



KKEG: Etik I

**VAD SKULLE DU GÖRA OM
DU VAR OSYNLIG?**

KP frågade några elever i skolan i Abisko.



Sonja Fjärdbo, 10 år

Jag skulle gå och ta saker på bus. Sen skulle jag kidnappa folk och begrava ner dem i jorden. Sen skulle jag knycka en massa pengar. Jag har alltid önskat att jag kunde vara osynlig.

Idag



- Filosofiska distinktioner
- Moraliskt dilemma
- Kognitionsvetenskapligt perspektiv på moral
 - **Hur bildas moraliska val/ omdömen?**
- (Psykologi om moral)

Nästa tillfälle (Om en månad)



- Diskussioner om moraliska dilemman
- Samarbete
- Utveckling av empati och etik

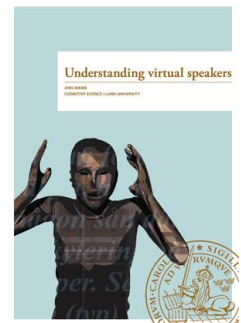
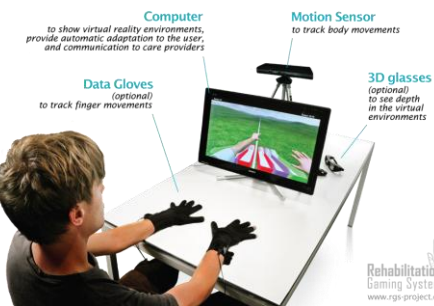
JENS NIRME

MSc Cog Sci + HCI, Universitat Pompeu Fabra
Stroke motor rehab. w low-immersive VR

PhD in Cognitive Science
- Educational Technology Group
- Multimodal interaction and communication

Post-doctoral researcher
- VR lab
- Humanities Lab

Associate lecturer in Interaction Design, LTH



Acknowledgement ...



Philip Pärnamets (philip.parnamets@lucs.lu.se)



Distinktioner I



Etik

- Vad är **rätt** eller **fel**? (**teori**, allmängiltigt)

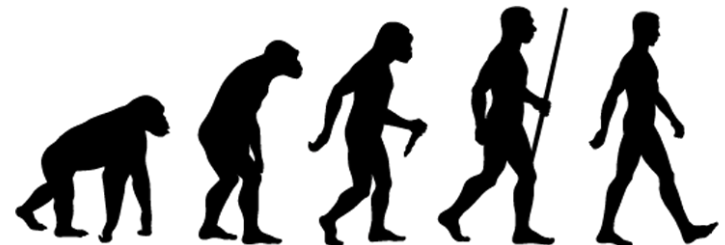
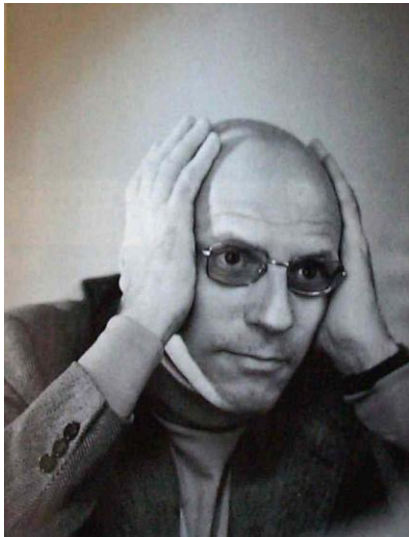
Moral

- Hur man faktiskt tar **beslut** och **beter** sig.
(I **praktik**, personligt)

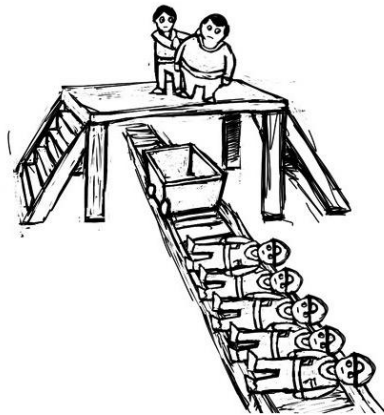
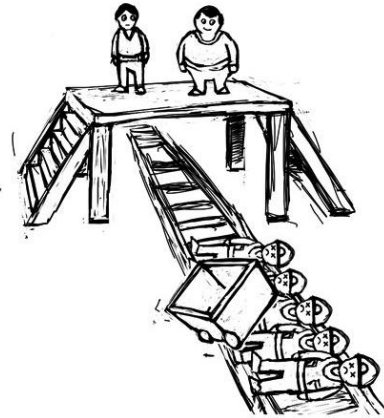
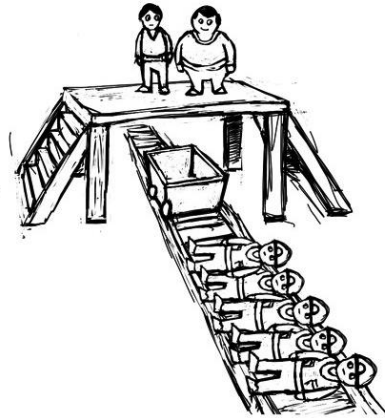
Perspektiv på moral



Man kan tänka på och studera moral och etik på olika sätt ...



Trolleyproblem I



Trolleyproblem I



De två svarsalternativen på "the trolley problem" motsvarar **två olika klassiska filosofiska teorier** om hur man kommer fram till vad som är moraliskt

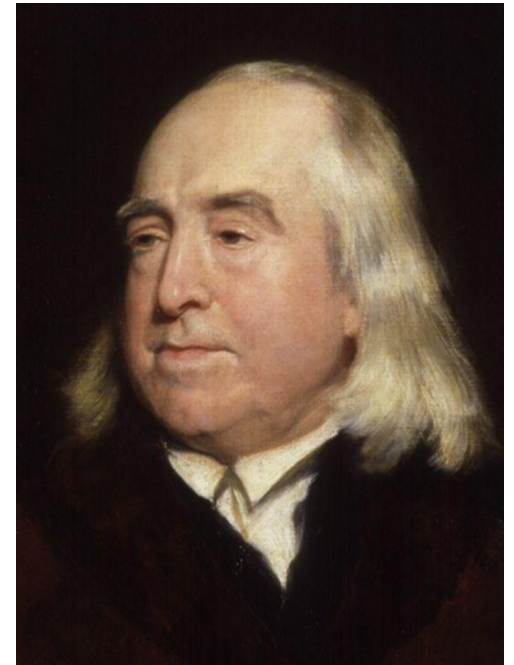
Distinktioner II



- **Plikt/regel-etik**
- **Utilitarism/konsekvensetik**

Utilitarism

- Bentham (1748-1832)
 - Största möjliga lycka åt störst antal individer
 - Njutning – motsats till lidande
 - “hedonic calculus”



Regeletik

- Immanuel Kant (1724-1804)
 - Vissa handlingar har inneboende moraliskt värde, oavsett syfte eller utfall!
 - Pliktetik vs dygdetik
 - Människor har inneboende lag om vad som är rätt och fel
 - ”Kategoriskt imperativ”:
”Handla endast efter den maxim genom vilken du tillika kan vilja att den blir en allmän lag”.

