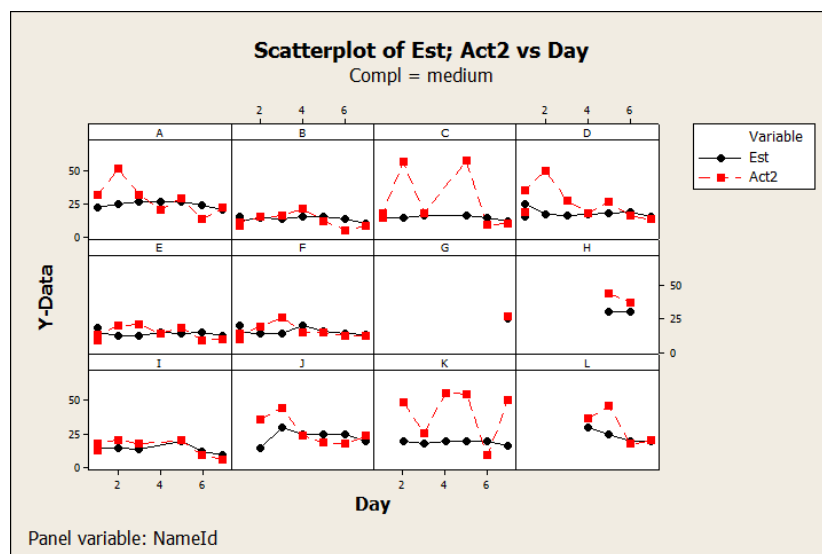
[**simula** . research laboratory]

Hva får vi når vi ber om et estimat?

- Planlagt arbeidsmengde (med risikobuffer)?
- Median arbeidsmengde (50% sannsynlighet for å overskride)?
- Mest sannsynlig arbeidsmengde?
- Mest sannsynlig arbeidsmengde under ideelle forutsetninger?

[**simula** . research laboratory]

Estimat synes typisk å gjenspeile “Hvis ingen ting går galt”



[**simula** . research laboratory]

Flere studier viser at ...

- Det er ofte liten eller ingen forskjell i prediksjoner gitt under idelle og realistiske scenarier.
- Dette har blant annet vært undersøkt innen:
 - Tidsestimering
 - Blodgiving
 - Treningsmengde
 - Sparing

[**simula** . research laboratory]

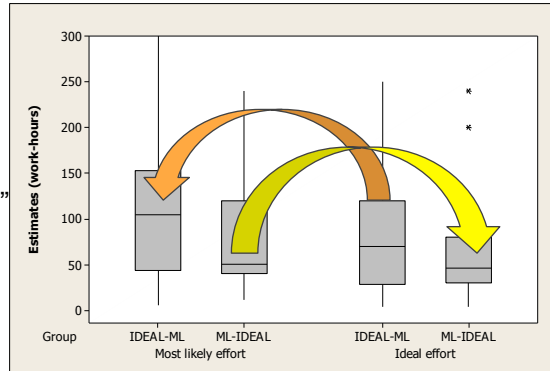
Feil spørsmål, eller feil svar?

- Hva gjør man når respondentene ikke svarer på det man spør om, men på noe annet?
 - Endrer respondentenes oppførsel gjennom bedre opplæring, bedre instruksjoner, etc?
 - Endrer spørsmålet til å spørre om det som respondentene uansett gir svar på?
- Vi har evaluert en metode basert på det siste alternativet.

[**simula** . research laboratory]

Eksperiment 1

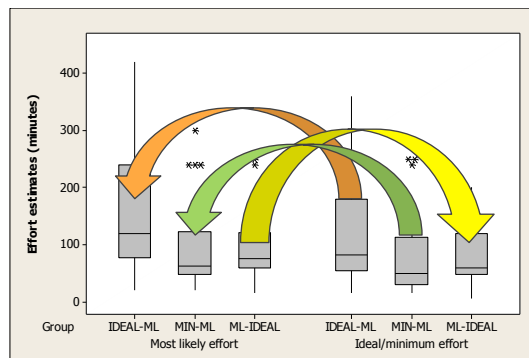
- IDEAL-ML: Først ideell tid, deretter mest sannsynlig tid.
- ML-IDEAL: Først mest sannsynlig, så ideell tid.
- Ideell tid: *“Antall timeverk antatt full konsentrasjon, uten forstyrrelser og fullt ut produktiv.”*
- Faktisk tid: Ca. 100 timeverk.
- IDEAL-ML mest realistisk.
- IDEAL samme som ML.
- 25% i ML-IDEAL gruppen økte ikke estimated fra ML til IDEAL.



