Computer Ethics - Philosophical Enquiry (CEPE) Proceedings

Volume 2019 *CEPE 2019: Risk & Cybersecurity* Article 6

5-29-2019

Autonomous Vehicles and the Ethical Tension Between Occupant and Non-Occupant Safety

Jason Borenstein
Georgia Institute of Technology

Joseph Herkert
North Carolina State University

Keith W. Miller
University of Missouri - St. Louis

Autonomous vehicles and the ethical tension between occupant and non-occupant safety

Jason Borenstein
Georgia Institute of Technology

Joseph Herkert
North Carolina State University

Keith W. Miller
University of Missouri - St. Louis

Abstract

Autonomous vehicle manufacturers, people inside an autonomous vehicle (occupants), and people outside the vehicle (non-occupants) are among the distinct stakeholders when addressing ethical issues inherent in systems that include autonomous vehicles. As responses to recent tragic cases illustrate, advocates for autonomous vehicles tend to focus on occupant safety, sometimes to the exclusion of non-occupant safety. Thus, we aim to examine ethical issues associated with non-occupant safety, including pedestrians, bicyclists, motorcyclists, and riders of motorized scooters. We also explore the ethical implications of technical and policy ideas that some might propose to improve non-occupant safety. In addition, if safety (writ large) is truly the paramount priority for autonomous vehicle advocates, we contend that autonomous public transportation should be considered as a more effective and less expensive way to improve public safety.

Keywords: Autonomous vehicles, ethics, occupants, non-occupants

On the evening of 18 March 2019, Elaine Herzberg (age 49) was walking her bike across Mill Avenue in Tempe, Arizona. An autonomous vehicle owned by Uber hit her, and she died of her injuries (Randsazzo 2019). At the time of the collision, there was an Uber employee at the controls of the car. According to the National Transportation Safety Board's preliminary report on the accident, "1.3 seconds before impact, the self-driving system determined that an emergency braking maneuver was needed to mitigate a collision" (NTSB 2018). Yet neither the car's autonomous controller nor the driver slowed the car before it hit Ms. Herzberg. The NTSB (2018) reports states that "According to Uber, emergency braking maneuvers are not enabled while the vehicle is under computer control, to reduce the potential for erratic vehicle behavior. The vehicle operator is relied on to intervene and take action. The system is not designed to alert the operator."

At the time of this writing, the legal dimensions of the case are still being argued in court with initial indications that Uber will not be held criminally liable (for example, see Shepardson and Somerville 2019). Although legal and ethical issues are often

interrelated, we contend they can and should be considered separately in the case of autonomous vehicles.

The Tempe Uber case is an illustrative example of a broad topic: ethical issues linking the developers of autonomous vehicles, people inside an autonomous vehicle, and people outside the vehicle. People outside a particular autonomous vehicle can include people in other vehicles (both autonomous and manual), pedestrians, bicyclists, skateboarders, and many others. In this paper, we primarily explore key ethical issues pertaining to non-occupant safety.

Problems with the Trolley Problem

During the past several years, the "trolley problem" as it intersects with autonomous vehicles has captured attention of both academics and the general public. For example, MIT researchers maintain a frequently visited website that allows people to choose between two actions for a vehicle encountering a perilous intersection in a collection of scenarios (Awad et al. 2018).

Public interest in philosophical quandaries can be beneficial in several ways, including by drawing attention to facets of an issue that may have been underdeveloped. But some scholars worry that an over-emphasis on simple-stated, somewhat contrived scenarios can lead to a trivialization of the ethical analysis necessary for more realistic life experiences (for example, see Gold et al. 2014). Along these lines, we have argued elsewhere that one of the shortcomings with the trolley problem is that it might shift attention away from broader, system level issues (Borenstein et al. 2019), and we continue to build on that argument in this paper.

Our Goal

If the creation and deployment of autonomous vehicles is going to continue, it is of paramount importance to explore the ethical responsibilities of designers, manufacturers, operators, and regulators of the technology. In this paper, we specifically focus on the ethical responsibilities surrounding autonomous vehicles that these stakeholders have to uphold the safety of non-occupants.