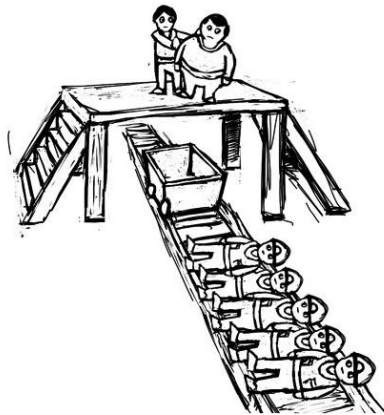
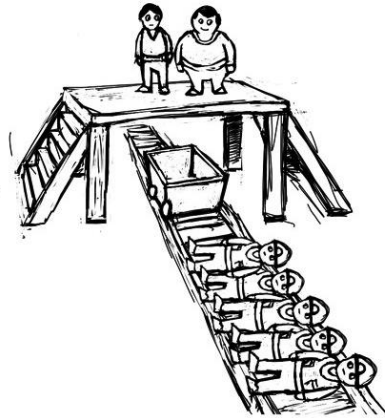


Distinktioner II



- **Regeletik**
 - **Pliktetik**
 - Dygdetik
- **Konsekvensetik**
 - **Utilitarism**
 - Egoism
 - Hedonism

Trolleyproblem I





Diskussion

- Exempel regel- vs konsekvensetiska ställningstagande ni själva tar?
 - Diskutera 5 min i par!

Moraliskt dilemma



Jan flickvän **Lisa** är skeppsbruten på en öde ö, utan mat eller drickbart vatten.

Lisa ringer **Jan** och undrar om han kan komma och rädda henne. Med **Jan** är också hans nära kompis **Jenny**.

Jans mamma heter **Ulla**. **Ulla** har väninnan **Siv**, som har en båt. **Jan** frågar **Siv** om han får låna hennes båt för att åka och rädda **Lisa**.

Siv svarar: *Ja om du ligger med mig.*

Jan går till sin mamma **Ulla** och frågar: *Hur ska jag göra?*

Ulla svarar: *Det får du ta ställning till själv.*

Jan ligger med **Siv**. Han lånar båten och åker tillsammans med **Jenny** för att rädda **Lisa**.

Jan berättar för **Lisa** vad som har hänt. **Lisa** gör slut omedelbart. **Jenny** blir så arg att hon smäller till **Lisa**, som faller så illa att hon avlider.

Jan

Lisa

Jenny

Ulla

Siv

Moralkognition



Kognitionsvetenskap är typiskt intresserad av **hur** ett moraliskt omdöme eller handling uppstår

Genom att studera vad som **påverkar** ett moraliskt omdöme

Förnuftig moral



Tidigt lades tyngdpunkten på **tänkande och argumentation** inom kognitionsvetenskap

Man tänkte sig att människan fungerade likt en **dator**

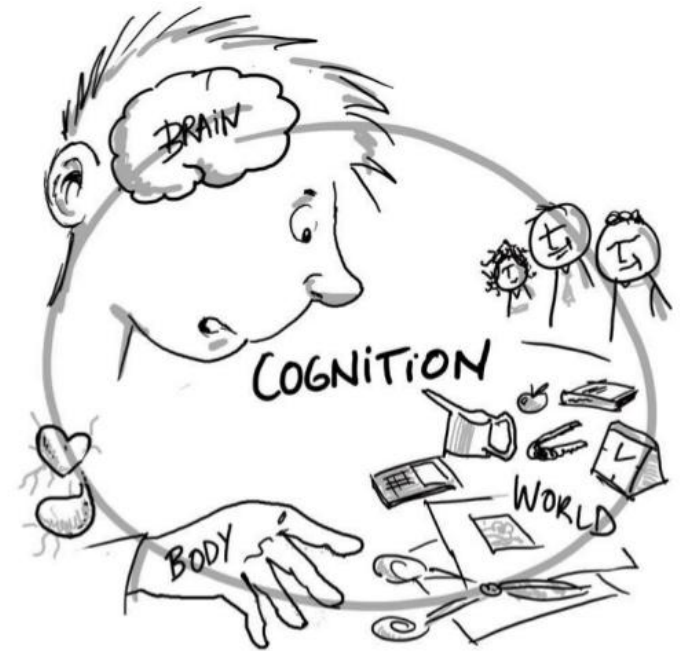
Barns utveckling som en process mot ständigt **mer avancerade resonemang**

Människor är inbäddade i världen



Vi är omgivna av våra kroppar, andra människor, miljöer och resultaten av våra egna handlingar.

Denna **rika inbäddning** har **format utvecklingen** av våra kognitiva system.



Testbara hypoteser



Om omdömen är
förnuftsbaserade så bör de:

- **kunna förklaras och motiveras**

Och inte påverkas av:

- **känslor**

Eller andra

- **irrelevanta faktorer**



Förnuftig moral, hypoteser

Om omdömen är
förnuftsbaseade så bör de:

- kunna förklaras och motiveras

Och inte påverkas av:

- **känslor**

Eller andra

- **irrelevanta faktorer**

En moralisk fråga ...



Skulle det var rätt för ett **syskonpar** att ligga med varandra?

