Robotics Reading Group
@ Instituto Superior Técnico

Session #7
31-01-2020

Raquel Oliveira

Reasons why I don’t drive

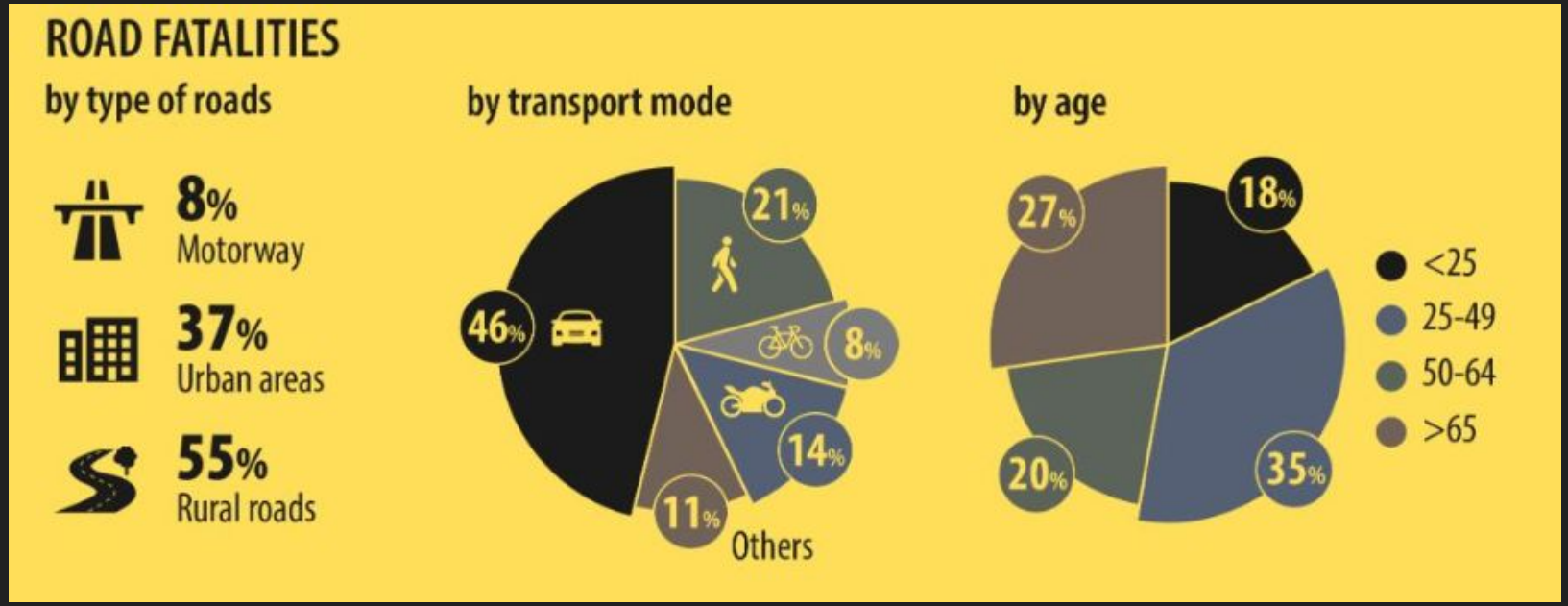
Reasons why I don’t drive

Driving is dangerous!

Although the number of road fatalities has decreased ~58% since 2001, it is estimated that 25.300 people die every year and 135.000 get seriously injured as a result of car accidents.

EPRS, European Commission, 2017

95% of all accidents involve human error



EPRS, European Commission, 2017

Technologies & improved road safety

"*Compulsory safety technologies could help save more than 25.000 lives and avoid at least 140.000 serious injuries by 2038* (...)."

EPRS, European Commission, 2017

Human decision-making biases in the moral dilemmas of autonomous vehicles

“The challenge in building machine morality based on people’s moral decisions [...] is accounting for the biases in human moral decision-making.”

How should autonomous vehicles make decisions about road safety when interacting with humans?

Human decision-making biases in the moral dilemmas of autonomous vehicles

How do people make these decisions in a daily context?

