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

Jason Borenstein  
Georgia Institute of Technology, borenstein@gatech.edu

Joseph Herkert  
North Carolina State University, jherkert@ncsu.edu

Keith Miller  
University of Missouri · St. Louis, millerkei@umsl.edu

Follow this and additional works at: https://digitalcommons.odu.edu/sociotechnicalcritique

Part of the Applied Ethics Commons, Artificial Intelligence and Robotics Commons, and the Science and Technology Studies Commons

Recommended Citation

This Research Article is brought to you for free and open access by ODU Digital Commons. It has been accepted for inclusion in The Journal of Sociotechnical Critique by an authorized editor of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.
Autonomous vehicles and the ethical tension between occupant and non-occupant safety

Cover Page Footnote
A version of this paper was presented at Computer Ethics—Philosophical Enquiry (CEPE) 2019, at Old Dominion University in Norfolk, Virginia, in the United States of America. It was selected to be included in a special issue devoted to papers from the conference, contingent on successful normal blind peer-review processes. Due to disruptions in article revision processes caused by the COVID-19 pandemic, and the very different ways these disruptions have affected different authors, JStC has chosen to publish articles from this special issue as they are ready, rather than waiting to publish them all together. Our thanks to the anonymous reviewers for their constructive criticisms.
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*

Given that the creation and deployment of autonomous vehicles is likely to continue, it is important to explore the ethical responsibilities of designers, manufacturers, operators, and regulators of the technology. We specifically focus on the ethical responsibilities surrounding autonomous vehicles that these stakeholders have to protect the safety of non-occupants, meaning individuals who are around the vehicles while they are operating. The term "non-occupants" includes, but is not limited to, pedestrians and cyclists. We are particularly interested in how to assign moral responsibility for the safety of non-occupants when autonomous vehicles are deployed in a complex, land-based transportation system.

*Keywords: autonomous vehicles, occupants, non-occupants, safety, ethics*

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 due to the injuries caused by the accident (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 resolved with initial indications that Uber will not be held criminally liable (for example, see Shepardson & 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 broader 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 and the potentially problematic shift of responsibility for safety of autonomous vehicles to non-occupants.

**Shortcomings of applying the trolley problem to autonomous vehicles**

During the past several years, the “trolley problem” as it relates to autonomous vehicles has frequently captured the 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; findings from the MIT study purportedly reveal different perceptions of safety priorities in different regions of the globe (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 simply-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 such as how different types of autonomous technology might interact with one another (Borenstein et al., 2019). We continue to build on that argument in this paper.

**Our goal**

Given that the creation and deployment of autonomous vehicles is likely to continue, it is of paramount importance to explore the ethical responsibilities of designers, manufacturers, operators, and regulators of the technology. Here, we specifically focus on the ethical responsibilities surrounding autonomous vehicles that these stakeholders have to protect the safety of non-occupants, meaning individuals who are around the
vehicles while they are operating. The term “non-occupants” includes, but is not limited to, pedestrians and cyclists. We are particularly interested in how to assign moral responsibility for the safety of non-occupants when autonomous vehicles are deployed in 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).

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, has been 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. Obviously, this should be a priority, but it raises questions about whose safety is being prioritized.

Sparrow & Howard (2017) point out that if statements by companies and others 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 & 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.

**Prioritizing private automated vehicles over public transportation**

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, ethical and otherwise. 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 by the transportation industry as a realistic alternative to rapid expansion of private autonomous vehicle ownership (although it probably should be). This is particularly striking when we recall that automated trains have been
operating since the 1980s, recording relatively safe passenger miles for millions of people (Lardennois, 1983).

A key reason for mentioning the issue of mass transit versus private transportation is that the latter adds much more complexity in terms of protecting non-occupants in many ways, including that it is harder to implement centralized control over the vehicles and presumably the sheer number of vehicles on the road will be greater. 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. While safety is our focus here, we mention in passing that other issues are impacted by the decision to prioritize public versus private transportation, including equitable access to transportation options, ecological impact of emissions, and the number of vehicles required per capita.

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 will be prioritized: Occupants or non-occupants?

The design decisions that autonomous vehicle companies make 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 draw specific attention to Mr. von Hugo’s statement because he is in a position of power to shape the design aims of Mercedes-Benz’s autonomous vehicle program.