The themes of responsibility, praise, and blame have a complex philosophical history. We will not review that history here. Instead, we will assume the reader has at least some familiarity with Kantian views of responsibility for actions, which relies on a notion of free will, and with ideas about just distribution of benefits and harms, which in some formulations may be influenced by Utilitarianism. In both cases, we are particularly interested in how this philosophical notions give insights into how to assign moral responsibility for the safety of non-occupants when autonomous vehicles are deployed into a complex, land-based transportation system. For the sake of simplicity, we will restrict our attention to the well-being of human non-occupants, while acknowledging that there are significant ethical considerations related to animals as well (Bendel 2016).

Public Transportation vs. Private Automated Vehicles

One way to examine questions about responsibility for human safety is to examine public statements by people leading efforts to automate vehicles. For example, John Krafcik, the CEO of the self-driving car company Waymo, was quoted as saying, “We'll continue to put our focus on safety…It is the overwhelming, number-one priority for the team at Waymo” (Kilgore 2018). The insistence on safety as a priority is common among proponents of autonomous vehicles.

Sparrow and Howard (2017) point out that if these statements are to be taken seriously, then the deployment of autonomous vehicles is fundamentally an ethical decision. Furthermore, they contend:

As long as driverless vehicles aren’t safer than human drivers, it will be unethical to sell them (Shladover 2016). Once they are safer than human drivers when it comes to risks to 3rd parties, then it should be illegal to drive them: at that point human drivers will be the moral equivalent of drunk robots.

Sparrow and Howard also point out that improved public transit may be a far more cost effective application of artificial intelligence to transportation if safety is really the number one priority.

The decision to pursue the development of privately owned and used autonomous vehicles (including one-family automobiles, taxis, and ride-shares) versus autonomous vehicles used for public transportation (including buses, mini-buses, and light-rail) raises interesting questions. We contend that although public transportation is mentioned in some of the literature about the ethics of autonomous vehicles (for example, see Beiker 2017), it is not sufficiently regarded as a realistic alternative to rapid expansion of private autonomous vehicle ownership (and it probably should be). This is particularly striking when we recall that automated trains have been operating since the 1980’s, recording relatively safe passenger miles for millions of people (Landennois 1983).

If advances in artificial intelligence (AI) do deliver on their promise and radically alter human transportation, perhaps that revolution should include a massive shift from private transportation to public transportation. At least some people, especially in urban areas, might be happy with this outcome (Jain et al. 2014), especially if it could help ease traffic congestion.

A serious discussion of the safety ramifications of a more comprehensive implementation of automated public transportation instead of increased automated private transportation is unlikely to be in the best financial interest of people and companies committed to selling automated cars. That does not mean that the rest of us should not seriously contemplate other transportation alternatives. While not our main focus in this paper, the macro-ethical questions raised by the pursuit of alternative transportation options, including public transit, should not be overlooked amid the enthusiasm for automated private vehicles.

Whose Safety wlll be Prioritized: Occupants or Non-Occupants?

The design decisions that autonomous vehicle companies are making are laden with numerous ethical dimensions, including how autonomous driving systems will prioritize safety. Christoph von Hugo is Mercedes-Benz's manager of driver assistance systems. He is quoted as saying, "If you know you can save at least one person, at least save that one. Save the one in the car … If all you know for sure is that one death can be prevented, then that's your first priority" (Dodgson 2016).

We find that a remarkable statement with respect to the car manufacturer's responsibility for the safety of non-occupants. We note that von Hugo's logic sets up a tension between data about an automated vehicle's passengers and its apparent lack of data about non-occupants. The implication is that this imbalance justifies explicitly prioritizing the safety of passengers over non-occupants. Since passengers are the likely customers of the car manufacturer, and non-occupants are not, this makes a certain amount of economic and marketing sense; however, it is an ethically problematic logic. It is particularly troubling since it could motivate a car manufacturer to continuously improve design features pertaining to the safety of passengers (whose data are more easily collected and analyzed by a vehicle) and to ignore or downplay the safety of non-occupants.

