



Trolley problem & autonomous cars

**Eva Zackova, Jan Romportl, Daniel
Soutner,
Man-Machine Interaction Dep.,
University of West Bohemia, Czechia**

Motivation

- Horse → car → AC?
- Unknown problems waiting
- AC already in testing; legislation in progress
- Our expectation of AC ethics?
- AC not in common experience: thought experiments

Questionnaire

- on-line survey
- from June 2015 to April 2016
- questions:
 - trolley dilemma & AC dilemma scenarios,
 - basic socio-demographic information,
 - Five Factor Model / Big Five & Temperament and Character Inventory test
 - ...

Sample

- 430 respondents (41% women, 59% men)
- median age 25 (average 28)
- all participants from the Czech Republic; mostly from the western part of the country
- our sample deviates from national average of higher educated people (49% vs 25%)

Trolley problem: switch

A train is speeding on a main track. In the distance, you can see five people tied and lying on the same track, unable to move. The train is speeding towards them. You are standing nearby next to a lever. If you pull the lever, the train switches to the side track, where there is another single man tied and lying on the track, unable to move.

- **What would you do in that situation?**

- You rather do nothing and the train is going to kill five people on the main track.
- You pull the lever and the train is going to kill one man in the side track.

- **To make that decision about the described situation was for you:**

very easy – somewhat easy – neither easy nor difficult – somewhat difficult – very difficult

Trolley problem: big man

A train is speeding on a main track right towards five people tied and lying on the same track, unable to move. You are watching the situation from the bridge above the track. On the bridge, there is a big man standing next to you, who could surely stop the train with his body.

- **What would you do in that situation?**
 - You rather do nothing and the train is going to kill five people on the main track
 - You push the big man from the bridge to the track, where the train will kill him and stop

AC problem: “sitting”

You are sitting in an autonomous car, which is going on a narrow mountain road, approaching a narrow tunnel. Just before your car enters the tunnel, a man suddenly gets on the road in front of your car, trying to cross it. He/she suddenly slips, falls down and blocks the path.

- **What should your autonomous car do in that situation?**
 - Run over that human and kill him/her.
 - Turn the steering out of the road, hit the wall and kill you.

AC problem: “voting”

An autonomous car is going on a very narrow mountain road with one passenger on board, approaching a narrow tunnel. Just before your car enters the tunnel, a man suddenly gets on the road in front of your car, trying to cross it. He/she suddenly slips, falls down in the middle of the road and blocks the path.

- **Imagine that you are voting in a referendum on a new legislation determining how autonomous cars will respond in such situation. What should the autonomous cars do according to your opinion?**
 - In such situation, autonomous cars should protect passengers. That means to run over that human and kill him/her.
 - In such situation, autonomous cars should protect pedestrians. That means to turn the steering out of the road, hit the wall and kill the passenger.

Results: gender

- 1-st person “sitting” scenario
 - moderately significant association ($p = 0.057$ in Fisher test) between gender and the dilemma outcome
 - easier for men to decide whether to kill themselves or a pedestrian
- 3-rd person “voting” scenario
 - no difference
- Gender bias manifest itself in the more suggestive 1-st person perspective scenario

Results: active driver

- 1-st person “sitting” scenario
 - **no significant difference** between drivers and non-drivers
- 3-rd person “voting” scenario
 - **more difficult** for drivers to decide than for non-drivers

Results: education

- education does not play any crucial role in the decision outcome of the AC dilemmas nor in the perceived level of difficulty of these decisions

Results: character traits

- 1-st person “sitting” scenario
 - significant association between *Cooperativeness* (more likely kill self), *Persistence* (less likely kill self), *Conscientiousness* (less likely kill self) and the decision whom to kill
 - statistically significant association between *Reward Dependence* and the difficulty of the decision (higher RD = more difficult)
- 3-rd person “voting” scenario
 - no statistically significant association

Results: character traits

- higher *Cooperativeness* is associated with how much people identify with and accept others
- higher *Persistence* generally means that people are less likely to give up in spite of difficulties, frustration or fatigue, hence probably not give up even in case of an autonomous car accident
- *Conscientiousness* associated with being organized and highly planned (even stubborn)
- *Self-transcendence* association with the dilemmas outcomes was not confirmed
- lower psychological bias in the “voting” scenario

Results: trolley problem

- trolley problem outcome and the AC dilemma outcome: no significant association
- positive association between the **difficulty of answering the first trolley problem** question and the **AC dilemma in both scenarios**

Results: predictive analysis

- From all input variables (gender, age, character traits,...)
- Random forest, logistic regression
- We can not predict AC dilemma outcome in both scenarios

Results: 1st “sitting” scenario vs 3rd “voting”

- no difference in the mean values of the difficulty rating in the scenarios (difficult for some groups of respondents, whereas for others it is easier)
- not a statistically significant difference ($p = 0.15$) in the probabilities of a pedestrian being killed (73%-80%)

Conclusion

- the evidence from our data is still too weak ($n=430$) to firmly confirm that no such a difference can exist
- pool of respondents was not gender-balanced
- 3-rd person scenario is less biased than the 1-st person scenario (only active drivers)