De flesta svarar **nej**

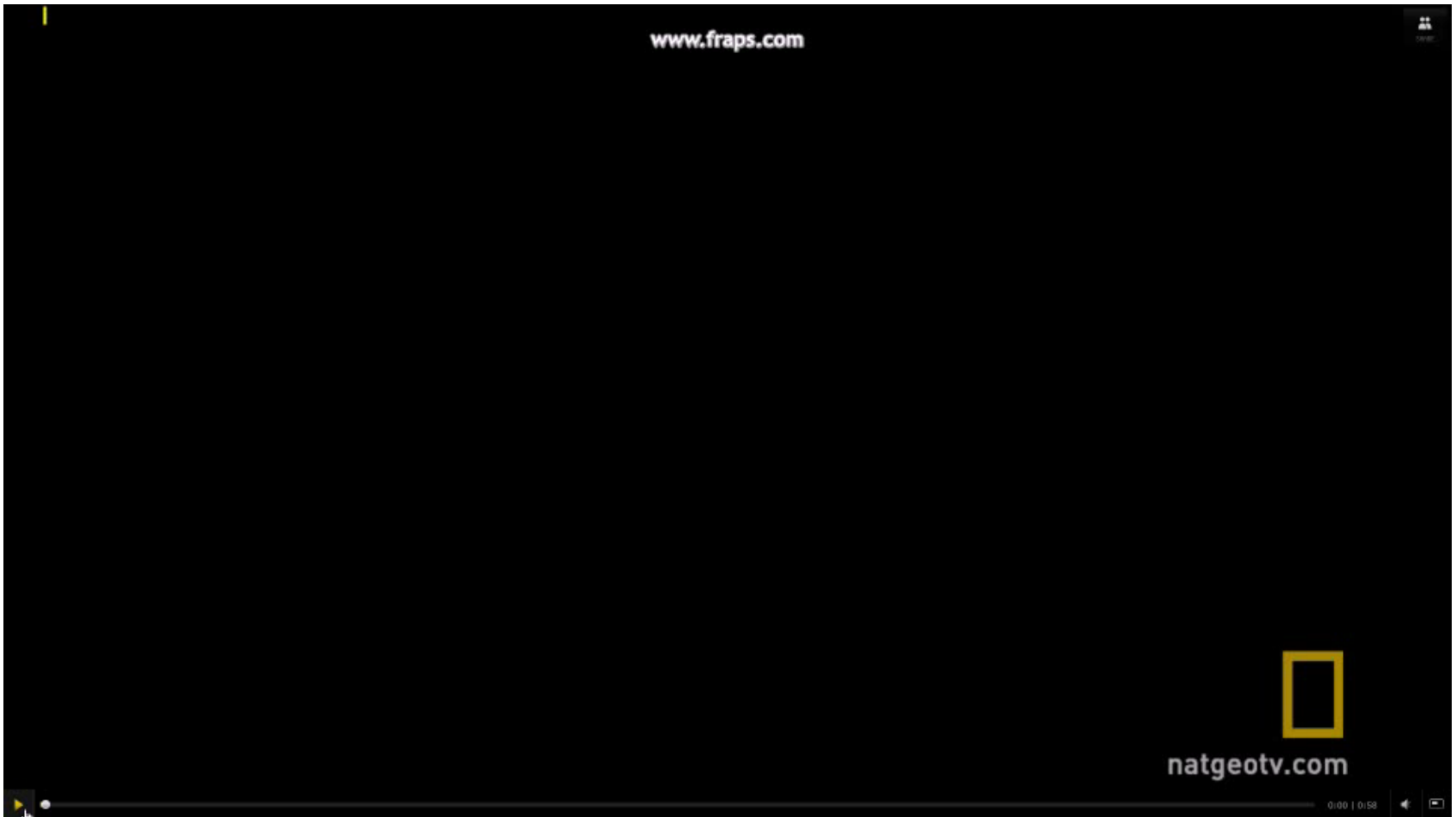
'moral dumbfounding'



Vad händer om man börjar ge **argument** för att det kan vara ganska **säkert** för syskon att ligga med varandra?

De flesta svarar fortfarande **nej**, men får allt **svårare att motivera** sina omdömen

Choice blindness



Motivera beslut som inte tagits

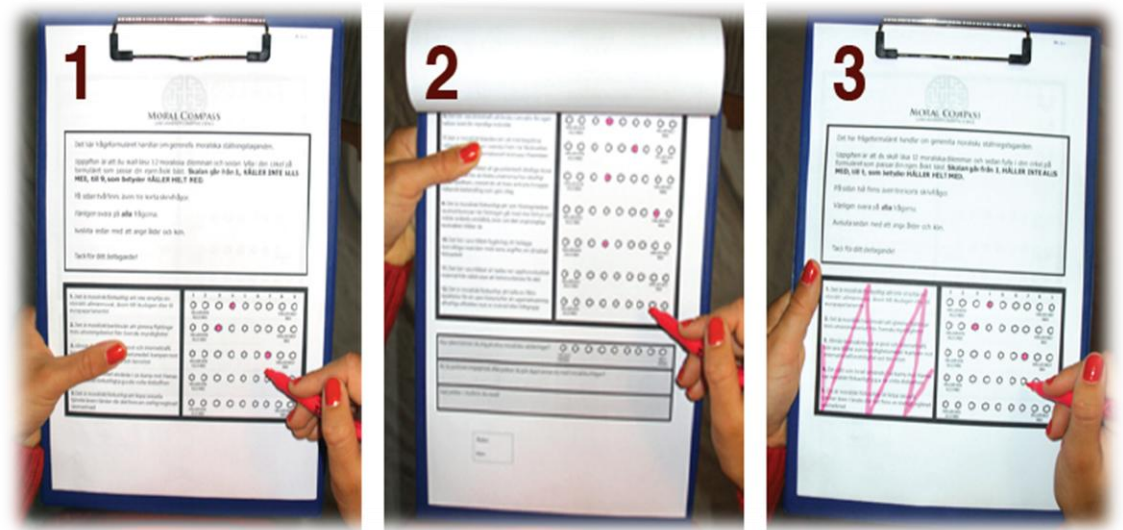
- **Retroaktivt.** Minnen återskapas, snarare än återkallas.



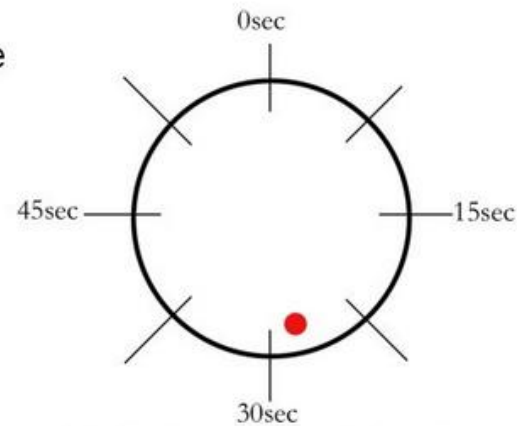
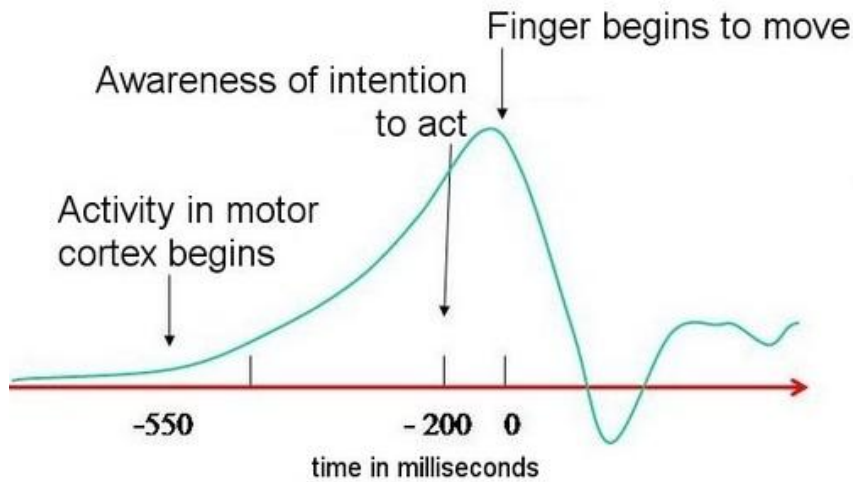
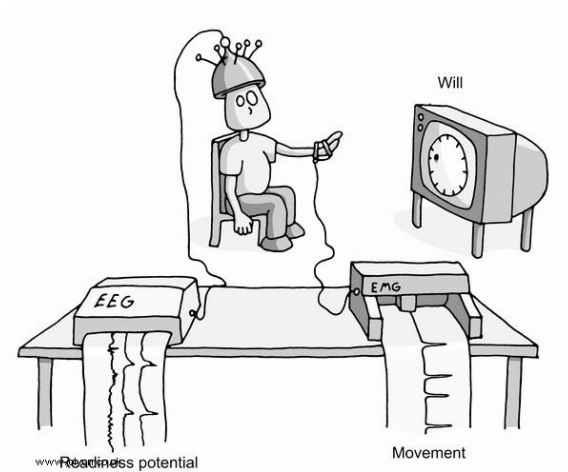
Beslutsblind moral