[**simula** . research laboratory]

Eksperiment 2

- IDEAL-ML: Først IDEAL, så ML.
- ML-IDEAL: Først ML, så IDEAL.
- MIN-ML: Først MIN, så ML.
- Realistisk tid: Ca. 140 minutter.
- IDEAL-ML mest realistisk.
- IDEAL samme som ML.
- MINIMUM først hadde IKKE samme effekt som IDEAL først.



[**simula** . research laboratory]

IDEAL->ML metoden

1. Be om et estimat under antagelsen av ideelle betingelser.
 - a. Ulike formuleringer av hva som menes med ideelle betingelser synes å gi noenlunde samme effekt.
 - b. Må skille seg tydelig fra realistiske betingelser.
2. Be deretter om et estimat av hva arbeidet realistisk sett krever, f eks i form av mest sannsynlig antall timeverk).
 - a. Vi har tidligere funnet at mer risikoanalyse kan gi mer optimistiske estimer. Denne metoden er trolig mer motstandsdyktig mot denne effekten, men det kan likevel være lurt å splitte denne estimeringen i to deler:
 - I. Fase 1: Estimer realistisk arbeidsmengde for kjente aktiviteter og kjent risiko.
 - II. Fase 2: Estimer tillegg for ukjente aktiviteter. Baser dette på historikk over andel timeforbruk på ikke-planlagte aktiviteter og ikke-identifiserte risiko.

[**simula** . research laboratory]

[**simula** . research laboratory]

**DU ER MER LIK MEG
ENN JEG ER LIK DEG**

**ASYMMETRI I RELATIV
ESTIMERING**

Estimering av arbeidsmengde er alltid relativt til noe annet (“Alt er relativt”)

Sammenligning kan være eksplisitt:

1. Hvor mye større/mindre (i timeverk eller %) er dette prosjektet i forhold til det vi gjorde tidligere av samme type?
2. Hvor mange “story points” er “user story” X? (sammenligning mot en referanse (baseline) user story)
3. Hvilken kategori oppgave er dette? (sammenligning mot en “arketype”, f.eks. en typisk “liten”, “middels” og “stor” oppgave)

Sammenligning kan være implisitt/ubevisst:

1. Estimering av en oppgave rett etter en annen fører ofte til at den første oppgaven blir en stor del av sammenligningsgrunnlaget (uten at vi nødvendigvis merker det)
2. Ledende teorier innen psykologi tilsier at vi vil ubevist aktivere og sammenligne med tidligere erfaringer, og at dette gir vår “magefølelse”. Et førsteinntrykk dannes ofte veldig raskt og kan være vanskelig å endre.

[**simula** . research laboratory]

Lønner det seg å hoppe etter Wirkola?



[**simula** . research laboratory]

Ja, det lønner seg å hoppe etter Wirkola (gitt at du kan hoppe på ski)

- Det gir stort sett bedre evalueringer å kommer etter, dvs å bli sammenlignet med, gode prestasjoner
 - Resultater blant annet fra muntlig eksamen og turn
- Den dominerende effekten er “assimilasjon” og ikke “kontrast”
 - MEN, noen ganger (gjerne ved ekstreme egenskaper hos den ene) blir det en kontrasteffekt, dvs en skjevhet mot større ulikhet
- Innen estimering av IT-oppgaver finner vi at å estimere en liten oppgave etter en stor gjør den lille oppgaven større. Tilsvarende blir en stor oppgave mindre dersom den sammenlignes med en liten oppgave.
- I denne presentasjonen er det **ikke** denne effekten jeg har fokus på, men om dragningen mot “likhet” (assimilasjonen) er større enn den andre veien ved sammenligninger, dvs om det er asymmetri.

[**simula** . research laboratory]

Ulike måter å spørre/tenke på

- Anta at vi har to systemutviklingsoppgaver: Utvikling av System A (lite system) og System B (større system). En sammenligning kan være basert på:
 - Alternativ 1: “Hvor mye større er B enn A?” (A er referanse)
 - Alternativ 2: “Hvor mye mindre er A enn B?” (B er referanse)
 - Alternativ 3: Estimer A, så B (A blir trolig referanse)
 - Alternativ 4: Estimer B, så A (B blir trolig referanse)
- Vil vil trolig få en “assimileringsseffekt” i alle tilfellene, MEN, gir alternativene 1-4 ulike svar på forskjellen mellom A og B?
- Det kan være gode grunner for å velge referanse/spørremåten som forventes å gi minst assimilasjonseffekt
 - Ref.: Dette er relatert til diskusjonen om vi bør bruke en liten, middels eller stor oppgave som referanse i Story Points-basert estimering.

