More Than Trolleys:

Plausible, Ethically Ambiguous Scenarios Likely to Be Encountered by Automated Vehicles

Noah J. Goodall

This is a pre-print version of an article published in Transfers: Interdisciplinary Journal of Mobility Studies. The definitive publisher-authenticated version [Noah J. Goodall, "More than Trolleys: Plausible, Ethically Ambiguous Scenarios likely to Be Encountered by Automated Vehicles," *Transfers: Interdisciplinary Journal of Mobility Studies* 9, no. 2 (2019): 45–58] is available online at: https://doi.org/10.3167/TRANS.2019.090204.

Abstract: As the act of driving becomes increasingly automated, vehicles will encounter situations where different objectives of safety, mobility, and legality will come into conflict. These situations require a vehicle to compare relative values of different entities and objectives, where the action of the vehicle has a moral component. While discussion of these scenarios often focuses on the "trolley problem" thought experiment, these types of life-or-death moral dilemmas may be rare in practice. This article identifies four far more common examples of routine driving that require decisions with some level of ethical reasoning about how to distribute risk. These scenarios may be useful for automated vehicle developers in assessing vehicle safety and responding to potential future regulations, as well as for regulators in developing performance requirements.

Keywords: autonomous vehicles, driverless vehicles, ethics of mobility, risk, vehicle automation, welfare

The act of driving has become increasingly automated. Several vehicles on the market today are able to stay within a lane and follow a lead vehicle ahead under certain conditions.[1] Most major automakers and several technology companies have invested in research aimed at developing a car that can automatically control throttle, brake, and steering across a wide range of roadways and conditions using data from various on-board sensors and highly detailed maps of the environment.[2] Simultaneously, governments and industry are investing in smart infrastructure such as wireless communication networks to support the automation of driving.[3] It is difficult to overstate the impact that reliable, capable, computer-driven vehicles will have on the economy, public health, and even social identities.[4] Google first publicly acknowledged its efforts to develop self-driving technology, including testing in live traffic, in October 2010.[5] Twenty months later, *The Economist* asked what might happen when a vehicle encounters a dilemma in which there was no safe outcome: "Should a driverless car swerve to avoid pedestrians if that means hitting other vehicles or endangering its occupants?"[6] This appears to be the first published acknowledgment of an ethical challenge in programming automated vehicle (AV) behavior. The ethics of artificial intelligence and robotics had been written about for far longer, as applied to medical assistant robots, autonomous weapons systems, and advanced AI.[7]

While these fields use a range of example scenarios to illustrate the needs and challenges for ethicists, early media articles and scientific literature on vehicle automation often employed variations of essentially the same problem: an automated vehicle is driving along and must make a choice between hitting one of two objects, one of which may or may not injure the automated vehicle's passenger. The objects vary: a school bus with forty children or a bridge railing, a public bus or a bridge railing, a cat/deer or a child, a motorcyclist with a helmet or one without a helmet, an eight-year-old girl or an eighty-year-old grandmother, ten pedestrians in the road or one pedestrian on the sidewalk.[8] These hypothetical situations seem to receive more attention in popular media than their likelihood of occurrence would justify. If automated vehicle developers and regulators consider trolley problems as the only scenarios in which automated vehicles must make decisions with moral consequences, they may miss many far more common and mundane scenarios with subtle yet significant ethical implications. This article reviews the trolley problem as it is applied to automated vehicles, the reasons it is useful in considering the ethics of driving, and industry criticisms. Four mundane yet morally significant driving situations are presented as far more common and pressing scenarios requiring ethical decision-making for automated vehicles.

The Trolley Problem

These discrete choice situations are often referred to as variations of the trolley problem, Philippa Foot's thought experiment, later expanded by Judith Jarvis Thompson.[9] In the original version, a driver of a runaway tram can steer toward either a track with one worker or a track with five workers, with anyone struck by the tram facing certain death. The similarities between the original trolley problem and the automated vehicle scenarios seem coincidental. One purpose of the trolley problem is to demonstrate the doctrine of double effect, that is, the idea that it seems morally acceptable for the one worker to die if his death was not required to save the five, but is merely an unintended consequence. If somehow the scenario required the worker to die in order to save the five (maybe he is pushed onto the track and his body derails the train and saves the five workers), this would be immoral because his death is required to save the five. What is clear is that Foot and Thompson were interested not in improving rail safety but rather in the foundational principles guiding our intuitions about right and wrong.