Several engineers and executives of companies building automated vehicles, including AI entrepreneur Andrew Ng, have made public statements about how pedestrians should behave to increase their safety as autonomous vehicles become more common (Norton 2018). Providing advice to pedestrians and other non-occupants interacting with autonomous vehicles, Ng states that "What we tell people is, 'Please be lawful and please be considerate'," (Kahn 2018). The shift of responsibility implicit in these pronouncements has not gone unnoticed. Kahn (2018) writes:

Rodney Brooks, a well-known robotics researcher and an emeritus professor at the Massachusetts Institute of Technology, wrote in a blog post critical of Ng's sentiments that "the great promise of self-driving cars has been that they will eliminate traffic deaths. Now [Ng] is saying that they will eliminate traffic deaths as long as all humans are trained to change their behavior? What just happened?"

The ethical significance of this shifting of responsibility is clear. Surely all of the people who share the road have responsibilities for their own safety and the safety of others. But it would be ethically problematic if the developers of the new technology suggest that pedestrians, not car manufacturers, are primarily responsible for pedestrian safety in situations when pedestrians, automated vehicles, and other non-occupants mingle. One could argue that such thinking is an extension of "blaming the operator (user)" (Holden 2009). In the case of autonomous vehicles, where there may be no human operator or user, this thinking has the effect of transferring the traditional role (and blame) of operator/user to the pedestrian.

How Responsibilities are Likely to Play Out

The Bicycle Problem

In this section, we introduce several specific issues that illustrate how the ethical responsibilities for non-passenger safety are likely to play out as autonomous vehicles become more common. For example, it is anticipated that bicyclists may be at particular

risk from an autonomous vehicle because they move faster than a pedestrian, and are more difficult to detect than other cars (Bonnington 2018). Fairley (2017) quotes Steven Shladover: "Bicycles are probably the most difficult detection problem that autonomous vehicle systems face."

One way to approach the “bicycle problem” would be to invest time and money so that automated vehicles are at least as safe as human drivers with respect to bicycles. This could, for example, involve efforts to improve computer vision and sensors. An alternative approach is to require bicyclists to become more easily recognized by automated vehicles; this could entail requiring bicyclists to carry electronic devices that automated vehicles could use to more effectively locate (and avoid) bicycles (Bonnington 2018). These devices could be incorporated into the bike, or in a helmet or other wearables. The bicycle problem is similar, but not identical, to problems with motorcycles and automated vehicles. Stock (2016) describes the relevant challenges with motorcycles.

Although these two strategies would not necessarily be mutually exclusive, requiring bicyclists to acquire, wear, and maintain a device in order to protect themselves from an automated vehicle can be problematic ethically in part because of the associated shift of responsibility to them. There are also technical complexities that would emerge. For instance, the addition of the device to the complex communications required in a system that will include multiple versions of complicated automated vehicle software and hardware systems will be another strain on an already difficult technical challenge (Borenstein et al. 2019).

Recent Addition: Motorized Scooters

During recent years, electric scooters (e-scooters) for curb-side rental have proliferated in urban areas (Irfan 2018). E-scooters have both advocates and detractors, but assuming the technology will continue to be used, their interaction with self-driving cars is likely to be problematic, often in ways similar to bicycles. Electric scooters are quick, hard to identify from the street, and often (in the writers’ experience) their riders do not strictly follow rules established for either other motorized vehicles or pedestrians. Another wrinkle is that there are plans for making some scooters self-driving (Blain 2018).

The rapid appearance of e-scooters illustrates a difficulty with automated vehicles: they are unlikely to be able to adapt quickly and safely to a new device that appears in a transportation system. Algorithms and implementations of complex control systems for automated vehicles are likely to be challenging, and frequent changes will make them more so. Testing such software in the face of changing conditions and equipment will be daunting (Kalra and Paddock 2016).

