

Myter og empiri

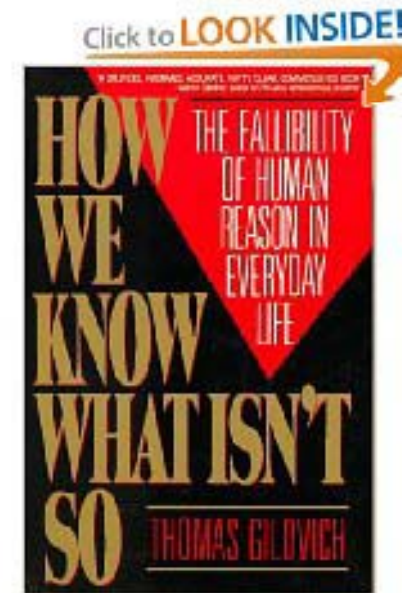
(og hvorfor vi aldri vil få svaret på om smidige metoder virker)

Magne Jørgensen
Simula Research Laboratory



Hvorfor og hvordan oppstår myter?

- Dekker et behov
 - Særlig for salg av tjenester og markedsføring av ideer
 - Vi liker å tro at det er sant
- Manglende presisjon i formidling/misforståelser
 - Kan godt være basert på gode studier
- Selvoppyllende
 - Ved å tro på myten så blir den (nesten) sann
- Det "passer godt" med vår opplevelse (lett å forestille seg)
 - Vår opplevelse er sjelden basert på representativitet
- Aktiv mytedannelse/feilinformasjon
 - Den offisielle "historien om" finnes ikke, den lages ...
- Gjentas mange ganger
- Autoriteter fremsetter dem
- Å forstå er å akseptere. De-akseptere er vanskelig (se neste side)



... dekker et behov

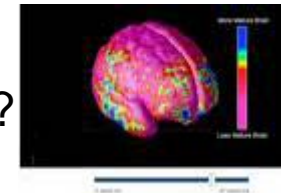
- **MYTE 1: Bindersen er en norsk oppfinnelse.**

- Binderser fantes allerede (også patenterte) da Vaaler patenterte den (se: no.wikipedia.org/wiki/Binders).
- Vaalers binders fungerer dårlig (test selv!)
- Spredning i USA styrket av forhold rundt nordmenn og 2. verdenskrig. (Symbol for samhold mot tyskerne).
- Brukes for å illustrere norsk oppfinnsomhet (Posten har f eks laget frimerke med binders, og BI, Sandvika hadde en bindersstatue).



- **MYTE 2: Vi bruker bare 10% av hjernen vår**

- Hva betyr utsagnet? 10% brukes av gangen? 90% som aldri blir brukt?
- Faktum er at vi bruker så å si 100% (om enn ikke på en gang)
- Opphavet til myten er noe uklart.
- Brukes til å selge "bruk-hjernen-bedre kurs" og til å forklare overnaturlig evner.



- **MYTE 3: Frosken som ikke hopper ut av gradvis varmere vann ("the boiling frog syndrome")**

- Opphav: Synes å komme fra et eksperiment utført i 1882.
- Ikke støttet av andre funn (og biologi-forskere synes å avvise fenomenet)
- Brukes til å selge miljøbudskap m.m.



... rene misforståelser eller feilinformasjon

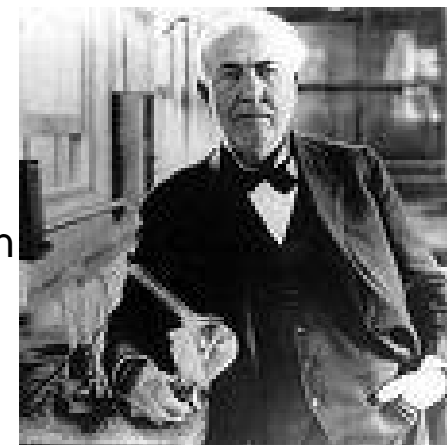
- **MYTE 4: Vi mister vi det meste av varmen gjennom hodet. (Bruk lue. Rumpa bar er ok?)**

- Hva betyr dette? At vi vil miste det meste av varmen gjennom hodet dersom vi står med kroppen godt påkledd, nakne, eller i en badstue med hodet ut vinduet en kald vinternatt?
- Det er antatt (British Medical Journal) at myten oppstod under en militærstudie hvor forskere iførte testpersoner arktiske overlevelsesdrakter - men ingen lue - og plasserte dem i ekstremt lave temperaturer. Ettersom hodet var den eneste kroppsdelen som ikke var tildekket, skjedde det meste av varmetapet der.
- Ekspertene hevder at dersom denne studien var blitt utført med badedrakter, ville varmetapet fra hodet vært kun 10 prosent.