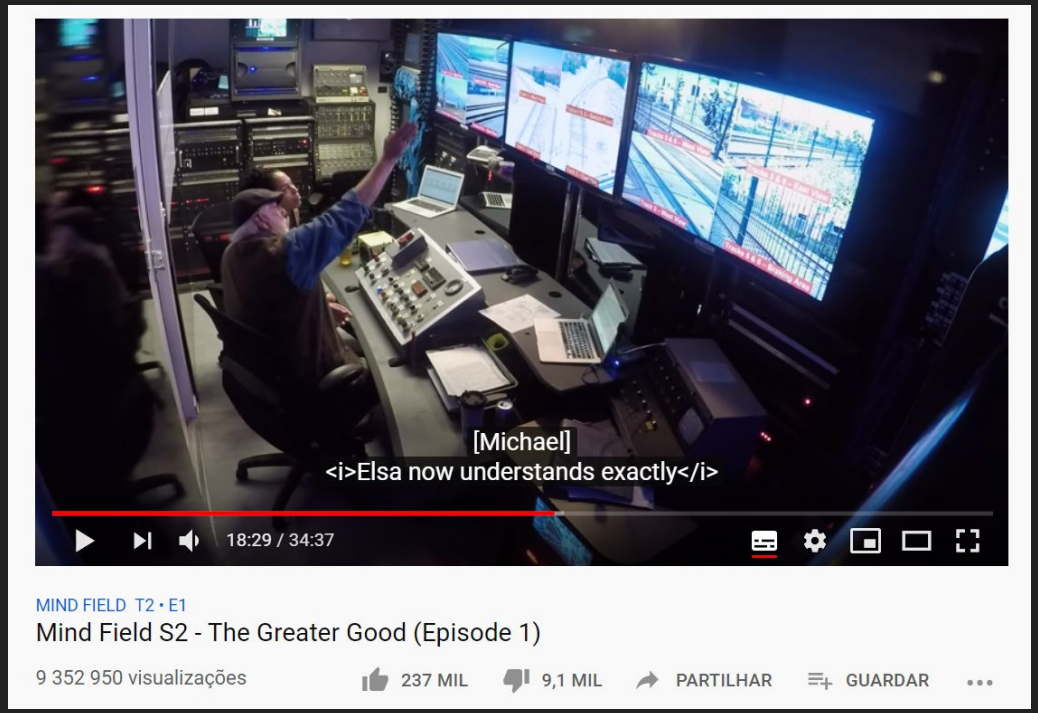
How would people like/expect autonomous cars to make decisions?

Human decision-making biases in the moral dilemmas of autonomous vehicles

How do people make these decisions in a daily context?

How would people like/expect autonomous cars to make decisions?

Human decision-making biases in the moral dilemmas of autonomous vehicles



You can check the full video here:

https://www.youtube.com/watch?v=1sl5KJ69qiA

How do we make decisions in this context?

It depends on the nature of the situation….

Deliberate decision-making is more likely to be guided by utilitarian principles

How do we make decisions in this context?

It depends on the nature of the situation….

Deliberate decision-making is more likely to be guided by utilitarian principles

Intuitive decision-making is more likely to be guided by emotions and easily accessible rules (i.e. heuristics)

Cf. Dual-process theory

How do we make decisions in this context?

It depends on the nature of the situation….

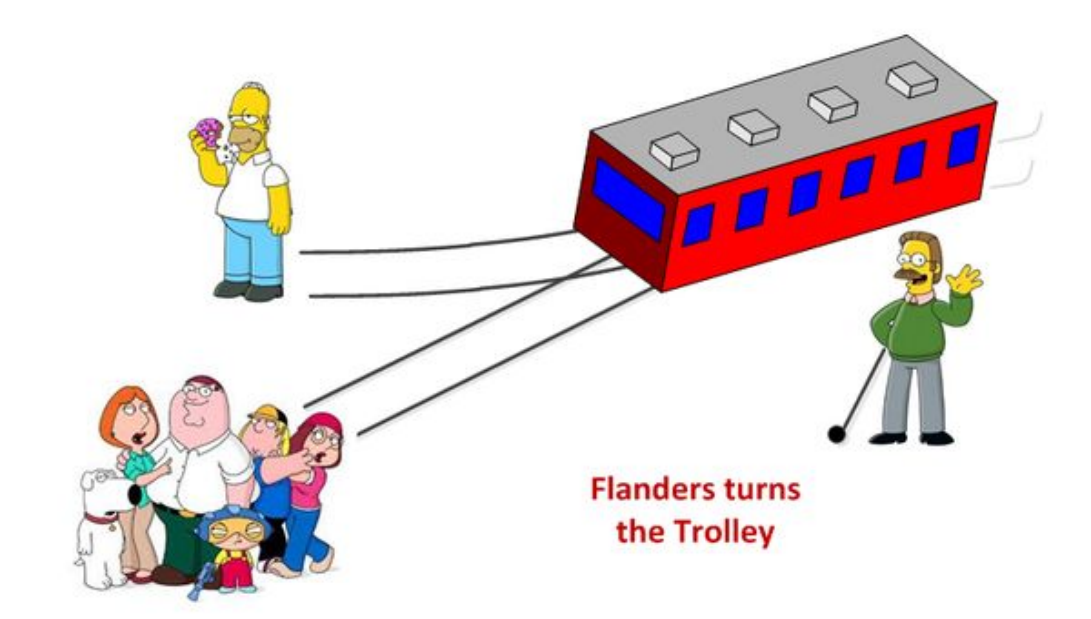
Deliberate decision-making is more likely to be guided by utilitarian principles

Intuitive decision-making is more likely to be guided by emotions and easily accessible rules (i.e. heuristics)

Cf. Dual-process theory

How do we make decisions in this context?