We find that a remarkable statement with respect to what it implies regarding 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 intentionally 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. For a more extended discussion of other opinions about the issue of tradeoffs between occupants and non-occupants, including a different analysis of the Hugo quote, see (Keeling et al., 2019).

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 automated vehicles mingle with pedestrians and other non-occupants. One could argue that such thinking is an extension of “blaming the operator (user)” (Holden, 2009). In the case of autonomous vehicles, where a human operator or user might be absent, this thinking has the effect of transferring the traditional role (and blame) of the operator or 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-occupant 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 may be more difficult for sensors to detect and identify than other cars (Bonnington, 2018). Fairley (2017) quotes Steven Shladover: “Bicycles are probably the most difficult detection problem that autonomous vehicle systems face.” Not only will autonomous vehicles need to detect that an object, a bicycle, is within its proximity, but will also have to decipher what the object is, or else a human being, the bicyclist, may be put at significant risk. False positives (something not a bicycle is identified as a bicycle) and false negatives (something that is a bicycle is not recognized as such) both have potentially significant safety risks for non-occupants.

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 some of 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 ethically problematic in part because of the associated shift of responsibility to bicyclists. 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 use of the technology will continue, 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 plans for making some scooters self-driving might be underway (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 & Paddock, 2016). Furthermore, is the expectation that the autonomous vehicles will have to be recalibrated and adjust to the new technology, in this case e-scooters, or will the new technology have to be designed in a way so that it is compatible with autonomous vehicles?

A technical approach to increasing the safety of non-occupants

In the cases of pedestrians and bicyclists, one technical method 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 is a serious problem with facial recognition applications (e.g., Benjamin, 2018). For example, if autonomous vehicles have been programmed in such a way that certain skin tones are more easily detected than others, sensors worn by bicyclists could perhaps help to correct for that problem.

Yet there is cause for worry about implementing the sensor idea. First, for the approach to be effective, there would need to be extensive standardization and/or cooperation between vendors. Achieving that kind of cooperation across industries and political entities, however, will be difficult. Second, the sensor idea may shift considerations of safety (and the associated responsibilities for it) 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 adherence 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 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 companies and vehicle occupants to non-occupants.

Another example of a proposed technical fix is provided 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 it is 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. Moreover, what happens if a person has a hearing impairment or disability? Assuming that the vehicle’s sounds can be heard, 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 at least some of the concerns surrounding non-occupant safety. 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 as compared to traditional human-operated vehicles. Yet we are not aware of any legislation that is 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 to have more clearly defined and divided spaces, 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 exceedingly 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 in some areas, but it would also likely impede traffic flow and vehicles may just seek alternative routes (which may just move safety-related problems to other areas of the city).

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 should be primarily responsible for the safety of non-occupants?

2. Should drivers of bicycles, motorcycles, and e-scooters that 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? Should software updates be scrutinized by an external agency to monitor safety?

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)? Before they are deployed, will such exceptions to local traffic laws be approved by the legislative or regulatory entities? Also, if an autonomous vehicle receives a ticket for a road violation while under software control, who pays the fine?

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? Who should pay for such equipment? Should the equipment be regulated to ensure safety? If so, who should enforce the regulations?

8. To what extent should traffic laws be changed to safely accommodate interactions between automated vehicles and non-occupants? Should such laws be limited to local, state, or national boundaries? If the laws differ in different jurisdictions, how should conflicts between those laws be integrated when changing jurisdictions? (Some cities straddle jurisdictions, and confusion could be severe both for humans and for autonomous systems.)

9. If autonomous vehicles are involved in incidents that, through no fault of their own, injure or kill non-occupants, or traffic violations that potentially cause such harm, who is ultimately responsible – vehicle manufacturers, owners, and/or 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, need to occur before deployment, not after.

Conclusions

One takeaway from the view we articulated here is that the public transportation alternative to private autonomous vehicles has not, in our opinion, received sufficient consideration. 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. Crashes may be blamed on the technology (Goodall, 2018), and sometimes humans (Brooks, 2017). In the Tempe case, over time, the victim was blamed, the driver was blamed, Uber was blamed, and the technology was blamed.

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, public acceptance of the technology is mixed (e.g., Liernert & 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 a paramount priority during the process of developing and deploying autonomous vehicles.

References


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