- **MYTE 5: Tomas Edison oppfant glødepæra**

- Selv Edison-museet hevder visstnok ikke dette.
- Han – og teamet hans - var imidlertid viktig i forbedring og produksjon av glødepæra.



... ”passer” + selvoppfyllende

- **MYTE 6: Barn blir mer hyperaktive av sukker**

- Forskerne (British Medical Journal) har funnet til sammen 12 uavhengige studier gjort på sukker og barns oppførsel. Ingen av disse studiene kunne vise forskjeller i adferd hos de barna som fikk sukker og de barna som ikke fikk sukker. Dette inkluderer sukker fra godterier, sjokolade og naturlig sukker.
- Det er også forsket på hvordan foreldre reagerer på sukkermyten. Når foreldrene tror at barna får sukkerholdig drikke (selv om den egentlig er sukkerfri), vurderer de barna sine som mer hyperaktive. **Forskjellene i barnas oppførsel er trolig kun i foreldrenes hoder!**



- **MYTE 7: Nervøsitet, vikende blikk, og usammenhengende tale indikerer at en person lyver.**

- Masse forskning viser at det motsatte oftere er tilfelle. De som lyver er opptatt av å ikke oppføre ihht løgner-stereotypen.
- Men, det fleste av oss har historier der vi har fått bekreftet våre fordommer.
- Vi glemmer (eller bortforklarer) de gangene vi tok feil, og vektlegger der vi fikk rett.

... dekker et behov + ”passer”

- **MYTE 8: Gjennomsnittlig kostnadsoverskridelse i IT-prosjekter er på 189% (Standish Group: 1994-studie, mest referert “software crisis” dokumentasjon)**
 - Alle andre studier viser overskridelser på gj.sn. 30-40%.
 - Ikke-randomisert utvalg (overhyppighet i kriseprosjekter)
 - Skulle trolig vært 89% (men selv det er for høyt)
 - Brukt av mange for å argumentere for mer forskning, salg av konsulentbistand, etc.
- **Hvorfor trodde/tror vi på dette tallet?**
 - Studien finner at IT-prosjekter i gjennomsnitt kostet 2,9 ganger mer enn estimert!
 - Slike overskridelser (2.9 ganger mer) er UNNTAKET!
 - Hvorfor aksepterer vi UNNTAKET som typisk?
 - Husker unntak best?
 - Passer til en følelse at estimering er vanskelig?
 - Gjelder sikkert de andre (som er dårligere enn meg)?
 - Passer til et formål? Selge konsulentttjenester. Selv om jeg er dårlig til å estimere, så er jeg bedre enn gjennomsnittet.



Jim Johnson
Standish Group

... dekker et behov + ”passer”

- **MYTE 9: 45% av funksjonaliteten brukes aldri (The Standish Group, igjen)**

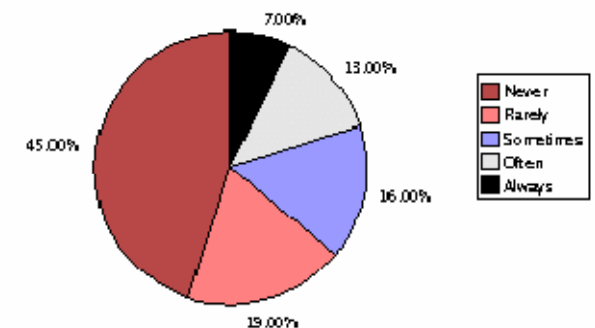
- Refereres svært mye i smidigmiljøet for å argumentere for å gå fra vannfalls til smidig utvikling.
- Originalstudie ikke tilgjengelig. Henvisning til en studie presentert av Jim Johnson på XP 2002.
- Ifht til www.agilemodeling.com/essays/examiningBRUF.htm var studien som følger:

- **The Standish Group looked at a subset of traditional teams which eventually delivered into production and asked the question, “Of the functionality which was delivered, how much of it was actually used?”**

- Hvordan ville du tolke spørsmålet? Gjelder spørsmålet funksjonalitet som ikke har vært bruk av noen, noen gang?
- Hvor god kunnskap tror du at de som svarte har om dette? Hva baserer de det på? Målinger? (Neppe)
- Virker det rimelig? Du vet trolig om minst et system av denne typen, men er det noe som **typisk** skjer?