It depends on the nature of the situation…. and of the individual characteristics of the person making the decision.



Human decision-making biases in the moral dilemmas of autonomous vehicles

How do people make these decisions in a daily context?

How would people like/expect autonomous cars to make decisions?

Human decision-making biases in the moral dilemmas of autonomous vehicles

The paradox:

The inherent problem of peoples' preferences in moral dilemmas (...) is that people seem to favor a utilitarian moral doctrine (...) but they simultaneously report preferring an autonomous vehicle that is preprogrammed to protect themselves (...).

Human decision-making biases in the moral dilemmas of autonomous vehicles

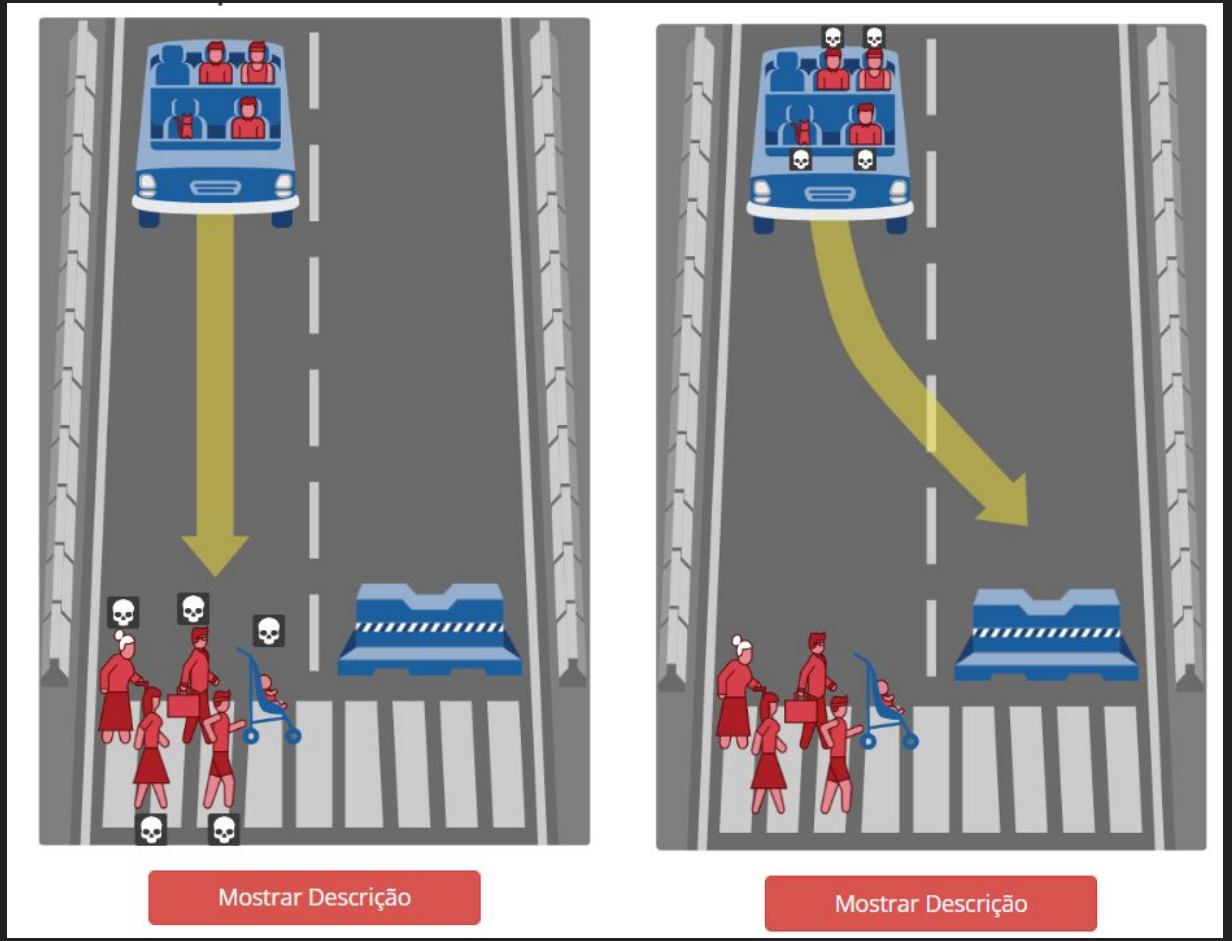
The approach:

Crowdsourcing moral norms for machine autonomous decision-making

The authors replicated the Moral Machine Dilemma

Collected ~12.000 decisions (U.S.A & Denmark)

Keep this in mind because we will
come back to this in the end



What should the autonomous car do?