[**simula** . research laboratory]

Participants should be able to register their participation at the summer party on the web by submitting their name and their email address. The layout should be as in the figure below. The system confirms (on web) that the data is registered. There is no data validation (duplicate check, etc). The data is stored in a database. Generation of reports, such as attendee lists, is done manually, i.e. by querying the database.

Summerparty 2011

Name 

Your E-Mail Address 



System B specification: Registration of Scientific Articles

Simula wants a web-based library system that contains information about scientific articles. Users and administrators must be able to view an information page about each scientific article that is registered in the system, search for articles, see a printer-friendly display of the search results, register new scientific articles (data validation must be done during registration), and perform user management (administrators can register, edit and remove other administrators and users).

simula research laboratory		
Main Page		
First Author	Year	Title
Holm-Mannstam, T. A.	1989	The impact of goals on software project management: the experimental investigation
Holm-Mannstam, T. A.	1993	The dynamics of software project scheduling
Holm-Mannstam, T. A.	1995	The impact of software project scheduling on software project management

Example of an article information page

[illegible]

Estimering av systemutviklingsoppgaver

- Deltagere: Ca. 100 utviklere fra Ukraina
- Reell forskjell i arbeidsmengde (B-A) var på minst 100 timeverk.
- Alternativ 1: Hvor mye større er B enn A? (A er referanse)
 - Svar (gjennomsnitt): 80 timeverk
- Alternativ 2: Hvor mye mindre er A enn B? (B er referanse)
 - Svar (gjennomsnitt): 40 timeverk
- Alternativ 3: Estimer A, så B (A er trolig referanse for B)
 - Svar (gjennomsnitt): Diff. på 77 timeverk
- Alternativ 4: Estimer B, så A (B er trolig referanse for A)
 - Svar (gjennomsnitt): Diff. på 68 timeverk

Mao, forskjellen blir større når vi ser en stor oppgave i forhold til en liten, enn motsatt. Assimileringsseffekt.

[**simula** . research laboratory]

Estimering av antall innbyggere

- Spørsmål om antall innbyggere i USA vs. Mexico (Riktig svar er ca. 200 mill forskjell i antall)
- Hvor mange flere innbyggere har USA enn Mexico? (Mexico referanse)
 - Svar (gjennomsnitt): 150 mill
- Hvor mange færre innbyggere har Mexico enn USA? (USA er referanse)
 - Svar (gjennomsnitt): 100 mill
- Antall innbyggere i Mexico, så antall innbyggere i USA (Mexico er trolig referanse)
 - Svar (gjennomsnitt): Diff på 150 mill
- Antall innbyggere i USA, så antall innbyggere i Mexico (USA er trolig referanse)
 - Svar (gjennomsnitt): Diff på 100 mill

[**simula** . research laboratory]

To oppgaver (A og B) av lignende størrelse

- Hvor mye mer/mindre (strek under det som passer) arbeid krever B sammenlignet med A? (A er referanse)
 - Gjennomsnitt: B krever mest (2 tv mer)
- Hvor mye mer/mindre (strek under det som passer) arbeid krever A sammenlignet med B? (B er referanse)
 - Gjennomsnitt: A krever mest (24 tv mer)
- Estimer A, så B (A er trolig referanse for B)
 - Gjennomsnitt: B krever mest (66 tv mer)
- Estimer B, så A (B er trolig referanse for A)
 - Gjennomsnitt: A krever mest (70 tv mer)

En oppgave øker i størrelse (målt mot en annen) i rollen som målobjekt og/eller minker i rollen som referansen. Dette er trolig den samme effekten som vi så på stor-liten oppgavene.

[**simula** . research laboratory]

To land med nokså likt antall innbyggere

(Polen vs Romania: Polen har 17 mill. Flere innbyggere enn Romania)

- Hvor mange flere/færre (strek under det som passer) innbyggere er det i Polen enn i Romania? (Romania er referanse)
 - Polen har 4 mill flere
- Hvor mange flere/færre (strek under det som passer) innbyggere er det i Romania enn i Polen? (Polen er referanse)
 - Polen har 2 mill flere
- Estimer Romania, så Polen (Romania er trolig referanse for Polen)
 - Polen har 10 mill flere
- Estimer Polen, så Romania (Polen er trolig referanse for Romania)
 - Polen og Romania like store

Som for IT-utvikling, så øker et lands innbyggere i antall (målt mot et annet) i rollen som målobjekt og/eller minker i rollen som referanse.