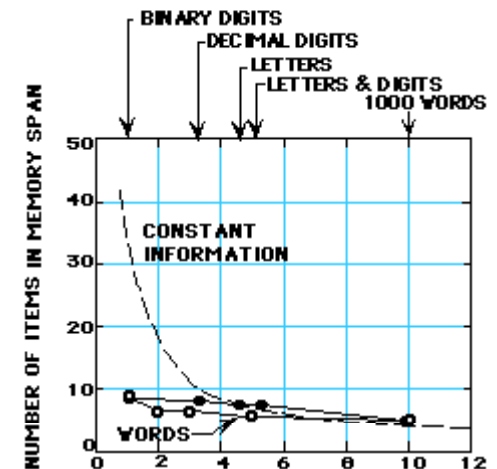


Features and Function Usage



... misforståelse

- **MYTE 10: Det optimale antall elementer (f eks i et brukergrensesnitt eller kall fra en rutine) er 7 pluss/minus 2, dvs mellom 5 og 9.**
 - *“One guideline that needs closer attention is the “Magic Seven, Plus or Minus Two” (7 ± 2) rule of thumb. This principle has often been applied to determine the number of items in a navigation menu on a web page. It arose to satisfy a tactical need to make quick design choices and to objectively justify navigation to site stakeholders.”*
 - Basert på studier av Miller om kognitiv kapasitet (dvs det maksimale vi klarer, f eks å oppfatte visuelt uten å dele opp eller huske, i gitte kontekster).
 - Studiene omhandler imidlertid **ikke** hva som er et optimalt antall og er neppe særlig relevant for antall menyvalg i utforming av brukergrensesnitt, kall fra programvarerutiner og mye av det andre den brukes til.
 - Studiens resultat klart misforstått. Sammen med behovet for enkle regler for utforming og at syv er litt magisk, har dette spredd seg.



... behov for enkle regler

- **MYTE 12: “Adding manpower to a late software project makes it later”** The Mythical Man-Month (Brook’s lov)
 - Beskrives av Brook selv som en: "outrageous oversimplification". Myten brukes blant annet for å argumentere for håpløse estimeringsmodeller.



- **MYTE 11: Forskjellen mellom dårligste og beste programmerer er 10:1.**
 - Studie fra 1968 på med erfarne utviklere på en konkret problemstilling.
 - Forholdet kan åpenbart variere mye, mye mer! (Minner om myten om at "vi mister det meste av varmen gjennom hodet")

... nyttig, høres riktig ut

- **MYTE 13:** “Lykkelige team er produktive team”
 - Veldig nyttig for å få arbeidsgiver til å bruke penger på arbeidsmiljø.
 - Svært lite entydig dokumentasjon av forholdet “happy”-”productive”, blant annet fordi det er vanskelig å måle. Stort sprik i resultatene avhengig av kontekst og begrepsbruk.
 - Trolig forskjell på å være fornøyd med oppgave, medarbeidere og ledelse. (Fornøyd med oppgave synes å være mest viktig.)
 - Studier indikerer at “happy teams” kan ha større spredning i produktivitet. De er både de mest og de minst produktive.
- **MYTE 14:** Brainstorming bør gjøres i grupper
 - Mange studier tyder på at vi er mer innovative alene (særlig når vi føler at vi blir ”observert” – Zajonc-studie)
 - MEN, informasjon bør deles og ideer bør diskuteres i grupper
- **MYTE 15:** Det meste av systemutvikling er vedlikehold (eller testing, eller ...)
 - Manglende presisjon i begreper

... enkle regler

- **MYTE 16:** Kostnadene ved endring (eventuelt retting av feil) stiger med en faktor 10 i hver fase
 - Påstand finnes i mange varianter. Den mildeste er at det koster alltid mer å rette en feil i produksjon.
- Alistair Coburn skriver for eksempel på www.xprogramming.com/xpmag/cost_of_change.htm
 - *The first thing is to establish that it really does cost x10 or similar to detect and handle mistakes as we proceed across r, p, t, f, (i.e. exponential) and that XP really is well suited to handle that exponential growth in cost. It is not the case that the x10 doesn't exist, or that XP nullifies it. XP deals with its existence, is "well aware of it" to use my preferred anthropomorphizing locution.*
 - *r=requirements; p=programming, t=test, f=field (10*10*10=1000 ganger høyere i produksjon enn i krav!)*
- Hva betyr påstanden om x10?
- Hvor kommer x10 fra?

Hva vi bør gjøre ...

1. Finn ut hva som menes med påstanden
 - a) Har den noe presist innhold?
 - b) Lar den seg falsifisere? (Hvis ikke, hva slags funksjon har de da?)
2. Vær kritisk mhp din egen tendens til å akseptere påstander uten evidens eller med ikke-nøytral evidens
 - a) Se grunner for mytedannelse. Hvorfor høres dette riktig ut?
 - b) Det å forstå innebærer (mener noen) å akseptere – så kan man eventuelt "de-akseptere" etterpå
 - c) Hvem kommer med påstanden? Aksepterer du den pga kilden, ikke pga innholdet?
 - d) Du glemmer fortære kilden enn innholdet (Påstander fra kilder med lav kvalitet kan dermed bli sterkere med tiden! Dette kalles "the sleeper effect").
3. Hva finnes av nøytral kunnskap (evidens)?
 - a) Finnes nøytrale og relevante studier eller erfaringer som støtter dette? (Erfaringer av en som selger kurs i samme metode er neppe nøytrale)
 - b) Hva sier totaliteten av evidens?
 - c) Når gjelder påstanden?