Available at: http://moralmachine.mit.edu/

Human decision-making biases in the moral dilemmas of autonomous vehicles

Study 1

Determining the effect of the participant's perspective

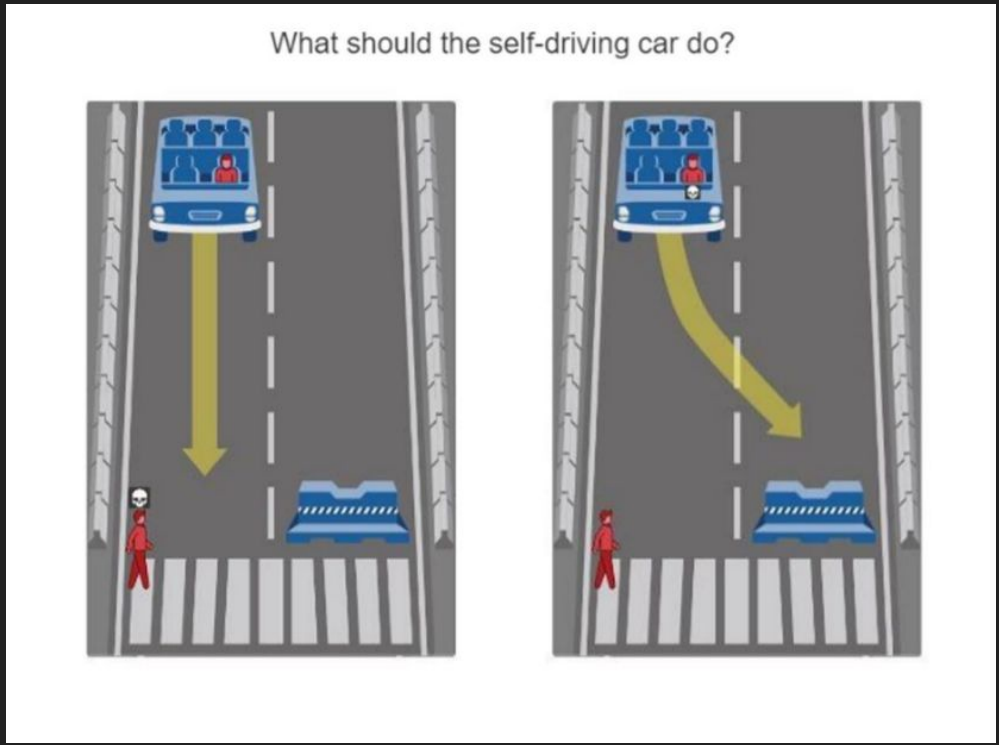
(n = 807, 46% female, avg age = ~32yo)

Imagine you are the passenger/driver …

Study 1

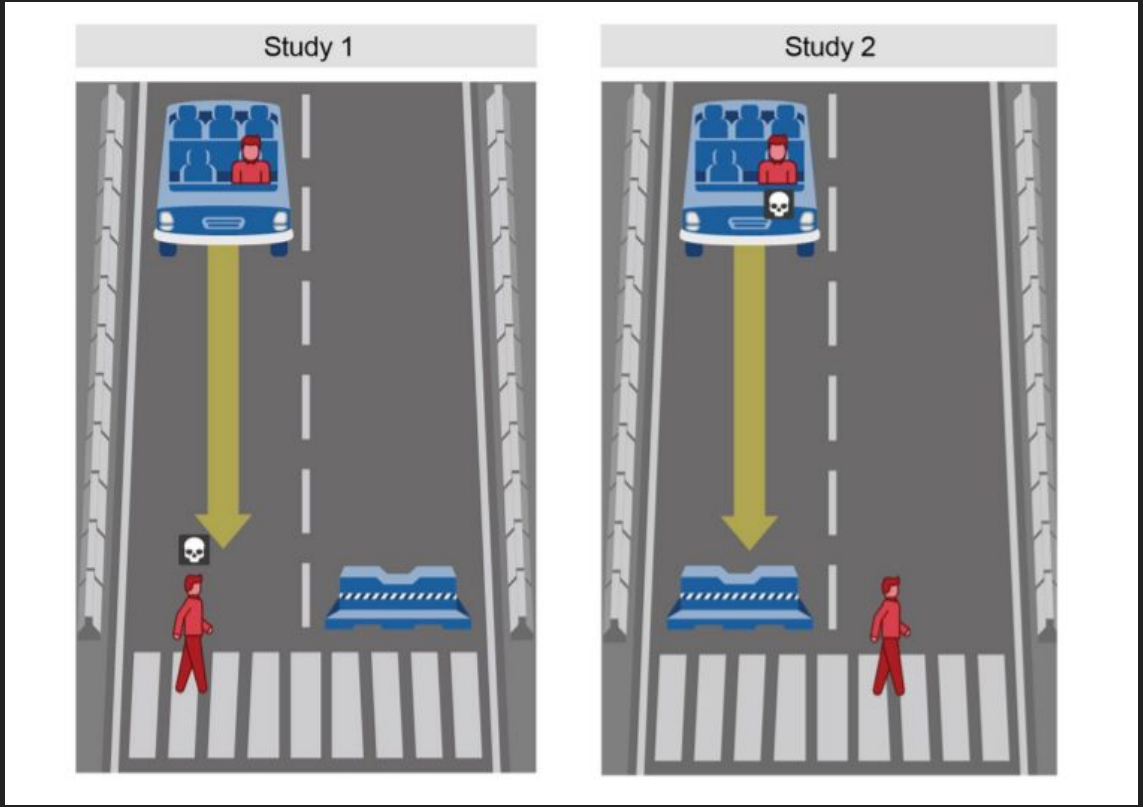
<u>Determining the effect of the participant's perspective</u>

