Self-driving car dilemmas reveal that moral choices are not universal

Survey maps global variations in ethics for programming autonomous vehicles.

When a driver slams on the brakes to avoid hitting a pedestrian crossing the road illegally, she is making a moral decision that shifts risk from the pedestrian to the people in the car. Self-driving cars might soon have to make such ethical judgments on their own — but settling on a universal moral code for the vehicles could be a thorny task, suggests a survey of 2.3 million people from around the world.

The largest ever survey of machine ethics1, published today in *Nature*, finds that many of the moral principles that guide a driver's decisions vary by country. For example, in a scenario in which some combination of pedestrians and passengers will die in a collision, people from relatively prosperous countries with strong institutions were less likely to spare a pedestrian who stepped into traffic illegally.

"People who think about machine ethics make it sound like you can come up with a perfect set of rules for robots, and what we show here with data is that there are no universal rules," says Iyad Rahwan, a computer scientist at the Massachusetts Institute of Technology in Cambridge and a co-author of the study.

The survey, called the Moral Machine, laid out 13 scenarios in which someone's death was inevitable. Respondents were asked to choose who to spare in situations that involved a mix of variables: young or old, rich or poor, more people or fewer.

People rarely encounter such stark moral dilemmas, and some critics question whether the scenarios posed in the quiz are relevant to the ethical and practical questions surrounding driverless cars. But the study's authors say that the scenarios stand in for the subtle moral decisions that drivers make every day. They argue that the findings reveal cultural nuances that governments and makers of self-driving cars must take into account if they want the vehicles to gain public acceptance.

At least one company working on self-driving cars — the German carmaker Audi — says that the survey could help prompt an important discussion about these issues. (Other firms with autonomous-vehicle programmes, including auto manufacturer Toyota and technology companies Waymo and Uber, declined to comment on the findings.) And Nicholas Christakis, a social scientist at Yale University in New Haven, Connecticut, is fascinated by the results.

"It's a remarkable paper," he says. The debate about whether ethics are universal or vary between cultures is an old one, says Christakis, and now the "twenty-first century problem" of how to programme self-driving cars has reinvigorated it.

The road not taken

Some of the world's biggest tech companies — including Google's parent, Alphabet; Uber; and Tesla — and carmakers now have self-driving car programmes. Many of these companies argue that the vehicles could improve road safety, ease traffic and improve fuel efficiency. Social scientists say the cars raise ethical issues, and could have unintended consequences for public safety and the environment.

In 2016, Rahwan's team stumbled on an ethical paradox about self-driving cars[2]: in surveys, people said that they wanted an autonomous vehicle to protect pedestrians even if it meant sacrificing its passengers — but also that they wouldn't buy self-driving vehicles programmed to act this way.

Curious to see if the prospect of self-driving cars might raise other ethical condundrums, Rahwan gathered an international team of psychologists, anthropologists and economists to create the Moral Machine. Within 18 months, the online quiz had recorded 40 million decisions made by people from 233 countries and territories.

No matter their age, gender or country of residence, most people spared humans over pets, and groups of people over individuals. These responses are in line with rules proposed in what may be the only governmental guidance on self-driving cars: a 2017 report by the German Ethics Commission on Automated and Connected Driving.

But agreement ends there. When the authors analysed answers from people in the 130 countries with at least 100 respondents, they found that the nations could be divided into three groups. One contains North America and several European nations where Christianity has historically been the dominant religion; another includes countries such as Japan, Indonesia and Pakistan, with strong Confucian or Islamic traditions. A third group consists of Central and South America, as well as France and former French colonies. The first group showed a stronger preference for sacrificing older lives to save younger ones than did the second group, for example.

The researchers also identified correlations between social and economic factors in a country and the average opinions of its residents. The team found that people from countries with strong government institutions, such as Finland and Japan, more often chose to hit people who were crossing the road illegally than did respondents in nations with weaker institutions, such as Nigeria or Pakistan.

Scenarios that forced survey participants to choose whether to save a homeless person on one side of the road or an executive on the other revealed another point of departure: the choices people made often correlated with the level of economic inequality in their culture. People from Finland — which has a relatively small gap between the rich and the poor — showed little preference for swerving one way or the other. But the average respondent from Colombia — a country with significant economic disparity — chose to kill the lower-status person.

Azim Shariff, a psychologist at the University of British Columbia in Vancouver, finds this result interesting because it suggests that the survey really does reveal people's moral preferences. "If you assume that places that have a lower level of income inequality have political policies that favor egalitarianism, this shows that the moral norms that support those policies are expressed in the way that people play these games."

Potholes ahead?

Although autonomous vehicles aren't yet for sale to the public, test versions are cruising through several US cities. By 2021, at least five manufacturers hope to have self-driving cars and trucks in wide use.

Bryant Walker Smith, a law professor at the University of South Carolina in Columbia, is sceptical that the Moral Machine survey will have any practical use. He says that the study is unrealistic because there are few instances in real life in which a vehicle would face a choice between striking two different types of people. "I might as well worry about how automated cars will deal with asteroid strikes," Walker Smith says.