- **Konkreta och abstrakta** moraliska frågor
- Bara **hälften** märkte att deras svar bytts ut
- **Gav motiveringar** för bytta åsikter som var lika 'äka' motiveringar



Fri vilja, en illusion?



Libet's electronic clock. The subject notes where the red dot is when his/her intention to act arises.

Libet, konsekvenser och kritik



Fri vilja ett “epifenomen” ?

Libet själv ansåg inte att hans
resultat förnekar fri vilja

Medveten “vetorätt” ?

Problematiskt med självbedömning

Förnuftig moral, hypoteser



Om omdömen är
förnuftsbaseade så bör de:

- ~~kunna förklaras och
motiveras~~

Och inte påverkas av:

- **känslor**

Eller andra

- **irrelevanta faktorer**

Så vad är **RELEVANTA** faktorer?



Ett beslut kan **ibland vara svårt** och **ibland vara enkelt**, men vi har en föreställning om att utfallet av våra beslut **reflekterar** vad vi faktiskt tycker:

våra ”**preferenser**”

Homo economicus



Homo economicus är tanken att vi fattar beslut med hjälp av **rationella resonemang** på en grund av **skarpa och stabila preferenser** som vi har en **oproblematiskt tillgång** till

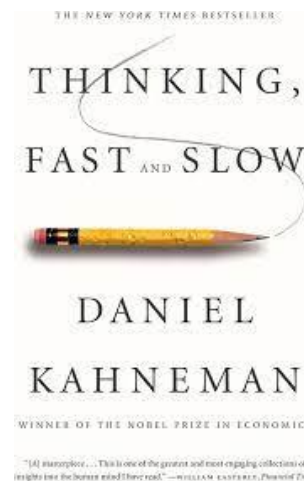
Bentham (1789): “hedonic calculus”

~~Homo economicus~~



Kahneman: Kognitiva “biaser”

- Avvikelser från objektivitet i beslut eller omdömen





Formuleringseffekter

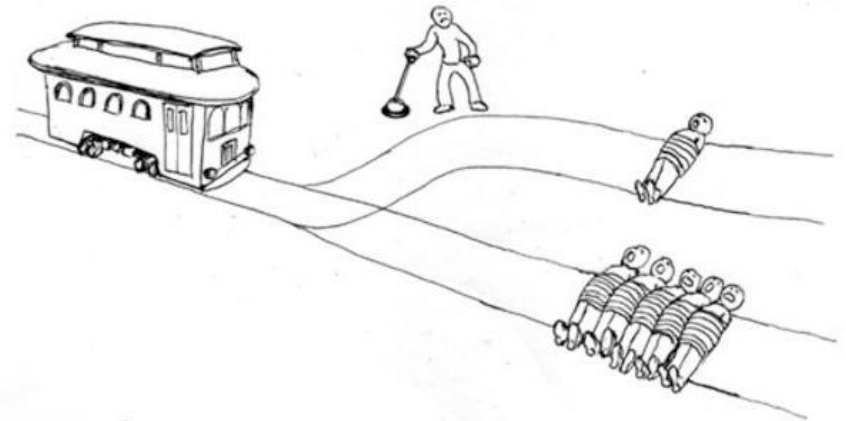
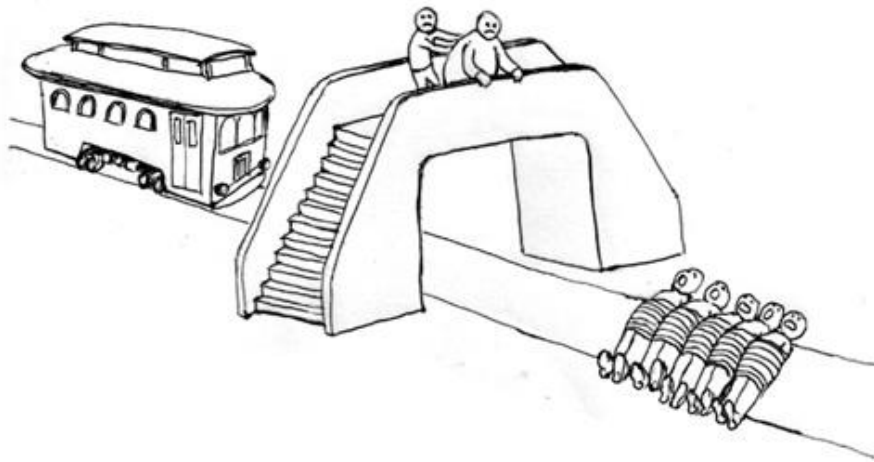
”**offra** en för att **rädda** fem personer”

”**döda** en för att fem personer ska **överleva**”

Olika formuleringar ger skillnader i vad som anses acceptabelt

- Borde vara irrelevant om utfallet är samma!

Formuleringseffekter



Vem skulle du låta adoptera ett barn?



Parent A: average income
average health
average working hours
reasonable rapport with the child
relatively stable social life

Parent B: above-average income
very close relationship with the child
extremely active social life
lots of work-related travel
minor health problems

Vem skulle du INTE låta adoptera ett barn?



