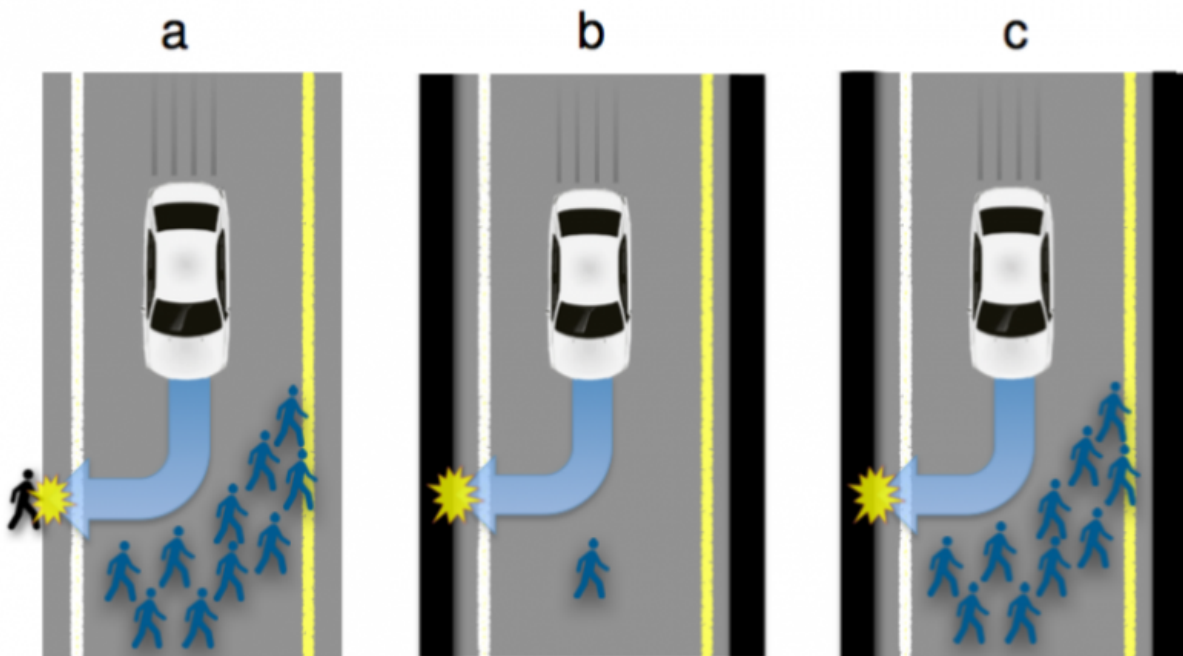


# When self-driving cars drive the ethical questions

October 24 2015, by Nancy Owano



Three traffic situations involving imminent unavoidable harm. (a) The car can stay on course and kill several pedestrians, or swerve and kill one passer-by. (b) The car can stay on course and kill one pedestrian, or swerve and kill its passenger. (c) The car can stay on course and kill several pedestrians, or swerve and kill its passenger. Credit: arXiv:1510.03346 [cs.CY]

Driverless cars are due to be part of day to day highway travel. Beyond their technologies and safety reports lies a newer wrinkle posed by three

researchers, in the form of ethical questions which policy makers and vendors will need to explore.

"Autonomous Vehicles Need Experimental Ethics: Are We Ready for Utilitarian Cars?" is by Jean-Francois Bonnefon, Azim Shariff and Iyad Rahwan. They are from Toulouse School of Economics, University of Oregon and MIT.

We are told that [self-driving cars](#) are capable of preventing road accidents and deaths in significant numbers. More recent discussions, though, point to the next chapter of driverless cars as to be more complex. When self-driving cars first appear on roads, the safety picture may be nuanced, say experts, where people are not necessarily prepared for the abundance of caution used by automated [drivers](#).

Now three researchers are adding to the mix of concerns to think about. That is, they are posing [ethical questions](#) raised by the presence of self-driving cars. Their paper on *arXiv* poses traffic situations involving unavoidable harm.

