

To crash or swerve? Study reveals which actions taken by self-driving cars are morally defensible

October 9 2018

A crash by one of Uber Technologies, Inc.'s self-driving cars earlier this year resulted in the first pedestrian death associated with self-driving technology. The incident highlighted the challenges technology companies are facing in developing software that can adequately detect and respond to hazards in the road and immediate surroundings. The vehicle could have come to a complete stop in three seconds but did not employ emergency braking until 1.3 seconds before impact. Should emergency braking have been the default action taken immediately following the detection of roadway hazard?

A new study, "How should autonomous cars drive? A preference for defaults in moral judgments under risk and uncertainty," published in *Risk Analysis: An International Journal* addressed this challenge by asking the public what they believed would be the most morally and ethically sound behavior for an autonomous vehicle (AV) faced with an oncoming collision. Even a perfectly functioning AV will not be able to avoid every collision and in some situations, every option will result in some type of crash.

The research team, comprised of Björn Meder, Nadine Fleischhut, and Nina-Carolin Krumnau of the Max Planck Institute for Human Development and Michael R. Waldmann of the University of Göttingen, addressed this question by asking participants to choose between staying in their lane (and braking), or swerving, where each action could lead to

a collision with another road user at varying degrees of uncertainty.

The key finding from this study is that people generally preferred for the car to stay in its lane and perform an emergency stop. This supports the idea that people consider the stay option a reasonable default, as it conforms to general rules of driving and provides a better degree of controllability, even if it does not minimize expected loss. Employing this action as a simple default rule requires less processing of information and will often lead to better results. The researchers also found that even if staying in the lane resulted in an accident, people were less likely to alter their moral evaluation of the action taken in retrospect, whereas a bad outcome heavily influenced a retrospective analysis of a decision to swerve out of the lane.

In their first experiment, participants were presented with a scenario in which an AV had to perform one of two maneuvers: stay in the lane or swerve. Staying in the lane puts a pedestrian in the street in danger while swerving puts a bystander on the sidewalk at risk. The likelihood of colliding with the pedestrian and the bystander were varied creating different scenarios with specified or unknown risks. Data were collected from 872 individuals online through the Amazon Mechanical Turk (AMT) platform. Participants were presented with a written description of a traffic scenario in which a car is traveling down a road when suddenly a pedestrian appears in its path. The car can either stay in the lane and perform an emergency stop, in which case it might collide with the pedestrian, or the car can swerve to the right and perform an emergency stop, in which case it might collide with a bystander on the sidewalk.

The likelihood of colliding with the pedestrian in the street was either 20 percent, 50 percent or 80 percent. In the 'risk condition' the likelihood of colliding with the bystander was 50 percent. In the 'uncertainty' condition, the likelihood was unknown because the car's systems were

unable to make the estimate. The results showed a general preference for staying in the lane with more than 85 percent of subjects opting to stay in the lane. When the likelihood of colliding with the bystander was unknown, about 70 percent of subjects still preferred to stay in the lane. When the likelihood of colliding with the pedestrian was 20 percent and the likelihood of colliding with the bystander was 50 percent no one opted to swerve. Even when the likelihood for both collisions was 50 percent, staying was considered more acceptable than swerving.

The second experiment examined how people morally evaluate AV behavior in retrospect when a collision has occurred. From a policy perspective, AVs should act in ways that society deems acceptable even if collisions do occur. The researchers once again recruited 766 subjects via the AMT platform. The participants were asked how an AV should perform in a specific situation and to evaluate the moral acceptability of both staying and swerving. The results showed that if the car stayed in its lane, the outcome of the situation (collision or no collision with another road user) did not affect participants' judgments of how an AV should behave. If the car swerved, however, the outcome was highly persuasive in retrospect. If no collision occurred, about 40 percent preferred to swerve, but less than 20 percent held that preference when a collision did occur. Thus, even when a [collision](#) occurred, staying in the lane was considered more acceptable.

"Our research highlights the importance of gaining a better understanding of how people think about the behavior of autonomous vehicles under different degrees of uncertainty," states Meder. "The findings will help to inform policy making and public discussion of the ethical implications of technological advances that will transform society in a variety of ways."

Overall, there is a general preference for staying in the lane and it is a morally acceptable default option in critical traffic situations even if it

does not minimize expected loss. This simple default requires no information to be gathered by the AV's systems about alternative actions or probabilities. Despite these findings, policy makers are faced with the challenge of developing policies that are morally sound but also appeal to the general public's desire for self-preservation. Previous studies (Bonneton, Shariff, & Rahwan, 2016, *Science*) have shown that subjects will show a preference for AVs that would sacrifice their own passengers to minimize the total number of casualties, but they wanted their own AVs to put a premium on passenger safety. So, while this study demonstrates a general acceptability for a default action to minimize potential losses, AV owners would prefer actions intended to save the vehicle passengers.

More information: Björn Meder et al, How Should Autonomous Cars Drive? A Preference for Defaults in Moral Judgments Under Risk and Uncertainty, *Risk Analysis* (2018). [DOI: 10.1111/risa.13178](https://doi.org/10.1111/risa.13178)

Provided by Society for Risk Analysis

Citation: To crash or swerve? Study reveals which actions taken by self-driving cars are morally defensible (2018, October 9) retrieved 18 April 2024 from <https://phys.org/news/2018-10-swerve-reveals-actions-self-driving-cars.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--