(n = 807, 46% female, avg age = ~32yo)



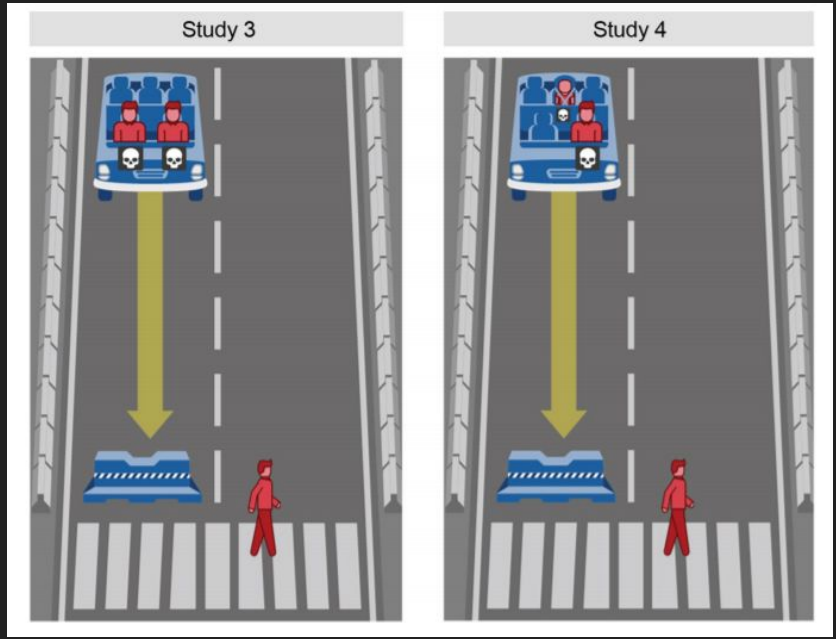
Study 1	Study 2
Determining the effect of the participant's perspective (n = 807, 46% female, avg age = ~32yo)	Inaction vs. Action (n = 848, 52% female, avg age = ~33yo)

Study 1	Study 2
Determining the effect of the participant's perspective (n = 807, 46% female, avg age = ~32yo)	Inaction vs. Action (n = 848, 52% female, avg age = ~33yo)



Study 1	Study 2	Study 3
Determining the effect of the participant's perspective (n = 807, 46% female, avg age = ~32yo)	Inaction vs. Action (n = 848, 52% female, avg age = ~33yo)	Moral doctrines (n = 393, 51% female, avg age = ~38yo)

Study 1	Study 2	Study 3
Determining the effect of the participant's perspective (n = 807, 46% female, avg age = ~32yo)	Inaction vs. Action (n = 848, 52% female, avg age = ~33yo)	Moral doctrines (n = 393, 51% female, avg age = ~38yo)



Study 1	Study 2	Study 3	Study 4	Study 5
Determining the effect of the participant's perspective (n = 807, 46% female, avg age = ~32yo)	Inaction vs. Action (n = 848, 52% female, avg age = ~33yo)	Moral doctrines (n = 393, 51% female, avg age = ~38yo)	The value of life (n = 428, 51% female, avg age = ~36yo)	Agency (n = 302, 56% female, avg age = ~37yo)

Study 1

Determining the effect of the participant's perspective

(n = 807, 46% female, avg age = ~32yo)

Study 2

Inaction vs. Action

(n = 848, 52% female, avg age = ~33yo)

Study 3

Moral doctrines

(n = 393, 51% female, avg age = ~38yo)

Study 4

The value of life

(n = 428, 51% female, avg age = ~36yo)

Study 5

Agency

(n = 302, 56% female, avg age = ~37yo)

In all of the studies, the authors manipulated the decision-making mode by implementing a 30s time limit (or no limit).