The question is over assessing the relative morality of different algorithms—who gets harmed and who gets spared. (a) The car can stay on course and kill several pedestrians, or swerve and kill one passer-by (b) The car can stay on course and kill one pedestrian, or swerve and kill its passenger (c) The car can stay on course and kill several pedestrians, or swerve and kill its passenger.

Is the passenger killed to save the other people? One may consider *MIT Technology Review's* calling it "an impossible ethical dilemma of algorithmic morality."

The authors' abstract stated, "It is a formidable challenge to define the algorithms that will guide AVs confronted with such moral dilemmas. In

particular, these moral algorithms will need to accomplish three potentially incompatible objectives: being consistent, not causing public outrage, and not discouraging buyers. We argue to achieve these objectives, manufacturers and regulators will need psychologists to apply the methods of experimental ethics to situations involving AVs and unavoidable harm."

Continued *MIT Technology Review*: "Should different decisions be made when children are on board, since they both have a longer time ahead of them than adults, and had less agency in being in the car in the first place? If a manufacturer offers different versions of its moral algorithm, and a buyer knowingly chose one of them, is the buyer to blame for the harmful consequences of the algorithm's [decisions](#)?"

To be sure, "Figuring out how to build ethical autonomous machines is one of the thorniest challenges in artificial intelligence today," the authors wrote. "As we are about to endow millions of vehicles with autonomy, taking algorithmic morality seriously has never been more urgent."

The authors believe answers are most likely to come from surveys employing the protocols of experimental ethics. Overall, they wrote, the field of experimental ethics offers key insights into the moral and legal standards that people expect from autonomous driving algorithms.

The researchers conducted three online surveys in June. The studies were programmed on Qualtrics survey software and recruited participants from the Mechanical Turk platform, for a compensation of 25 cents.

Results? They were "interesting," said *MIT Technology Review*, "if predictable. In general, people are comfortable with the idea that self-driving vehicles should be programmed to minimize the death toll."

The authors said, "Three surveys suggested that respondents might be prepared for [autonomous vehicles](#) programmed to make utilitarian moral decisions in situations of unavoidable harm. This was even true, to some extent, of situations in which the AV could sacrifice its owner in order to save the lives of other individuals on the road."

Offering his reflections on the research, Dave Gershgorn in *Popular Science* wrote, "Sure, [driverless cars](#) can reduce traffic fatalities by up to 90 percent. And like the field of ethics itself, what happens in the other 10 percent is still up for [debate](#)."

**More information:** Autonomous Vehicles Need Experimental Ethics: Are We Ready for Utilitarian Cars? arXiv:1510.03346 [cs.CY]  
[arxiv.org/abs/1510.03346](http://arxiv.org/abs/1510.03346)

### **Abstract**

The wide adoption of self-driving, Autonomous Vehicles (AVs) promises to dramatically reduce the number of traffic accidents. Some accidents, though, will be inevitable, because some situations will require AVs to choose the lesser of two evils. For example, running over a pedestrian on the road or a passer-by on the side; or choosing whether to run over a group of pedestrians or to sacrifice the passenger by driving into a wall. It is a formidable challenge to define the algorithms that will guide AVs confronted with such moral dilemmas. In particular, these moral algorithms will need to accomplish three potentially incompatible objectives: being consistent, not causing public outrage, and not discouraging buyers. We argue to achieve these objectives, manufacturers and regulators will need psychologists to apply the methods of experimental ethics to situations involving AVs and unavoidable harm. To illustrate our claim, we report three surveys showing that laypersons are relatively comfortable with utilitarian AVs, programmed to minimize the death toll in case of unavoidable harm. We give special attention to whether an AV should save lives by sacrificing

its owner, and provide insights into (i) the perceived morality of this self-sacrifice, (ii) the willingness to see this self-sacrifice being legally enforced, (iii) the expectations that AVs will be programmed to self-sacrifice, and (iv) the willingness to buy self-sacrificing AVs.

© 2015 Tech Xplore

Citation: When self-driving cars drive the ethical questions (2015, October 24) retrieved 16 April 2024 from <https://techxplore.com/news/2015-10-self-driving-cars-ethical.html>

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.