

How people think that self-driving cars should behave in an accident

In a paper just published in Nature, a team of psychologists and computer scientists describe a different approach. They created the “Moral Machine”, a website which presents visitors with a series of choices about whom to save and whom to kill.

The strongest preferences, expressed by respondents from all over the world, were for saving human lives over animal ones, preferring to save many rather than few and prioritizing children over the old. There were weaker preferences for saving women over men, pedestrians over passengers in the car and for taking action rather than doing nothing. Criminals were seen as literally subhuman—ranking below dogs in the public’s priority list, but above cats. It is easy to imagine the **utilitarian** argument for preserving the lives of doctors over others. Humanity’s (weak) preference for saving athletes seems less intuitive.

utilitarian 功利主义的

Many people, says Dr Rahwan, a computer scientist at MIT and one of the paper’s authors, dismiss the trolley problem as a piece of pointless hypothesizing that is **vanishingly** unlikely to arise in real life. He is **unconvinced**. The specific situations posed by the website may hardly ever occur, he says. But all sorts of choices made by the firms producing self-driving cars will affect who lives and who dies in indirect, statistical ways. He gives the example of **overtaking** cyclists: “If you stay relatively near to the cycle lane, you’re increasing the chance of hitting a cyclist, but reducing the chance of hitting another car in the next lane over,” he says. “Repeat that over hundreds of millions of trips, and you’re going to see a skew in the [accident] statistics.”

vanishingly 趋于零地；难以察觉地

unconvinced 不信服的；怀疑的

overtaking 超车