Study 6

Social norms violation

(n = 608, 59% female, avg age = ~35yo)

Study 7

Systematic Validation

Study 1

Determining the effect of the participant's perspective

(n = 807, 46% female, avg age = ~32yo)

Study 2

Inaction vs. Action

(n = 848, 52% female, avg age = ~33yo)

Study 3

Moral doctrines

(n = 393, 51% female, avg age = ~38yo)

Study 4

The value of life

(n = 428, 51% female, avg age = ~36yo)

Study 5

Agency

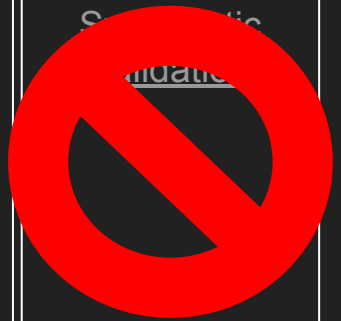
(n = 302, 56% female, avg age = ~37yo)

Study 6

Social norms violation

(n = 608, 59% female, avg age = ~35yo)

Study 7



Summarizing….

2 (Deliberate vs. Intuitive) X

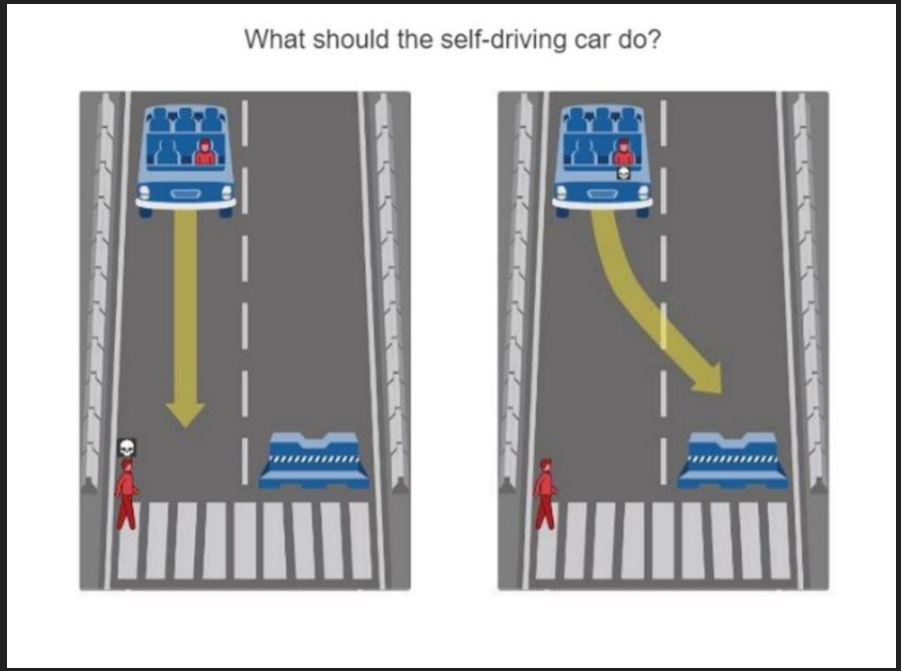
-3 (Perspective: Passenger vs. Pedestrian vs. Observer);
-2 (Taking action: Action vs. No action);
-2 (Moral doctrines: Utilitarian vs. deontological) ;
- 2 (The value of life: With child vs. No child);
-2 (Agency: Driver vs. Backseat);
-3 Social Norms Violation (High norm violation vs. Low norm violation vs. No violation)

Human decision-making biases in the moral dilemmas of autonomous vehicles

Perspective (S1)

People tend to sacrifice other's more often when they see them from a different perspective and vice-versa.

This difference is larger when participants were asked to make a deliberate choice than when they were asked to make an intuitive choice.