Parent A: average income
average health
average working hours
reasonable rapport with the child
relatively stable social life

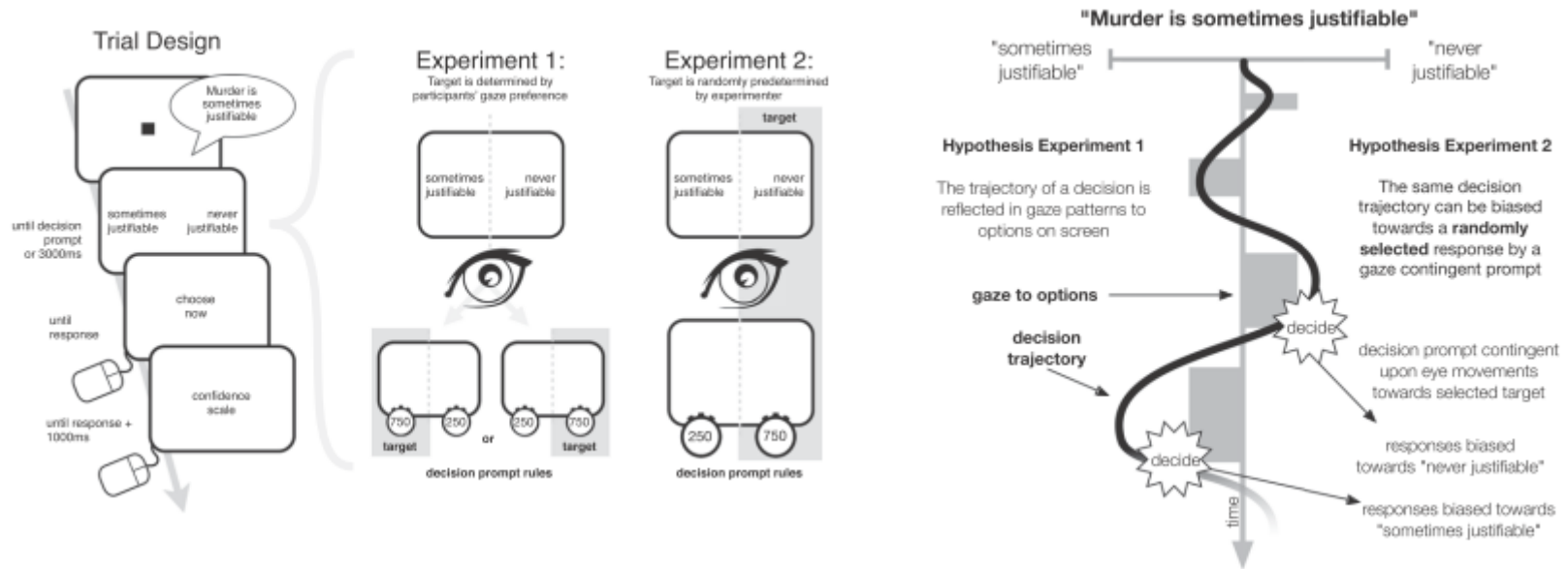
Parent B: above-average income
very close relationship with the child
extremely active social life
lots of work-related travel
minor health problems

lättillgängliga anledningar



	<i>Award</i>	<i>Deny</i>
<i>Parent A:</i> average income average health average working hours reasonable rapport with the child relatively stable social life	36%	45%
<i>Parent B:</i> above-average income very close relationship with the child extremely active social life lots of work-related travel minor health problems	64%	55%

Biasing moral decisions by exploiting the dynamics of eye gaze (P. Pärnamets, 2015)



Experiment 2: The target is randomly determined before trial onset. Participants' gaze is measured while they view alternatives and once the target alternative has been viewed for at least 750 ms and the other alternative has been viewed for at least 250 ms the decision prompt is activated and the trial is interrupted.

Biaser

- Biaser kan vara gynnsamma eller nödvändiga
 - Genvägar till beslut
 - Bidra till samarbete (nästa föreläsning)



Förnuftig moral, hypoteser



Om omdömen är
förnuftsbaseade så bör de:

- ~~kunna förklaras och motiveras~~

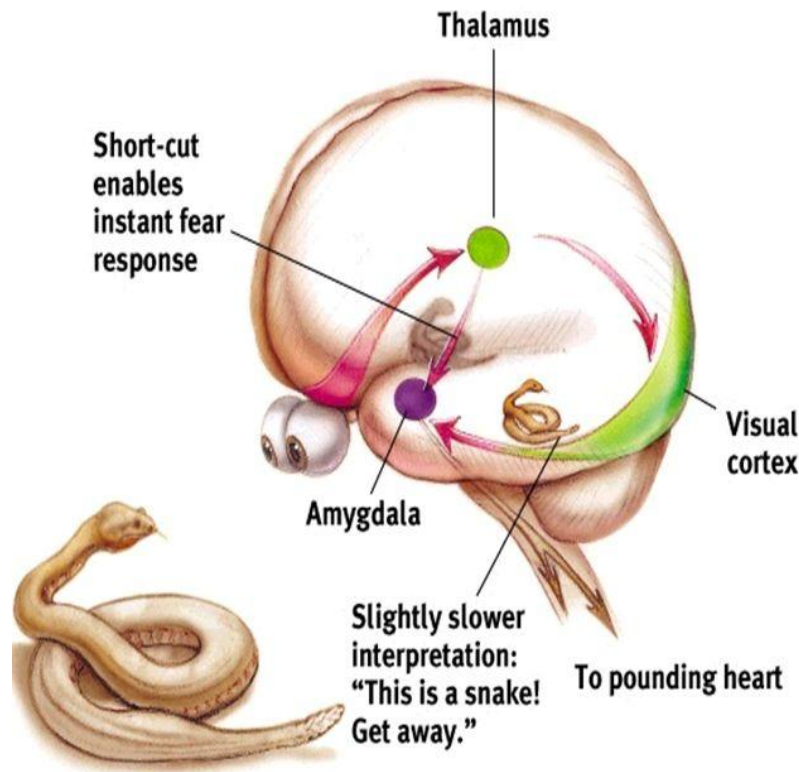
Och inte påverkas av:

- känslor

Eller andra

- ~~irrelevanta faktorer~~

Rädsla oberoende av kognitiv kontroll



Auditory and visual stimuli: sights and sounds are processed first by the thalamus, which filter the incoming cues and shunts them either to the amygdala or the cortex.

Olfactory and tactile stimuli: smells and touch sensations bypass the thalamus, taking a shortcut to the amygdala. Smells, often evoke stronger memories or feelings than do sights or sounds.

Amygdala: the emotional core of the brain, primary roles of triggering the fear response. Information that passes through the amygdala is tagged with emotional significance.

Hippocampus: vital to storing the raw information coming in from the senses, along with the emotional baggage attached to the data during their trip through the amygdala.



hjärtrytmfeedback

personer fick **lugn eller snabb**
hjärtrytmfeedback (60/90 slag per
minut) och läsa om en organisation
som jobbar med homosexuellas
rättigheter

39.5% av de med snabb feedback
men bara **17.1%** av de med lugn
feedback kunde tänka sig arbeta
som frivillig

Äckel och hypnos



Personer hypnotiserades att känna **äckel** när de läste vissa specifika (neutrala) ord

hypnotiserade och ohypnotiserade deltagare fick bedöma olika moraliska scenarion

Äckel och hypnos



Hypnotiserade deltagare fällde
mycket hårdare omdömen om
alla moraliska scenarior
(men inte andra scenarior)



Wheatley & Haidt (2005)

Förnuftig moral



Om omdömen är
förnuftsbaseade så bör de:

- ~~kunna förklaras~~

samt inte påverkas av:

- ~~irrelevanta faktorer~~

- ~~känslotillstånd~~

~~Förnuftig moral~~ ????