A Technical Approach to Increasing the Safety of Non-Occupants

In the cases of pedestrians and bicyclists, one technical idea is to “light up” these non-occupants with equipment that will alert an autonomous vehicle to their presence and location. The equipment could also alert non-occupants to the presence of an autonomous vehicle in the area.

The approach of placing sensors on non-occupants creates both potential opportunities and vulnerabilities. On the plus side, having autonomous vehicles and non-occupants more aware of each other could be advantageous to both. If the overall transportation system is also aware of these stakeholders and their location, then perhaps there can be system-wide adjustments that will increase safety. For example, traffic could be routed (or advised) away from congested areas (where congestion could reference both vehicles and non-occupants). Sensors might reduce the effect of algorithmic bias, which seems to be a serious problem with facial recognition applications. For example, if autonomous vehicles have been programmed in such a way that certain skin tones are more easily detected than others, sensors could perhaps overcome that problem.

But there is cause for worry about the sensor idea. First, in order for the approach to be effective, there would need to be extensive standardization and/or cooperation between vendors. However, it may be difficult to achieve that kind of cooperation across industries and political entities. Second, the sensor idea may shift safety considerations too far in the direction of non-occupants, since it requires humans to adapt to the technology of autonomous vehicles in a way that may seem intrusive, and likely damaging to privacy.

Also, the sensor idea requires significant user compliance, and it is unlikely that universal compliance will be achieved, especially if users have to purchase the sensors. Pedestrians and others may forget to wear the sensor on a particular day or be visiting a different city that has different norms about sensor use (or a whole host of other related problems). If compliance is spotty, this may increase rather than decrease safety risks for some non-occupants since autonomous vehicle designers may heavily depend on the presence of the sensors. Moreover, sensor malfunctions will eventually occur, and malicious actors might disrupt sensors for mischief or personal gain. Furthermore, the sensor approach might give non-occupants a false sense of security if, for example, an autonomous vehicle does not actually have sufficient time to stop even though a person outside the vehicle has been detected.

Concerns with Technical Approaches

Technical solutions to perceived challenges do not always take into consideration ethical issues that are inherent in the solutions. We have already discussed the example of requiring non-occupants to wear devices that could facilitate more efficient and effective identification by autonomous vehicles. Yet, as previously mentioned, this would shift responsibilities that perhaps should be placed on autonomous vehicles companies and occupants to non-occupants.

Another example of a proposed technical fix is proved by Lee (2018) who suggests that “human intuition” should be built into automated vehicle software. However, it is not clear that a deep understanding of how human intuition works is currently available, and even less clear that artificial intelligence could be made reliable and safe if it attempts to be “intuitive” (whatever that means for a computer program). As Sanctuary (2017) asks, “is reliable artificial intelligence possible?” If not, this technical approach will not pass ethical muster.

A third suggestion for a technical fix is to require autonomous vehicles to make “distinctive sounds” (Norton 2018). First of all, many technical problems could arise with

this “fix.” In a crowded situation, the added noise of hundreds of automated vehicles would probably yield more confusion than increased safety. Furthermore, would non-occupants be required to respond thereby indicating that the sound was heard and identified? This would again shift responsibility from the vehicles to non-occupants. Habibovic et al. (2018) describe a similar idea of having autonomous vehicles communicate their intent to pedestrians, but analogous problems might plague that approach.

Regulatory and Policy Approaches

Regulations and policies could perhaps mitigate non-occupant safety concerns. One policy idea is to require the developers of autonomous vehicles to demonstrate that non-occupant safety will be increased by the use of autonomous vehicles. Yet we are not aware of any legislation that is even being proposed which specifies the nature of that demonstration, and who will be the final judge of whether or not that requirement is met.

Another approach is to designate lanes and perhaps entire routes that will be either all-autonomous vehicles, or all-non-autonomous vehicles. This would help non-occupants as well as occupants, since presumably the non-occupants would be able to better predict the behavior of vehicles when the vehicles are separated. Yet the associated infrastructure investments might make such separation difficult to achieve in practice.