Human decision-making biases in the moral dilemmas of autonomous vehicles

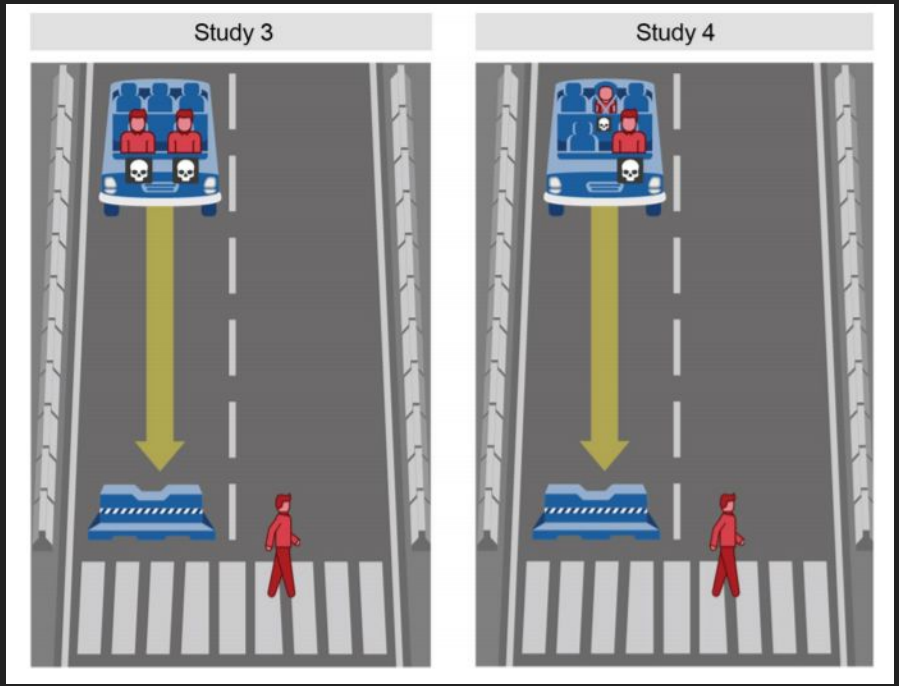
Action vs. Inaction (S2)

The authors found no effect of the outcome of the default choice in the participant's decisions in the dilemma.

Human decision-making biases in the moral dilemmas of autonomous vehicles

Moral doctrine (S3)

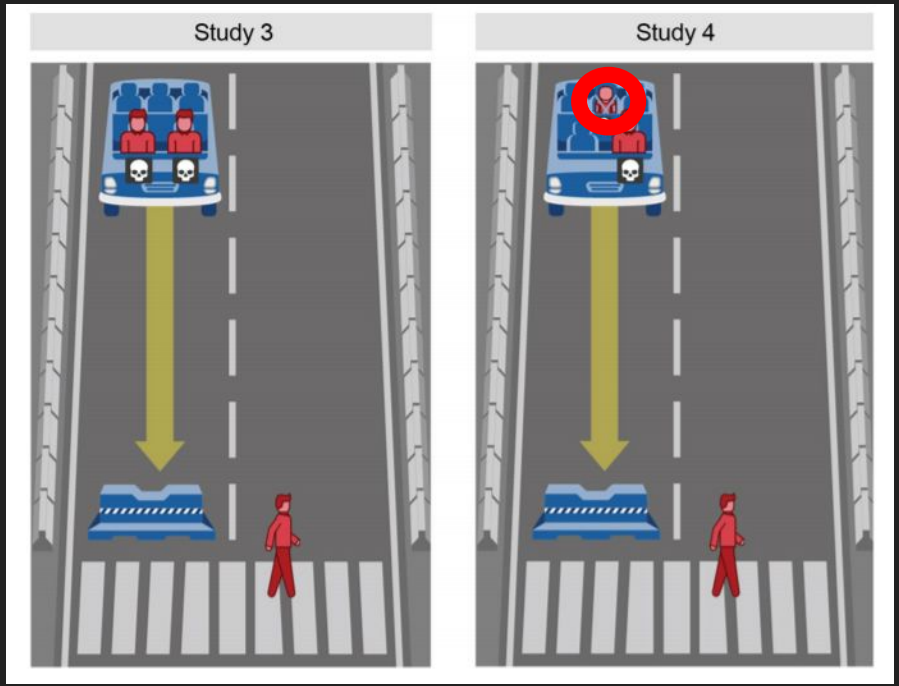
The likelihood of people sacrificing the pedestrian increases as the number of people in the car increases.



Human decision-making biases in the moral dilemmas of autonomous vehicles

The value of life (S4)