Om omdömen är
förnuftsbaseade så bör de:

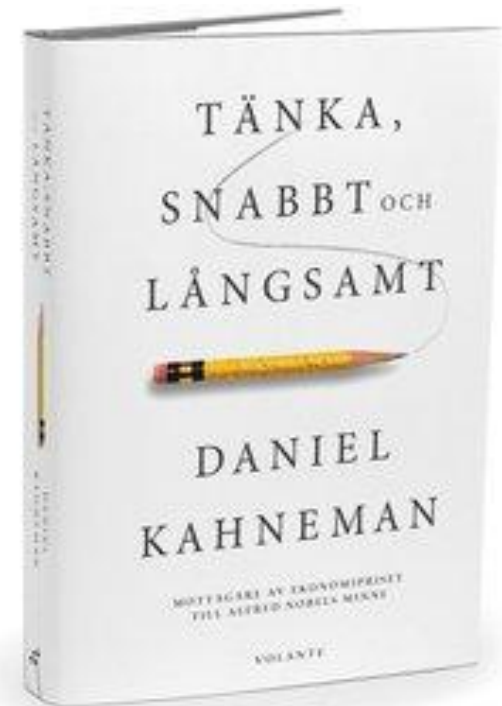
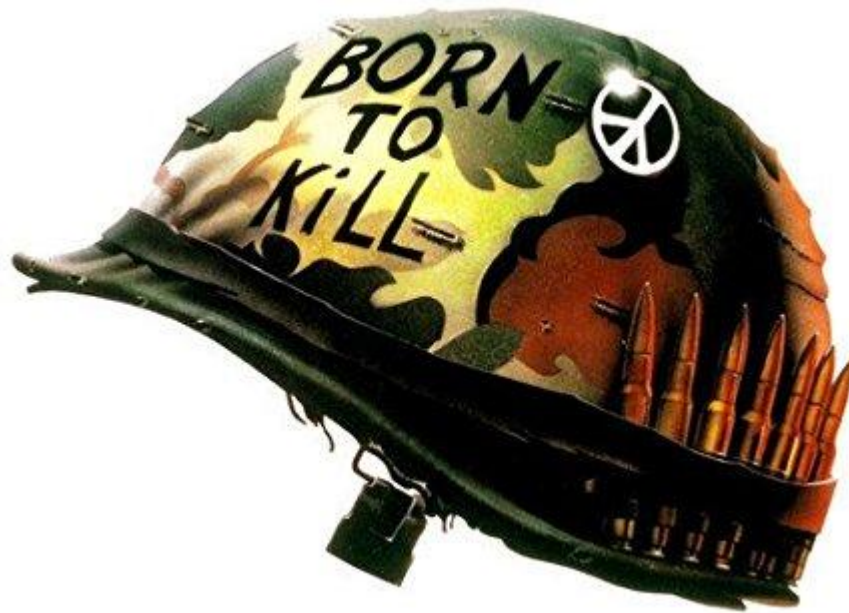
- ~~kunna förklaras~~

samt inte påverkas av:

- ~~irrelevanta faktorer~~

- ~~känslotillstånd~~

Moral verkar inte vara
endast förnuftsbaserat

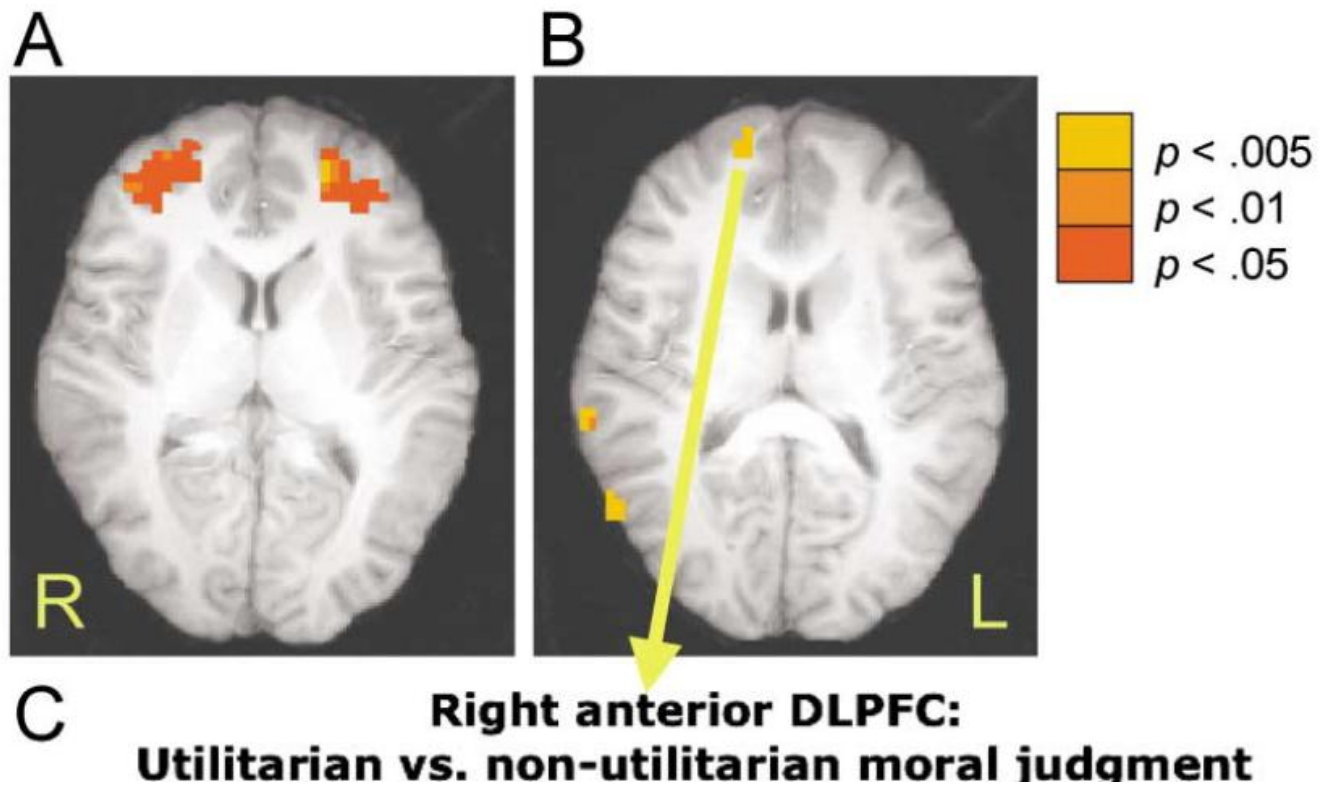


Snabba känslor, långsamt förnuft



Fler svarar regel/känslobaserat om
de får lite tid på sig (**8s**) än om de
får lång betänketid (**3min**)

