Self-Driving Cars: A Safer Future?

42,514 people died in (reported) traffic accidents in 2022 (NHTSA). While that number is low, it can always be closer to 0. Enter the fully self-driving car. The vision is that "[self-driving cars] are already safer than human drivers in many circumstances. Equipped with 360-degree vision, programmed to adhere to traffic laws, and immune to human distraction, they will soon surpass human driving competency in almost all circumstances" (McGillis). The vision takes many forms, including a fleet of self-driving cars that communicate with each other.

> "[Vehicle-to-vehicle] communications enable a vehicle to detect threats and hazards with a 360-degree awareness of the position of other vehicles, as well as the threat or hazard those vehicles present; calculate risk;... and take pre-emptive actions to avoid or mitigate crashes" (Narla).

The National Highway Traffic Safety Administration goes even further, estimating that communications between vehicles (and infrastructure) would enable safety applications that "could eliminate or mitigate the severity of up to 80 percent of non-impaired crashes" (Thomas). The belief is that self-driving cars would eliminate the human error of present-day drivers.

Safety beats at the heart of the push for self-driving cars. Why have needless traffic accidents? Why let people die in car accidents? If we can eliminate unnecessary accidents and needless deaths, we *should* do so. Moreover, self-driving cars could be an opportunity to better connect us. Kids could ride in a self-driving car to see their grandparents, and a blind man would no longer rely on the bus to get to the grocery store. We could have a safer, freer society. Who wouldn't want that?

The problem is we will only get there if we do self-driving cars *right*. History shows promises of progress can quickly sour when taken at face value– look only at how the Emancipation Proclamation and Reconstruction turned into Jim Crow and segregation within twenty years. If we want to do self-driving cars right and make that safer freer society, we need to take a hard, long look at where we are and how to get there.

To start, where is self-driving technology at the moment? In 2013, David Friggard said, "The technology works today" (Narla). Tesla also sells cars with a feature called Autopilot– which still requires human drivers to be ready to take over. Truthfully, we do not have fully automated cars outside of ads and demo tracks. We don't even have an agreement on models for artificial intelligence (AI) in self-driving cars, nor have we managed to demonstrate that self-driving cars *are* safer than human drivers on an equal playing field.

One group of engineers found that "the existing self-driving strategies focus too much on the 'correctness,' and, to some extent, overlook the human personality and social intelligence" (Li). Learning from a crash with a self-driving car and municipal bus, these engineers sought to make an algorithm that makes human-like decisions. Humans constantly make speculative decisions "from determining whether the car behind is going to yield to avoiding a driver who seems drunk or tired" (Li). This sort of AI seeks to solve issues in a mixed-market situation, where a self-driving car is navigating alongside human unpredictability. Other self-driving cars are more about calculations and theory where everyone is a perfectly correct driver.

This need to handle the unpredictability of driving also calls into question how safe self-driving cars are. Many companies offer statistics on how many miles their self-driving cars have driven without incident. However, one researcher writes that

> "more data and more miles produce better systems, but they also reveal more 'edge cases'-- circumstances that the model cannot account for. Engineers recount the unusual things– balloons, ducks, wheelchairs, kangaroos– that their sensors have seen but which their software has struggled to make sense of. (Stilgoe)

Humans can figure out to stop when an unidentified object bounces or runs into the road, but an algorithm needs to be told what to do and an AI needs to be taught what to do. While the humanlike AI did outperform human drivers in simulations (where it did not need to 'perceive' the road conditions), what will it do with an unidentified object? Are we planning on having self-driving cars stop for all unidentified objects? What effect will that have? Since driving– even without human drivers– is unpredictable, we must *guarantee* our prospective self-driving cars are ready to handle unpredictability.

Interestingly, one researcher tried talking to the general public. He writes "During the public dialogue exercise, participants were quick to highlight complexities and identify what engineers would call 'edge cases'" (Stilgoe). In other words, he found engineers were the experts in designing self-driving cars, but the public could bring in their unusual experiences that self-driving cars may need to be prepared for. Talking to the public may prove to be an important step in preparing for the unpredictable.