Early media coverage of automated vehicle ethics focused on trolley problem–influenced forced-choice scenarios because they represent such stark examples of ambiguous decisions with obvious moral components. It is very difficult to compare the value of a child in the road, an adult on the sidewalk, and a person in a car. Each has contributed to the incident in different ways (and in some cases, not at all). The child may not be legally responsible for her actions, and perhaps deserves special protection. The automated vehicle's sensors may be highly confident that a person is traveling inside the car but slightly less confident that the pedestrian is not actually a tree or debris or glare. Should it matter? In forced-choice scenarios, these decisions matter because the outcome is often a death sentence. Forced-choice scenarios, if you think they will occur with some frequency, are distinguishable yet clear choices with important consequences, and therefore of most interest to a general audience. Researchers have their own reasons for using forced-choice scenarios. In trying to answer some of these fundamental questions about how to compare values of different objects, it can be helpful to ask people what they would do in similar situations. If 51 percent agree it is permissible to swerve into one thousand dogs to save a human pedestrian, but only 49 percent agree it is permissible with two thousand dogs, then we are starting to get some kind of idea of the how people compare the value of dog and human life. Information like this is useful when deciding how hard an automated vehicle should brake to avoid a dog running into the road, given that sudden braking increases the probability of a rear-end collision. Introducing ambiguity into the hypothetical scenario only encourages participants to invent ways to avoid the crash altogether, which does not support discovering relative values of objects.[10]

Industry Responses to the Trolley Problem

Engineers working on vehicle automation are often asked about the trolley problem. The most common response seems to be that trolley problems are avoidable, implausible, rare, and distractions from more productive efforts.[11] They are considered avoidable because in many trolley problems, the automated vehicle must decide how best to crash when, with the right sensors and algorithms, the situation should have been avoided entirely.[12] An advanced automated vehicle would have slowed down before that blind turn, seen that animal before it leaped into the road, or known this neighborhood has young children and adjusted its speed accordingly. Developers find trolley problems implausible and rare for other reasons. Most people have trouble remembering a situation in which they had time to decide which way they should crash. Because of how these forced-choice scenarios are presented in the media and literature, they are easy to mock by focusing on some of the more outlandish examples, like a car colliding with a criminal instead of (specifically and consistently) a nun.[13] Focusing resources on unlikely edge cases seems like a waste of resources that could be better spent on general collision avoidance. These concerns are not unreasonable, and many have been responded to in the literature.[14] The criticisms fail to address the issue of an oncoming vehicle crossing a double yellow line and into an automated vehicle's path; this scenario is impossible to foresee, difficult to avoid safely, and fairly common. With other cars nearby, deciding how to minimize crash damage is difficult, and can be ethically complex.

General Automated Vehicle Ethics

In September 2016, the National Highway Traffic Safety Administration (NHTSA) released its automated vehicle policy.[15] As the federal agency responsible for writing and enforcing the US Federal Motor Vehicle Safety Standards, NHTSA has the authority to prohibit or mandate certain technologies in vehicles sold in the United States. In the policy document, NHTSA requests AV developers describe how they plan to address fifteen areas of a safety assessment, one of which is Ethical Considerations.[16] NHTSA describes two types of situations an automated vehicle may encounter where ethics might be relevant:

1. When balancing conflicting objectives of safety, mobility, and legality: "Most States have a law prohibiting motor vehicles from crossing a double yellow line in the center of a roadway. When another vehicle on a two-lane road is double-parked or otherwise blocking a vehicle's travel lane, the mobility objective (to move forward toward an intended destination) may come into conflict with safety and legality objectives (e.g., avoiding risk of crash with oncoming car and obeying a law). An HAV [highly automated vehicle] confronted with this conflict could resolve it in a few different ways, depending on the decision rules it has been programmed to apply, or even settings applied by a human driver or occupant."
2. When balancing the safety needs of one or more people: "It may be that the safety of one person may be protected only at the cost of the safety of another person. In such a dilemma situation, the programming of the HAV will have a significant influence over the outcome for each individual involved."[17]

The scenarios described are much broader than the trolley, and far more common. A full defense of the relevance of ethics in automated vehicle design is beyond the scope of this article but can be found in the literature.[18] The objective of this article is to identify examples of decision-making in routine driving that have significant ethical components. These situations can be used as test cases for validating an automated vehicle's programming and as examples for responding to potential future regulations or voluntary safety assessments.

