

To help autonomous vehicles make moral decisions, researchers ditch the 'trolley problem'

December 1 2023, by Matt Shipman



Credit: Unsplash/CC0 Public Domain

Researchers have developed a new experiment to better understand what people view as moral and immoral decisions related to driving vehicles,

with the goal of collecting data to train autonomous vehicles how to make "good" decisions. The work is designed to capture a more realistic array of moral challenges in traffic than the widely discussed life-and-death scenario inspired by the so-called "trolley problem."

The paper, "Moral judgment in realistic [traffic](#) scenarios: Moving beyond the trolley paradigm for ethics of autonomous vehicles," is [published](#) open access in the journal *AI & Society*.

"The trolley problem presents a situation in which someone has to decide whether to intentionally kill one person (which violates a moral norm) in order to avoid the death of multiple people," says Dario Cecchini, first author of a paper on the work and a postdoctoral researcher at North Carolina State University.

"In recent years, the trolley problem has been utilized as a paradigm for studying [moral judgment](#) in traffic," Cecchini says. "The typical situation comprises a binary choice for a [self-driving car](#) between swerving left, hitting a lethal obstacle, or proceeding forward, hitting a pedestrian crossing the street.

"However, these trolley-like cases are unrealistic. Drivers have to make many more realistic moral decisions every day. Should I drive over the speed limit? Should I run a red light? Should I pull over for an ambulance?"

"Those mundane decisions are important because they can ultimately lead to life-or-death situations," says Veljko Dubljević, corresponding author of the paper and an associate professor in the Science, Technology & Society program at NC State.

"For example, if someone is driving 20 miles over the speed limit and runs a red light, then they may find themselves in a situation where they

have to either swerve into traffic or get into a collision. There's currently very little data in the literature on how we make moral judgments about the decisions drivers make in everyday situations."

To address that lack of data, the researchers developed a series of experiments designed to collect data on how humans make moral judgments about decisions that people make in low-stakes traffic situations. The researchers created seven different driving scenarios, such as a parent who has to decide whether to violate a traffic signal while trying to get their child to school on time.

Each scenario is programmed into a virtual reality environment, so that study participants engaged in the experiment have audiovisual information about what drivers are doing when they make decisions, rather than simply reading about the scenario.

For this work, the researchers built on something called [the Agent Deed Consequence \(ADC\) model](#), which posits that people take three things into account when making a moral judgment: the agent, which is the character or intent of the person who is doing something; the deed, or what is being done; and the consequence, or the outcome that resulted from the deed.

Researchers created eight different versions of each traffic scenario, varying the combinations of agent, deed and consequence. For example, in one version of the scenario where a parent is trying to get the child to school, the parent is caring, brakes at a yellow light, and gets the child to school on time.

In a second version, the parent is abusive, runs a red light, and causes an accident. The other six versions alter the nature of the parent (the agent), their decision at the traffic signal (the deed), and/or the outcome of their decision (the consequence).

"The goal here is to have study participants view one version of each scenario and determine how moral the behavior of the driver was in each scenario, on a scale from 1 to 10," Cecchini says. "This will give us robust data on what we consider moral behavior in the context of driving a vehicle, which can then be used to develop AI algorithms for moral decision making in autonomous vehicles."

The researchers have done pilot testing to fine-tune the scenarios and ensure that they reflect believable and easily understood situations.

"The next step is to engage in large-scale data collection, getting thousands of people to participate in the experiments," says Dubljević. "We can then use that data to develop more interactive experiments with the goal of further fine-tuning our understanding of moral decision making. All of this can then be used to create algorithms for use in autonomous vehicles. We'll then need to engage in additional testing to see how those algorithms perform."

More information: Dario Cecchini et al, Moral judgment in realistic traffic scenarios: moving beyond the trolley paradigm for ethics of autonomous vehicles, *AI & Society* (2023). [DOI: 10.1007/s00146-023-01813-y](https://doi.org/10.1007/s00146-023-01813-y)

Provided by North Carolina State University

Citation: To help autonomous vehicles make moral decisions, researchers ditch the 'trolley problem' (2023, December 1) retrieved 29 April 2024 from <https://techxplore.com/news/2023-12-autonomous-vehicles-moral-decisions-ditch.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.