Manifesto for Agile Software Development

We are uncovering better ways of developing software by doing it and helping others do it. Through this work we have come to value:

Individuals and interactions over processes and tools

Working software over comprehensive documentation

Customer collaboration over contract negotiation

Responding to change over following a plan

That is, while there is value in the items on the right, we value the items on the left more.

Kontrollspørsmål:

- a) Hva betyr utsagnene? Har de noe presist innhold? Er det mulig å falsifisere utsagnene?
- b) Hvorfor har manifestet slått an?
- c) Dette er "verdier". Gjør det noen forskjell på om det hadde vært påstander?

The Principles

We follow these principles:

- *Our highest priority is to satisfy the customer through early and continuous delivery of valuable software.*
- *Welcome changing requirements, even late in development. Agile processes harness change for the customer's competitive advantage.*
- *Deliver working software frequently, from a couple of weeks to a couple of months, with a preference to the shorter timescale.*
- *Business people and developers must work together daily throughout the project.*

Kontrollspørsmål:

- Hva betyr prinsippene? Har de noe presist innhold? Er det mulig å falsifisere noe her?
- Hvorfor har de slått an?
- Hva er objektiv kunnskap om prinsippene?

Hva vet vi om smidige metoder? (og hvorfor vet vi ikke mer?)

- Dybå et al. (publisert i Information and Software Technology)
 - *“The results of the review suggest that it is possible to achieve improved job satisfaction, productivity, and increased customer satisfaction. The strongest, and probably most relevant, evidence for practice is from the studies of mature agile teams, which suggests that it is necessary to focus on human and social factors in order to succeed. Specifically, it seems that a high level of individual autonomy must be balanced with a high level of team autonomy and corporate responsibility. It also seems important to staff agile teams with people that have faith in their own abilities combined with good interpersonal skills and trust.”*
 - *“The evidence also suggests that agile methods not necessarily are the best choice for large projects. Thus, consistent with recommendations provided by others [11, 12, 15], we suggest that practitioners carefully study their projects’ characteristics and compare them with the relevant agile methods’ required characteristics.”*
 - *“Due to the limited number and relatively poor quality of the primary studies in this review, it is impossible to offer more definitive and detailed advice.”*

Hva vet vi om smidige metoder? (og hvorfor vet vi ikke mer?)

- Spørsmålet som stilles er ofte: **Lønner smidige metoder seg?**
- Er spørsmålet riktig stilt?
 - JA: Det er det vi ønsker å vite noe om.
 - NEI: Spørsmålet er for upresist til å la seg besvare meningsfullt!
- Hva mener vi med smidige metoder?
 - Hele XP, Lean, ...?
 - Er vi smidige selv om vi kun bruker deler av en smidig metode?
 - Forutsetter det riktig bruk av XP, Scrum, ...?
 - Forutsetter det egnede prosjekter?
- Hva mener vi med "lønner seg"?
 - Utviklingskostnader?
 - Totalkostnader kunde?

Hva kan vi vite noe om?

- Vi trenger et noenlunde presist og avgjørbare spørsmål for å kunne gi noenlunde presise svar.
- Spørsmålet bør (særlig i sammenhenger med mange variable) være "Når lønner det seg med X?" og ikke "Lønner det seg med X?"
- I størst mulig grad bør svarene komme fra nøytrale parter som etterstreber representativitet og uavhengighet.
- For mange problemstillinger vedkommende må vi gå fra å spørre om kunnskap om "totalpakka" (f eks smidige metoder) til kunnskap om "reduerte" problemstillinger som bygger på forskningsbasert kunnskap om kognisjon, arbeid i grupper, problemløsning m.m..
 - Eksempel: Hvilke typer feil i programvaren er det mer sannsynlig å oppdage ved at man bruker to programmere som jobber i parallell, hvorav den ene kun inspiserer det den andre gjør, gitt erfaringsnivå XX og applikasjoner av typen YY? I hvilken grad avhenger dette av hvor lenge disse har jobbet sammen.
 - Spørsmålet bygger på forskning om gruppedynamikk, psykologi, m.m.
- Reduserte problemstillinger er ikke optimale (synergieffekter), men trolig det lengste vi kommer. Krav om enkle svar på svært komplekse, dårlig definerte problemer gir liten framgang i kunnskap.