Kognitiv kontroll



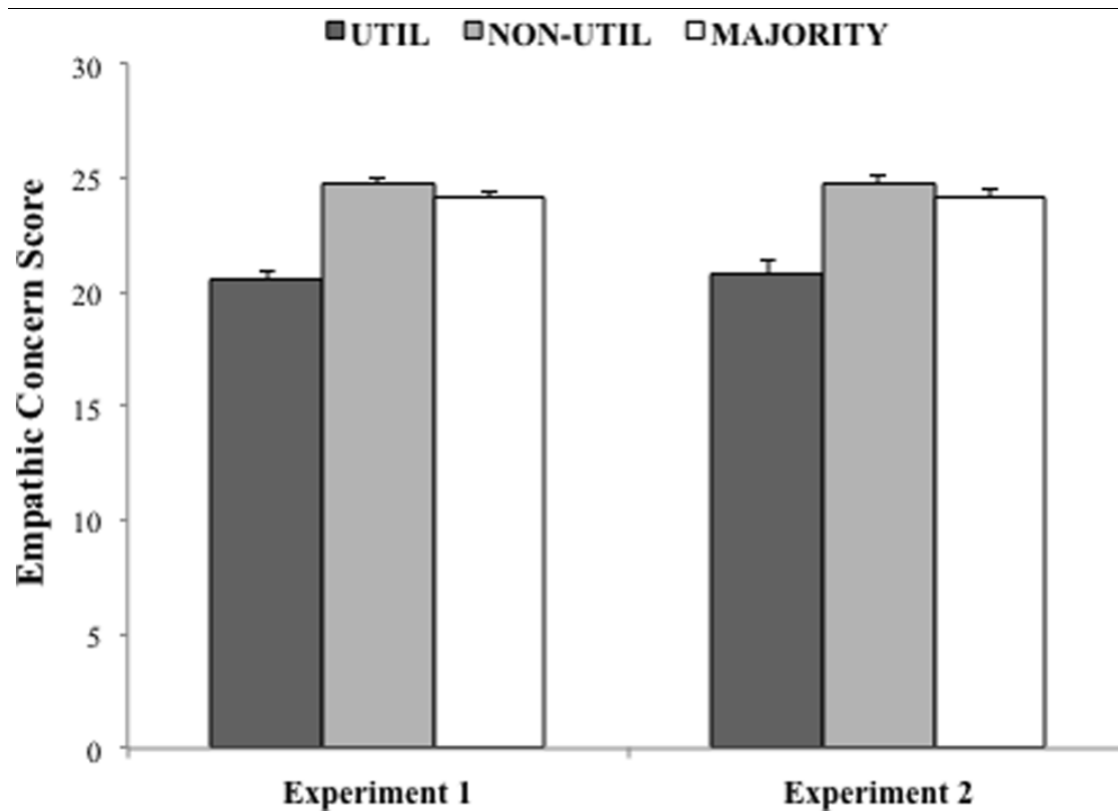
Personlighetstyper och omdömen



Deltagare får svara på två trolley dilemman
(bro och spårbyte)

Fylla i personlighetsformulär
empathic concern score

Empati och utilitaristiska omdömen



Sammanfattningsvis



Moraliska omdömen är **känsliga för påverkan** i situationen

Både "**känsla**" och **förnuft** är inblandade även om "**biaser**" **ofta styr**

- (Biaser kan vara gynnsamma)

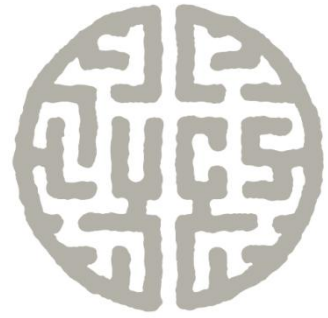
Moraliska omdömen baseras **inte endast** på etiska **resonemang**

Validitet

- Hur väl stämmer resultat från kontrollerade och hypotetiska situationer med *verkligheten*?



**Vad spelar detta för
roll för er?**



???



Vad spelar detta för roll för er?

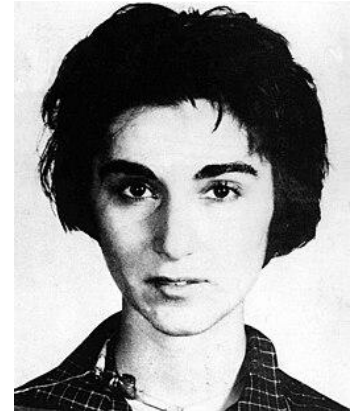
Förutsättningar kan ha stor betydelse

Argument mot att lägga allt moraliskt ansvar på användare
... eller på designer!

Nedslag i psykologi

Bystander effect

- Kitty Genovese mördad i New York 1964
- 37 grannar bevittnade utan att ingripa enl. media
 - (Visat sig vara fel i efterhand)
- The greater the number of bystanders, the less likely it is for any one of them to provide help to a person in distress. Darley, J. M. & Latané, B. (1968).



Darley, J. M. & Latané, B. (1968). "Bystander intervention in emergencies: Diffusion of responsibility". *Journal of Personality and Social Psychology*. 8 (4, Pt.1): 377–383. doi:10.1037/h0025589.

Kontrollfokus

Internal vs. External Locus of Control

Internal

Believe they are in control and take responsibility for their own actions



External

Blame external forces for their own circumstances



Etiska problem

Hypotetiska fallstudier

Dilemma: självkörande bilar



Hi Seth, I'm the driver of this vehicle. No I'm not sleeping, I'm very relaxed. Driving on Autopilot can be very relaxing. At the last second of this video you can see my hand is on the bottom of the wheel and my head turns to look at a mirror. Please mind your own business.

Dude is straight snoozing going 75mph on the interstate, letting his @Tesla do the work. 🤪🤪



126 5,297 17.6K

Dilemma: självkörande bilar

Level 3 fatalities [\[edit \]](#)

A Level 3 autonomous driving system would occasionally expect a driver to take over control.