An additional regulatory proposal to consider is adjusting laws in order to lessen the likelihood and magnitude of harm when autonomous vehicles, non-autonomous vehicles, and non-occupants are in close proximity to one another. For example, speed limits might be lowered in any area where the interactions are likely to be frequent. However, this notion is not problem-free either; for instance, it may increase safety, but it would also likely impede traffic flow.

Questions for Future Research

After embracing a systems level view of a transportation system, it becomes clear that many different stakeholders need to be taken into account (not just the occupants of autonomous vehicles) and that many ethical questions need resolution (not just the ones raised in this paper). Here are several examples of such questions:

1. When autonomous vehicles are deployed, who is primarily responsible for the safety of non-occupants?
2. Do drivers of bicycles, motorcycles, and e-scooters that share (or will share) the road with automated vehicles bear a greater responsibility for their own safety than do pedestrians?
3. Will autonomous vehicle ride-sharing services, which place a premium on route efficiency, potentially create an unsafe environment for non-occupants? (For

example, will ride-sharing services take into account the volume of bicycle traffic when scheduling routes?)
4. How much responsibility do companies have to re-evaluate and update autonomous vehicle operating systems with the introduction of new technologies used by non-occupants?
5. How much flexibility should autonomous vehicles have to bend or violate traffic laws if it may help preserve the safety of non-occupants (for example, going over the speed limit to avoid a bicyclist)?
6. Does shifting the unit of analysis from the individual autonomous vehicle to systems of autonomous vehicles (Borenstein et al. 2019) render solutions to such concerns more tractable?
7. To what extent should non-occupants be required to adapt to autonomous vehicles? For example, should non-occupants be required to carry or wear equipment that simplifies their detection?
8. To what extent should traffic laws be changed to safely accommodate interactions between automated vehicles and non-occupants?

Assuming that the momentum towards integrating autonomous vehicles into various transportation systems continues, we suggest that the exploration and analysis of research questions, like the ones above, will need to occur.

Conclusions

One takeaway from the view we articulated here is that the public transportation alternative to private autonomous vehicles has not, in our opinion, been sufficiently considered. The list of specific problems related to autonomous vehicle technology, including technical and ethical, is daunting (and we only highlighted a subset of them here). Yet unless a fatal crash occurs, these problems do not normally receive much attention. And even then, as in the Tempe case, the initial inclination is rarely to call into question the technology. In the Tempe case, the victim was blamed, then the driver was blamed, and only later was the technology called into question.

We contend that before widespread autonomous vehicles become routine, the entire enterprise should receive serious ethical analysis and criticism taking into account safety, equity, and cost effectiveness. In many cases, it is not immediately obvious who should be responsible for safety and security concerns. This has both legal and ethical ramifications.

In some sense, a large-scale autonomous vehicle experiment is occurring on public roads without anything close to informed consent from the relevant cities' citizens, including non-occupants who use roads, bike paths, and sidewalks. Add to this that at best, there is mixed public acceptance of the technology (e.g., Liernert and Caspani 2019). The use of autonomous vehicles is a case in point of a life-altering technology being introduced into society without sufficient opportunity for public input. Yet we voice the hope that it's not too late to ensure that non-occupant safety is prioritized during the process of developing and deploying autonomous vehicles.

References

Awad, E., Dsouza,S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.-F., Rahwan, I. (2018). *The Moral Machine Experiment*. *Nature*.

Sven, B. (2016) Deployment scenarios for vehicles with higher-order automation. *Autonomous Driving*. Springer, Berlin, Heidelberg, 2016. 193-211.

Bendel, O. (2016). Considerations about the relationship between animal and machine ethics. *AI & society*, *31*(1), 103-108.

Loz, B. (2018). Self-driving scooter to save riders from self-driving cars. *New Atlas*, May 29.

Christina. B. (2018). The cyclist problem. *Slate*, February 3.

Borenstein, J., Herkert, J. and Miller,K. (2019). Self-driving cars and engineering ethics: The need for a system level analysis. *Science and Engineering Ethics*, 25(2), 383-398.