Routine Driving Requiring Ethics

Forced-choice scenarios always result in a crash. Ethics, however, is relevant to a wide range of crash and non-crash scenarios, especially when safety is viewed as a part of a vehicle's risk management system. For example, almost all driving introduces small amounts of risk into the environment. Crashes can happen when equipment fails, when drivers make mistakes or drive poorly, or when unforeseen, low-probability events occur. A vehicle's crash risk for a given situation is a function of the probability of crashing, the severity of the crash, and the uncertainty of its knowledge of the environment. Any vehicle that can cause damage and has a nonzero probability of colliding with another object is introducing risk. An automated vehicle may drive in such a way that it is never responsible for an accident, yet it is still creating risk simply by moving and being a large object capable of inflicting damage.

Generally, ethics becomes relevant either when a vehicle must distribute that risk among road users, some of whom may not have consented to the risk or gain any benefit from the risk, and when competing values such as safety, efficiency, and legality come into conflict. This section describes four scenarios where conflicting objectives or distribution of risk requires a consideration of the ethical impacts of a vehicle's decision. These scenarios are far more common and plausible than forced-choice scenarios. This list is not exhaustive, but merely some of the more immediate issues of significance. Other researchers may identify similar issues, and new scenarios may be discovered as vehicle automation begins to deploy.

Following Distance

When driving in congested traffic, automated vehicles must determine the distance at which to follow the vehicle ahead. This can be expressed as space headway, that is, the distance between the front of the leading vehicle and the front of the following vehicle, or as time headway, the length of time between when the front of the leading vehicle and the front of the following vehicle pass the same location. (An alternative metric is space gap or time gap, which instead measures between the back of the lead vehicle and the front of the following vehicle.) Driver education courses in the United States often recommend a time headway of two to four seconds, although this is largely to account for human reaction time and drivers' inability to perceive the precise rate of deceleration of the lead vehicle.[19] Because computers using distance measurements from radar or lidar sensors can adjust vehicle speed more quickly and precisely than humans, many assume automated vehicles will be able to safely follow a lead vehicle at headways of one second or less. Closer headways are preferred, as they decrease fuel consumption needed to overcome wind resistance and increase a road's traffic capacity.[20] Closer following distances can increase the likelihood of crashing, although not for obvious reasons. In common law, the assured clear distance ahead (ACDA) doctrine specifies a driver can "regulate his speed so that he can stop within the range of his vision."[21] The result is that rear-end collisions are generally the fault of the following vehicle. Exceptions due to unforeseen objects entering the roadway are covered under the sudden emergency doctrine, which excuses a driver from negligence when striking a person, object, or vehicle that unexpectedly or suddenly moves laterally into his path. In the absence of such an exemption, the ACDA suggests, in the event of a rear-end collision, the driver of the following vehicle has the presumption of negligence.

Scott Le Vine and colleagues considered the impact of ACDA doctrine on the safety and efficiency of automated vehicles on freeways. They note the interpretation of the ACDA doctrine will greatly affect minimum allowable headways:

> ACDA generally requires vehicle operators to avoid colliding with both the vehicle ahead (the "weak" ACDA interpretation) as well as stationary objects (the "strong" ACDA interpretation). If the following vehicle strictly observes the ACDA requirement to be able to "*stop within the range of his vision*," the vehicle must avoid striking a possible stationary object in the travel lane (e.g. a moderate-sized piece of road debris) which only becomes "*within the range of his vision*" once the leading vehicle (which blocks the forward line of sight) has passed over the foreign object.

The following vehicle can operate under the assumption that the lead vehicle will decelerate at some percentile of the range of known deceleration rates, at the maximum deceleration rate of that particular vehicle, at the maximum deceleration rate of all vehicles, or at an infinite deceleration rate (i.e., the vehicle immediately becomes static). Likewise, the following vehicle may assume the lead vehicle could pass over or quickly avoid some debris that the following vehicle would be unable to avoid. These assumptions drastically change the desired following distance, with significant impacts on safety and capacity. Le Vine et al. demonstrate that automated vehicles operating under the assumption that the lead vehicle can stop instantly or pass over dangerous debris will require a following headway of 2.44 seconds (assuming 0.4 seconds to react). If AV's merely assume the lead vehicle can only stop using braking alone, the following headway drops to 0.69 seconds. These following distances translate into massive changes in capacity for an AV-only facility, from 4,108 to 1,367 vehicles per lane per hour at the conservative end for each assumption.[22] The decision for how to interpret the ACDA is an ethical issue because it requires balancing the competing values of safety and efficiency, where efficiency is expressed as travel time, fuel economy, and capacity. This is the same conflict behind much safety regulation, from building codes (safer buildings are more expensive) to workplace safety, although they might not be expressed using moral language. States may eventually define allowable following distances, or the insurance industry may assign relative costs to safety risk and make the two values more directly comparable. In either case, someone must decide how to explicitly compare safety and efficiency, a task that has clear moral implications.