[**simula** . research laboratory]

Tversky's forklaring

- Sammenligning skjer vha "pattern matching". Denne tenderer til neglisjering/undervektning av egenskaper ("features") som finnes hos referansen, men ikke hos målobjektet.
- Dersom målobjektet er mindre enn referansen (kanskje til og med omfattes av referansen mhp features) eller vi har mindre kunnskap om målobjektet, så vil likheten til referansen øke.
 - Dette fordi det meste av ulikheten ligger i de neglisjerte/undervektede egenskapene.
- Dersom målobjektet er større enn referansen eller vi har mer kunnskap om målobjektet, så vil ulikheten til referansen øke.
 - Ulikheter som neglisjeres/undervektes vil være mindre fremtredende her.

[**simula** . research laboratory]

Noen konsekvenser

- Du er mer lik meg, enn jeg er lik deg
- Nord-Korea er mer lik Kina, enn Kina er Nord-Korea
- Din ektefelle/kjæreste vil komme bedre ut dersom du sammenligner henne med "idealkvinnen" enn om du sammenligner "idealkvinnen" med henne
- Årsak X (f eks til prosjektfiasco) blir viktigere i forhold til Y når vi sammenligner X mot Y, enn Y mot X
 - Dersom du vil at svaret skal være i din favør som prosjektleder bør du sørge for at vurderingen er: "Hvor mye av problemene skyldes kundesiden i forhold til prosjektledelse hos leverandøren?" og ikke "Hvor mye av problemene skyldes prosjektledelse hos leverandøren i forhold til kundesiden?"
- Jo mindre vi vet om det vi estimerer, jo mer ligner det på det vi mer eller mindre bevisst sammenligner det med
- Jo mer vi vet om det vi estimerer, jo mindre ligner det på det vi mer eller mindre bevisst sammenligner det med

[**simula** . research laboratory]

Hva så ...

- Assimilerings-effekten tilsier at vi bør være forsiktige med å sammenligne med oppgaver som er svært ulike de vi skal gjennomføre.
 - Tilsvarer at vi bør bruke en middels stor (eller flere) "baseline" i story point-basert estimering.
- Asymmetrien i sammenligner (Tversky's forklaring) tilsier at vi vil få den største assimilerings-effekten når vi bruker en stor oppgave som referanse og vurder hvor stor en liten oppgave er i forhold til denne.
 - Tilsvarer at bruk av en stor baseline (referanse) vil gi systematisk over-estimering av små oppgaver. På den annen side har vi ofte en sterk over-optimisme (særlig tidlig i prosjektet), slik at det kan argumenteres at en stor baseline-oppgave vil kunne "rette opp en feil med en annen".
- Følelse av at noe ligner kan være svært villedende når du vet lite om det du skal estimere og mye om det du sammenligner med.
 - "Du er likere meg, enn jeg er deg", skyldes at vi vet mye mer om oss selv enn om andre. Vi neglisjerer egenskaper vi selv har som ikke de andre har når vi selv er referansen.

[**simula** . research laboratory]

Relativ estimering

- **Sammenligning gjør ting stort sett likere enn de er**
 - Bruk middels store oppgaver som "baseline" i Story Point-basert estimering
 - Bruk prosjekter av samme størrelse som referanse i tidligfase-estimering
- **Lite kunnskap om det du skal estimere gjør at det ligner mer på alt annet (referansene)**
 - Neglisjering av egenskaper som finnes hos referansen, men ikke hos det du skal estimere

[**simula** . research laboratory]

Hvilke godt dokumenterte forbedringsmuligheter har vi?

- Bruk av historiske data (“looking back”)
 - F eks: Analogi-basert estimering og usikkerhetsvurderinger basert på tidligere fordeling av estimeringsfeil (**NB**: Generelle estimeringsmodeller fungerer dårlig)
 - Ansvarlig for innsamling og bruk av historiske data: Cost engineer? “Special Estimation Force”?
- Kombinering av UAVHENGIGE estimeringsmetoder/modeller og estimeringseksperter
- Finne relevant ekspertise (som er SMAL!)
- Gruppe-estimeringsprosesser (gitt at den ikke er dominert av en person)
- Unngå villedende og irrelevant informasjon, f eks ved å fjerne nøytralisere irrelevant og villedende informasjon før estimering.
- Unngå “Winner’s curse” (lettere sagt enn gjort, men noe kan gjøres)
- Identifikasjon av “skumle prosjekter” og gi disse særlig oppmerksomhet – som f eks sjekklister, maler, enkle estimeringsmodeller, ...

[**simula** : research laboratory]