List of known autonomous car fatalities (occurring while autonomous-system acknowledged to have been engaged)

Date	Incident no.	Country	City	State/county/province	No. of fatalities	System manufacturer	Vehicle Type	Distance driven by the system at time of incident	Notes
18 March 2018	3	United States of America (USA)	Tempe	Arizona	1	Uber	'Refitted Volvo' ^[10]	—	Pedestrian fatality. ^[11]

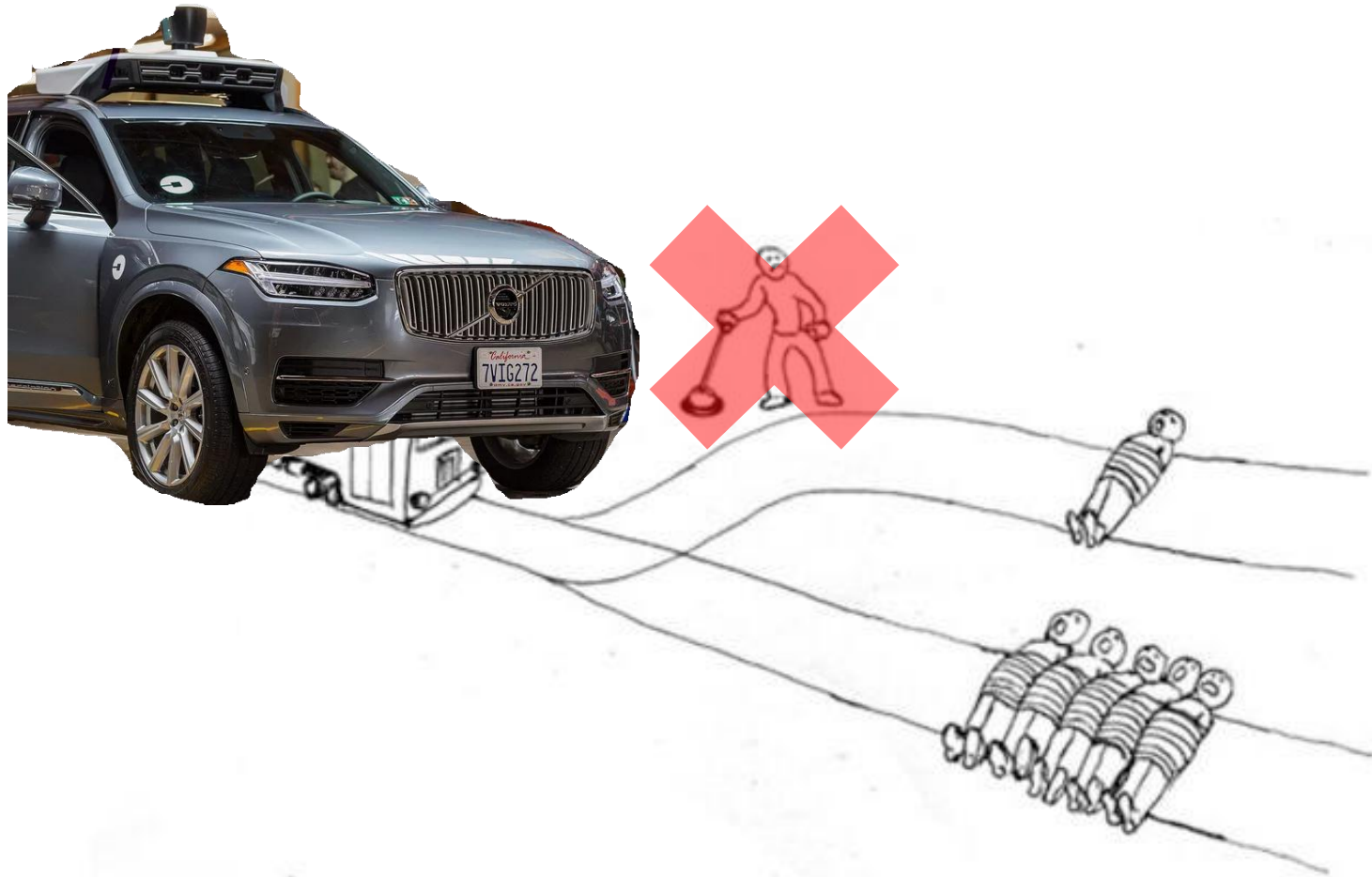
Level 2 fatalities [\[edit \]](#)

Level 2 is considered automated driving, but not autonomous driving. A Level 2 driving system expects a driver to be fully aware at any time of the driving and traffic situation and be able to take over any moment.

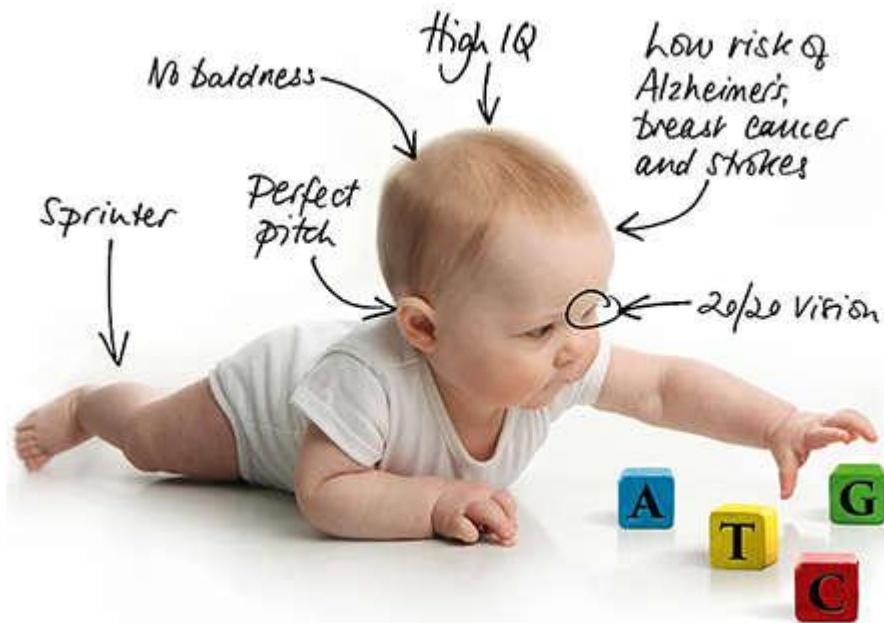
List of known automated driving system car fatalities (occurring while automated driving-system acknowledged to have been engaged)

Date	Incident no.	Country	City	State/county/province	No. of fatalities	System manufacturer	Vehicle Type	Distance driven by the system at time of incident	Notes
20 January 2016	1	China	Handan	Hebei	1	Tesla (Autopilot)	Model S ^[12]	—	Driver fatality. ^{[13][14]}
7 May 2016	2	United States of America (USA)	Williston	Florida	1	Tesla (Autopilot)	Model S ^[10]	130,000,000 mi 210,000,000 km ^{[15][16]}	Driver fatality. ^{[17][18]}
23 March 2018	4	United States of America (USA)	Mountain View	California	1	Tesla (Autopilot)	Model X ^[10]	—	Driver fatality. ^[19]

Dilemma: självkörande bilar



Dilemma: "designer babies"



Shutterstock

Till nästa gång

- Kan ni komma på ett ***möjligt*** moraliskt dilemma som kan uppstå i utvecklingen eller användandet av era designprojekt?
 - Vilka argument för eller emot ett visst agerande eller ställningstagande finns?
 - Grundar sig argumenten i ett regel- eller konsekvens-etiskt förhållningssätt?
- Kan ni dra en klar gräns om när ansvar faller på användare eller designer?

Också nästa gång

- Utveckling av empati
- Samarbete