By no means have developers, engineers, and proponents only been imagining perfect self-driving cars. The people at MIT Technology Review published an article, in 2015, titled "Why Self-Driving Cars Must Be Programmed to Kill." In it, they argue that a self-driving car programmed to sacrifice its occupant in favor of pedestrians would not be commercially viable because people don't want to buy a car that will sacrifice them. A later study did verify this hunch, finding "from the perspective of the passenger/driver,

participants were more likely to choose *stay* (i.e., harm to the pedestrian) in the [self-driving car] versus [being the driver] (52% vs. 23%)" (Gill). This study also finds that the more autonomous a self-driving car gets, the more people are willing to see it as an independent actor and assign blame to it. In simple words, self-driving cars would absolve people of the guilt they feel when prioritizing themselves above others.

This isn't the only murky ethical situation self-driving cars can get into either. The very nature of using an algorithm to decide morality possibly introduces bias. As one professor at Georgia Institute of Technology writes "In such an algorithm, what is included as a category and how it is included can be a source of deeply problematic biases" (JafariNaimi). Young developers could easily agree amongst themselves that the moral algorithm should prioritize protecting the young. Moreover, developers by no means are representative of the population as a whole, and you only need to look at Facial Recognition Technology to see what problems that can cause. One study into Facial Recognition Technology found that "the Western algorithm recognized Caucasian faces more accurately than East Asian faces and the East Asian algorithm recognized East Asian faces more accurately than Caucasian faces" (Phillips). If developers are simply made up of young people, they may address the concerns of the younger generations without taking stock of the concerns of all stakeholders.

In short, self-driving cars are not ready yet. The technology is new and proprietary (see: hidden from the public), and there are far too many questions about what self-driving cars would look like and what they need to account for. Are all cars going to be self-driving? Are they going to talk to the infrastructure? Do they need to be driving 'correctly', or do they need to be able to predict other drivers? Moreover, we need to understand how the

algorithms work to test coverage. If an accident is unavoidable, who should a self-driving car prioritize? Leaving that question open to the market would simply mean prioritizing the rich. Finally, we need to see that self-driving cars are, in *fact*, safer.

If we want to make a freer, safer society, we need to make self-driving cars *right*. We need to talk amongst ourselves, sharing our problems and concerns with the realities of developing this technology. We need to think about the challenges self-driving cars would face and how to solve them. To make a freer, safer society, we need to bring *everyone* to the decision-making process.

Works Cited

Emerging Technology from the arXiv. “Why Self-Driving Cars Must Be Programmed to Kill.” MIT Technology Review, 22 October 2015.

Gill, Tripat. “Blame It on the Self-Driving Car: How Autonomous Vehicles Can Alter Consumer Morality.” Journal of Consumer Research, vol. 47, 2020.

JafariNaimi, Nassim. “Our Bodies in the Trolley’s Path, or Why Self-driving Cars Must *Not* Be Programmed to Kill.” Science, Technology, & Human Values, vol. 43, no. 2, March 2018.

Li, Liangzhi, et al. “Humanlike Driving: Empirical Decision-Making System for Autonomous Vehicles.” IEEE Transactions on Vehicular Technology, vol. 67, no. 8, August 2018.

McGillis, Jordan. “Autonomous Now: Why We Need Self-Driving Technology and How We Can Get It Faster.” Manhattan Institute, 27 July 2023.

Narla, Siva R. K. “The Evolution of Connected Vehicle Technology: From Smart Drivers to Smart Cars to… Self-Driving Cars.” ITE Journal, Jul 2013.

National Highway Traffic Safety Administration (NHTSA). “Traffic Safety Facts 2022 Data.” U.S. Department of Transportation, September 2024.

Phillips, Jonathan P., et al. “An other-race effect for face recognition algorithms.” ACM Transactions on Applied Perception, vol. 8, issue 2, 2 February 2011.

Stilgoe, Jack. “How can we know a self-driving car is safe?” Ethics and Information Technology, 30 June 2021.

Thomas, Bryan. “U.S. DOT Advances Deployment Of Connected Vehicle Technology To Prevent Hundreds Of Thousands Of Crashes.” US Department of Transportation, 13 December 2016.