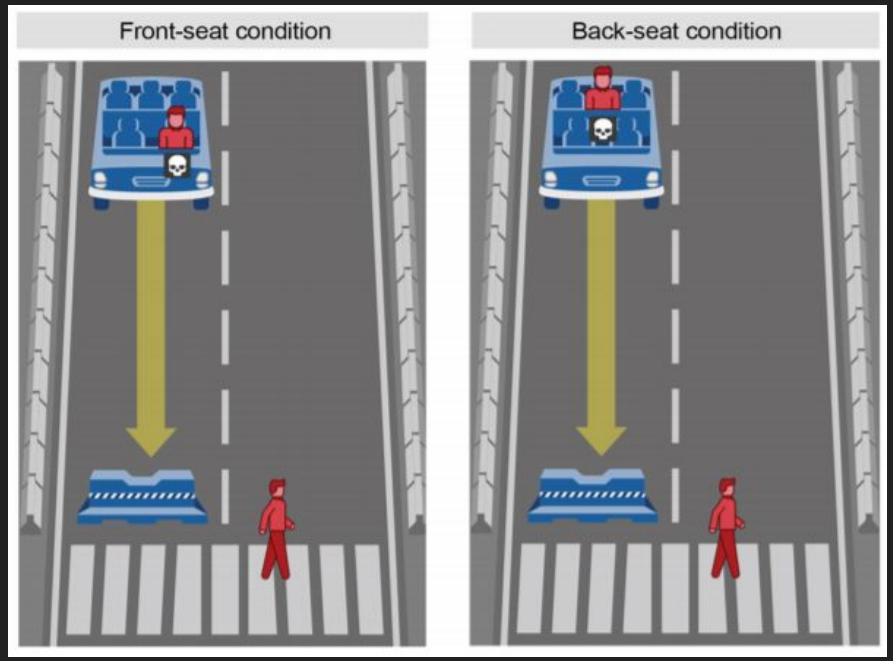
Participants are less protective of the child when they are asked to take a decision from the pedestrian perspective and vice-versa. Deliberate decisions tend to favour the child in comparison to intuitive decisions.



Human decision-making biases in the moral dilemmas of autonomous vehicles

Agency (S5)

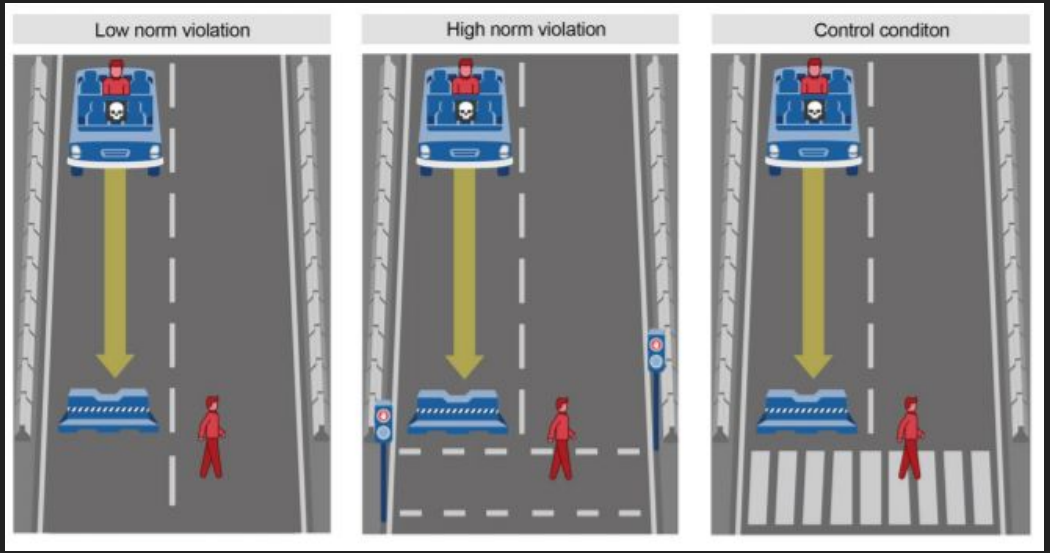
Participants are more likely to sacrifice the pedestrian when the passenger is in the back seat, compared when (s)he is on the drivers' seat.



Human decision-making biases in the moral dilemmas of autonomous vehicles

Social Norms Violation (S6)

Participants are less likely to sacrifice the pedestrian in the high norm violation condition than in the other two conditions.



Discussion and limitations

Cultural differences - Denmark & U.S.A. are highly individualistic countries (Hofstede scores > 70) - How would less individualistic countries perceive the trade off between pedestrian and passenger lives?

Discussion and limitations

Cultural differences - Denmark & U.S.A. are highly individualistic countries (Hofstede scores > 70) - How would less individualistic countries perceive the trade off between pedestrian and passenger lives?

The reaction time people have in real-life scenarios is much shorter than 30s.

Discussion and limitations

Cultural differences - Denmark & U.S.A. are highly individualistic countries (Hofstede scores > 70) - How would less individualistic countries perceive the trade off between pedestrian and passenger lives?

The reaction time people have in real-life scenarios is much shorter than 30s.

Difference between anticipated behaviour and actual behaviour

Discussion and limitations

Cultural differences - Denmark & U.S.A. are highly individualistic countries (Hofstede scores > 70) - **How would less individualistic countries perceive the trade off between pedestrian and passenger lives?**

The reaction time people have in real-life scenarios is much shorter than 30s.

Difference between anticipated behaviour and actual behaviour

How does this decision-making affect persons' decision to accept and use autonomous cars?

Thank you all for coming!