Braking Strategies

An automated vehicle engineer, when asked about the trolley problem in vehicle automation, noted that hard braking, not swerving, is often the safest option in a dilemma.[23] Yet, a sudden deceleration of a lead vehicle can increase the risk of being struck by the following vehicle for several reasons: the following driver may be inattentive, the following vehicle may have been following too closely, or the following vehicle may have worse braking capabilities than the lead vehicle. For the lead vehicle, deciding whether and how hard to apply the brakes should take into account the risk of and potential damage from being struck by the following vehicle. Consider an automated vehicle on a rural road. If a dog begins running toward the road on a collision course with the AV, the AV will probably brake to avoid striking the dog. Ideally, an automated vehicle will assess the traffic behind when determining how hard to brake. If a large truck is following closely, an AV may brake only slightly or not at all, willing to risk striking the dog to avoid a collision with the truck. Similarly, the calculation changes if, instead of a dog running toward the road, it is a child, a rabbit, or an object that looks like a cardboard box but the AV cannot be certain. These situations happen daily and have clear ethical components, as they require comparing the relative value of two very different outcomes. Both outcomes are probabilistic, but this does not suggest they are incomparable: the field of risk management incorporates value (referred to as magnitude), probability, and uncertainty in deciding what course of action to take.

A less dramatic but more common situation is a vehicle decelerating in traffic, either in response to traffic control like a stop sign or traffic signal, or in response to the vehicle ahead stopping or slowing. This appears to be a common source of crashes for AVs. A study of Google's automated vehicle crashes through October 2015 found eleven crashes while the vehicle was operating under automation.[24] Of these crashes, nine involved the automated vehicle being struck from the rear. In four of these, the automated vehicle had come to a full stop, while the vehicle in the others was still in motion. Automated vehicle developers have long understood that human drivers do not apply steady pressure to the brakes as a computer might but instead build pressure gradually and then back off as the vehicle nears a complete stop.[25] While varying brake pressure improves passenger comfort, changing a vehicle's method of deceleration may also reduce the prevalence and severity of rear-end collisions. Consider three vehicles in a line, all approaching a traffic signal that has just turned red. The middle vehicle, operating autonomously, may decelerate at a rate that is most comfortable to its passenger or, if headway to the leading vehicle is too small, by keeping a minimum headway to the lead vehicle. This ensures passenger comfort without causing a collision.

There are other strategies an automated vehicle may use to decelerate safely. The automated vehicle may instead prioritize collision risk avoidance over passenger comfort. There are several ways to do this:

1. The AV can follow the lead vehicle closely, leaving the largest possible gap for the following vehicle.
2. The AV can act as a pace car for the following vehicle, allowing the following vehicle to close the gap to the AV quickly while maintaining a minimum safety distance. This increases the likelihood that the following driver will notice the AV's deceleration, as drivers perceive their relative velocity to a vehicle ahead from changes in the visual angle of the vehicle's shape.[26] As a driver gets closer to a lead vehicle, the visual angle changes more dramatically, and relative velocity becomes easier to judge. This also lets the AV maintain a large gap to the lead vehicle, in case the AV needs to accelerate to avoid a distracted driver in the following vehicle.
3. The AV can try to split the distance between the lead and following vehicle, balancing distance from the potentially distracted following vehicle while also acting somewhat as a pace car.

Each of these possibilities will have various impacts on safety and comfort, which are themselves ethical issues. More importantly, these different strategies may transfer risk between the lead and following vehicles. For example, the first strategy, where the AV closely follows the lead vehicle, may increase the likelihood that a collision with the following vehicle pushes the AV into the lead vehicle, creating a three-car collision. It is tempting to dismiss these types of problems, as human drivers generally do not monitor traffic behind them and would be unable to make these kinds of calculations in real time. Computer-driven vehicles, with more precise measurements and control, allow these kinds of previously impossible safety enhancements.

Still, when safety improvements are distributed inequitably, the decisions behind them need to be carefully considered and defended.