Dodgson, L. (2016). Why Mercedes plans to let its self-driving cars kill pedestrians in dicey situations. Business Insider, October 12. https://www.businessinsider.com/mercedes-benz-self-driving-cars-programmed-save-driver-2016-10

Fairley, P. (2017). Self-driving cars have a bicycle problem [News]. *IEEE spectrum*, *54*(3), 12-13.

Gold, N., Pulford, B. D., & Colman, A. M. (2014). The outlandish, the realistic, and the real: contextual manipulation and agent role effects in trolley problems. *Frontiers in psychology*, *5*, 35.

Habibovic, A., Lundgren, V.M., Andersson J., et al. (2018). Communicating intent of automated vehicles to pedestrians. *Frontiers in Psychology*. 2018; 9:1336. doi:10.3389/fpsyg.2018.01336.

Holden, R. J. (2009). People or systems? To blame is human. The fix is to engineer. Professional Safety, 54(12), 34.

Irfan, U. (2018) Electric scooters' sudden invasion of American cities, explained. *Vox.* https://www.vox.com/2018/8/27/17676670/electric-scooter-rental-bird-lime-skip-spin-cities

Jain, S., Aggarwal, P., Kumar, P., Singhal, S., & Sharma, P. (2014). Identifying public preferences using multi-criteria decision making for assessing the shift of urban

commuters from private to public transport: A case study of Delhi. *Transportation Research Part F: Traffic Psychology and Behaviour*, *24*, 60-70.

Johnson, L.and Fitzsimmons, M.. (2018). Uber self-driving cars: everything you need to know. *Techradar*, May 25.

Kahn, J. (2018). To get ready for robot driving, some want to reprogram pedestrians. *Bloomberg News,* August 16, https://www.bloomberg.com/news/articles/2018-08-16/to-get-ready-for-robot-driving-some-want-to-reprogram-pedestrians, accessed 6 July 2019.

Kalra, N. and Paddock, S.M. (2016) Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability? *Transportation Research Part A: Policy and Practice, 94:* 182-193.

Kilgore, T.. (2018). Waymo CEO says safety is 'overwhelming' priority for its self-driving cars. MarketWatch, March 27.

Lardennois, R. (1993). VAL automated guided transit characteristics and evolutions. *Journal of advanced transportation*, *27*(1), 103-120.

Liernert, P. and Caspani, M. (2019). Americans still don't trust self-driving cars, Reuters/Ipsos poll finds. *Reuters*, April 1.

National Transportation Safety Board (NTSB). (2018). Preliminary Report Highway: HWY18MH010. May 24. https://www.ntsb.gov/investigations/AccidentReports/Pages/HWY18MH010-prelim.aspx.

Norton, P. (2018). Self-driving car developers should put pedestrians first. *Wired*, September 15.

Randazzo, R. (2019). Family of woman killed in crash with self-driving Uber sues Arizona, Tempe. *Arizona Republican*, March 19, https://www.azcentral.com/story/news/local/tempe/2019/03/19/arizona-city-tempe-sued-family-uber-self-driving-car-crash-victim-elaine-herzberg/3207598002/, accessed 6 July 2019.

Sanctuary, H. (2017). Is reliable artificial intelligence possible? *PHYS.ORG.* Retrieved from https://phys.org/news/2017-03-reliable-artificial-intelligence.html.

Shladover, S. (2016) The truth about “self-driving” cars. *Scientific American,* 314 (6), 52-57

Shepardson D.and Somerville, H. (2019). Uber not criminally liable in fatal 2018 Arizona self-driving crash: prosecutors. *Reuters,* March 5.

Sparrow, R. J., & Howard, M. (2017). When human beings are like drunk robots: Driverless vehicles, ethics, and the future of transport. *Transportation Research Part C: Emerging Technologies*, *80*, 206–215. https://doi.org/10.1016/j.trc.2017.04.014

Stock, K. (2016). Self-Driving cars will be the best thing to happen to motorcycles. *Bloomberg*, October 11.

Taylor, M. (2016). Self-Driving Mercedes-Benzes will prioritize occupant safety over pedestrians. *Car and Driver*, October 7.