MORAL COMPASS

A survey of 2.3 million people worldwide reveals variations in the moral principles that guide drivers' decisions. Respondents were presented with 13 scenarios, in which a collision that killed some combination of passengers and pedestrians was unavoidable, and asked to decide who they would spare. Scientists used these data to group countries and territories into three groups based on their moral attitudes.
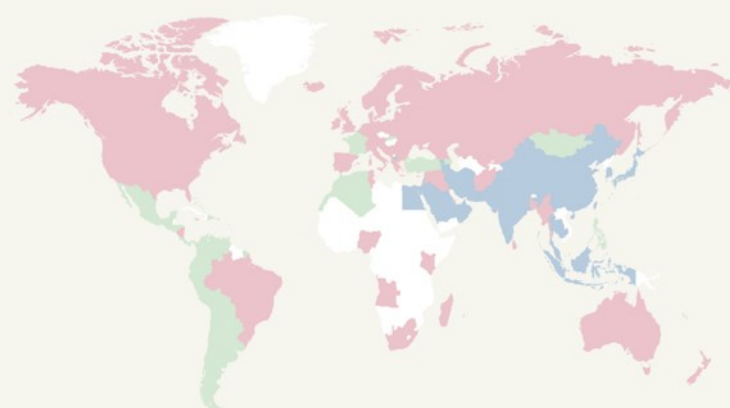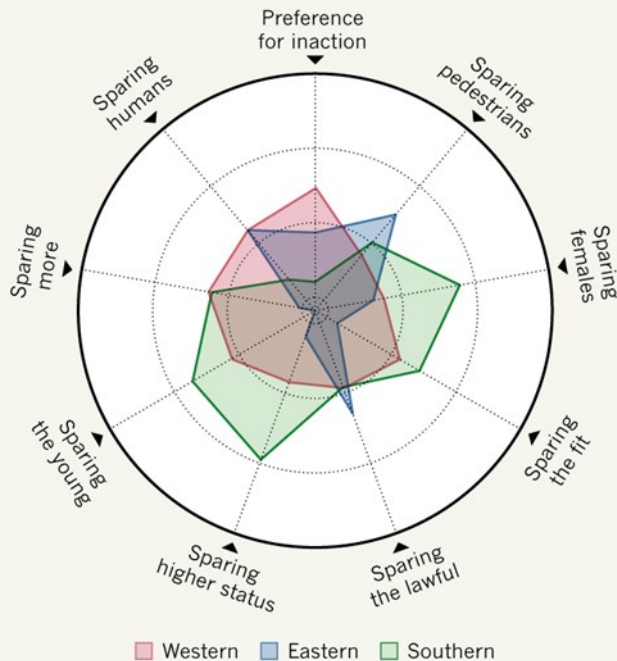
But the study's authors say that their scenarios represent the minor moral judgements that human drivers make routinely — which can sometimes be fatal. A driver who veers away from cyclists riding on a curvy mountain road increases her chance of hitting an oncoming vehicle. If the number of driverless cars on the road increases, so too will the likelihood that they will be involved in such accidents.

Some car companies are listening. Barbara Wege, who heads a group focused on autonomous-vehicle ethics at Audi in Ingolstadt, Germany, says such studies are valuable. Wege argues that self-driving cars would cause fewer accidents, proportionally, than human drivers do each year — but that events involving robots might receive more attention.

Surveys such as the Moral Machine can help to begin public discussions about these inevitable accidents that might foster trust. "We need to come up with a social consensus," she says, "about which risks we are willing to take."

Nature **562**, 469-470 (2018)

Ethics Commission on Automated Driving presents report

Dobrindt: First guidelines in the world for self-driving computers

Today, the Ethics Commission on Automated Driving set up by Federal Minister Alexander Dobrindt presented its report. In this report, the body of experts, headed by Professor Udo Di Fabio, a former Federal Constitutional Court judge, has developed guidelines for the programming of automated driving systems.

Mr Dobrindt said:

"In the era of the digital revolution and self-learning systems, human-machine interaction raises new ethical questions. Automated and connected driving is the most recent innovation where this interaction is to be found across the board. The Ethics Commission at the Federal Ministry of Transport and Digital Infrastructure has performed absolutely pioneering work in this field and developed the first guidelines in the world for automated driving. This means that we remain at the forefront of Mobility 4.0 on the international stage".

Prof. Di Fabio said:

"The Ethics Commission on Automated and Connected Driving has developed initial guidelines for policymakers and lawmakers that will make it possible to approve automated driving systems but that set out special requirements in terms of safety, human dignity, personal freedom of choice and data autonomy".

The Ethics Commission's report comprises 20 propositions. The key elements are:

- Automated and connected driving is an ethical imperative if the systems cause fewer accidents than human drivers (positive balance of risk).
- Damage to property must take precedence over personal injury. In hazardous situations, the protection of human life must always have top priority.
- In the event of unavoidable accident situations, any distinction between individuals based on personal features (age, gender, physical or mental constitution) is impermissible.
- In every driving situation, it must be clearly regulated and apparent who is responsible for the driving task: the human or the computer.
- It must be documented and stored who is driving (to resolve possible issues of liability, among other things).
- Drivers must always be able to decide themselves whether their vehicle data are to be forwarded and used (data sovereignty).

The Federal Ministry of Transport and Digital Infrastructure's Ethics Commission comprises 14 academics and experts from the disciplines of ethics, law and technology. Among these are transport experts, legal experts, information scientists, engineers, philosophers, theologians, consumer protection representatives as well as representatives of associations and companies.

The Ethics Commission's complete report can be found here:

https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.pdf?__blob=publicationFile