

A moral map for AI cars

Survey reveals global variations in ethical rules of the road for autonomous vehicles.

BY AMY MAXMEN

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 around the world.

The largest-ever survey of machine ethics¹, published this week 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, such as law enforcement, 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 study co-author Iyad Rahwan, a computer scientist at the Massachusetts

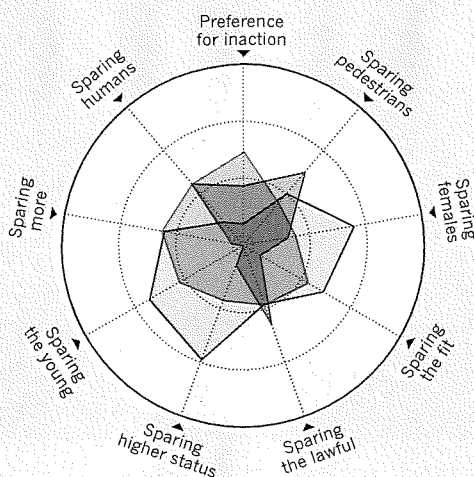
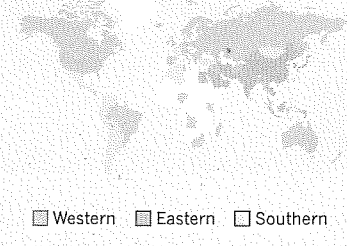
Institute of Technology in Cambridge.

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 ask whether the scenarios posed in the quiz are relevant to the ethical 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. The findings reveal ▶

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.



analysed answers from people in the 130 countries with at least 100 respondents, they found that the nations could be divided into three groups (see 'Moral compass'). One contains North America and several European and other nations where Christianity has historically been the dominant religion; another includes countries such as Japan, Indonesia and Pakistan, which have 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.

Test versions of autonomous cars 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, 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 person. “I might as well worry about how automated cars will deal with asteroid strikes,” he says.

But Barbara Wege, who heads a group focused on autonomous-vehicle ethics at the car manufacturer Audi in Ingolstadt, Germany, says that 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 prompt public discussions about inevitable accidents, and so might foster trust. “We need to come up with a social consensus,” she says, “about which risks we are willing to take.” ■

1. Awad, E. *et al.* *Nature* <https://doi.org/10.1038/s41586-018-0637-6> (2018).

2. Bonnefon, J. *et al.* *Science* **352**, 1573–1576 (2016).

▶ cultural nuances that governments and makers of self-driving cars must take into account if they want the vehicles to gain public acceptance, they say.

“It's a remarkable paper,” says Nicholas Christakis, a social scientist at Yale University in New Haven, Connecticut. The debate about whether ethics are universal or vary between cultures is an old one, he says, and now the “twenty-first-century problem” of how to program self-driving cars has reinvigorated it.

Some of the world's biggest tech companies — including Google, Uber and Tesla — and car-makers now have self-driving-car programmes. Many of these companies argue that the vehicles could improve road safety and ease traffic, but social scientists say the cars raise complex ethical issues.

In 2016, Rahwan's team stumbled on a paradox about self-driving cars²: in surveys, people

say they want an autonomous vehicle to protect pedestrians, even if it means sacrificing its passengers — but also that they wouldn't buy self-driving vehicles programmed to act in this way.

Curious to see whether the prospect of self-driving cars might raise other ethical conundrums, Rahwan gathered psychologists, anthropologists and economists to create the online Moral Machine quiz. Within 18 months, it 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 might 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