Lateral Positioning within a Lane

Vehicles can generally drive anywhere within a travel lane, as long as they remain within the lane markings. Lateral position is generally unregulated, with at least one exception in the Commonwealth of Virginia, which requires drivers about to turn right to position their vehicles "as close as practicable to the right curb or edge of the roadway."[27] Moving laterally within a lane is a safe, legal method to avoid potential crashes. Even with vehicles on all sides, one can move to the left to avoid a car that seems to be encroaching from the right, creating some extra time to decelerate and change lanes. This decision transfers risk among travelers without their consent. When encoded in AV software, small things like lateral movements within a lane will be repeated millions of times over trillions of miles, and will eventually affect crash rates in measurable ways. Developers may need to be clear about the circumstances under which a vehicle would move laterally within a lane, and more importantly when it would not. Consider a patent approved in 2014 that describes a method where an automated vehicle may decide to change its lateral position within a lane. The authors give two examples of how this might be applied. In the first, the AV would move away from a motorcyclist and toward the shoulder of the road, giving the motorcyclist more of a buffer. This seems easy to defend on moral grounds, as the AV takes on the risk of running off the shoulder in order to reduce the risk to the motorcyclist. The second example is harder to defend:

> In another example, the vehicle may be laterally adjacent to two objects in a given portion of the road for a given period of time, where a first object of the two objects may be a bus or a truck or any large vehicle moving on one side of the vehicle, and a second object of the two objects may be a small car moving on the other side of the vehicle. In this example, the computing device may be configured to modify the trajectory of the vehicle such that the vehicle has a larger lateral distance with the first object than with the second object. Thus, the modified trajectory may be biased, relative to a center of a lane of the vehicle, towards the small vehicle.[28]

In this example, the AV moves away from the large truck or bus and toward the small car. This transfers crash risk from the truck/bus to the small car while also lessening the risk for the AV, as, all things being equal, crashing into a small car is safer than crashing into a large bus or truck.[29]

Were the AV to be driving empty, a collision with the truck would seem preferable to one with the small car, as the greater mass of the truck would protect the truck driver from injury. Depending on crash severity models, the AV might actually be increasing the total magnitude of a potential crash, that is, the collective injury severity of all parties, by moving closer to the small car.[30] If this is the case, then the decision to prioritize the life of the passenger (even to a small degree) is embedded in the AV's code, whether or not it is explicit or put there with intent. This is not a minor decision: a representative from Mercedes generated significant controversy in 2016 by suggesting AVs should prioritize the lives of their passengers over others, remarks that Mercedes later apologized for and clarified.[31] Controversy aside, Geoff Keeling and colleagues argue it may be morally permissible for an automated vehicle to give a slight priority to its own passengers.[32] Passengers, knowing they are prioritized by the AV, may be more likely to use AVs, hastening deployment of relatively safe automated driving and improving safety for all road users. Automated vehicles also act as agents for their passengers and therefore may have a duty to carry out their preferences for safety, just as an automated weapons system has a duty to its owner. There may be other ways to justify driving closer to the small car: perhaps trucks sway more within a lane, or historically are more likely to encroach into the AV's lane because of the truck's poor blind spot visibility. The developers could also argue the slight risk increase to the small car is offset by the reduced risk involvement in the AV-truck collision. One could even argue the risk increase is so small as to be negligible. The transfer of risk, however, should be acknowledged and justified. Similar situations in routine driving include deciding when to pass a cyclist or pedestrian.

Permitting Violations of the Law

The objectives of driving are safety, mobility, and legality. These objectives often conflict, especially if one strictly interprets legal statutes. An early example of this conflict was speeding. Most roads have posted speed limits, and it is a violation of the law to exceed those speed limits. Enforcement, however, is somewhat discretionary, and the evidence suggests matching the speed of surrounding traffic may actually be safer than adhering to the speed limit.[33] Google has violated the speed limit while testing their automated vehicles, and defended their speeding citing safety concerns.[34] Any violation of the motor vehicle code calls for a justification using an argument on safety, efficiency, or moral grounds.

Conclusion

Vehicle automation continues to advance, facing new challenges beyond crash avoidance. The act of driving involves meeting objectives of safety, mobility, and legality. When those objectives conflict, or when there is conflict within a single objective, the decision process requires the values be weighted by their importance. These values are dependent on specific circumstances, and may never have universal agreement. The ethics of driving goes far beyond the dualistic, contrived nature of the trolley problem. This article identifies several types of situations that are common in routine driving yet still require ethics in the decision process. Examples included in this article are following distance, braking strategies, lateral positioning within a lane, and violating the law. When determining the desired following distance to a leading vehicle, an automated vehicle must consider both the crash risk and the efficiency of traffic, as every increase in headway translates into a (sometimes unsustainable) reduction in throughput. A vehicle braking for an object must consider the optimal rate of deceleration to minimize the probability of being struck from behind, thereby balancing the risk of colliding with an object both in front and in back. This calculation of crash risk involves not only the probability of crashing but also the magnitude or severity of each crash. Because a traffic cone is less valuable than a human, any measure of severity will force an automated vehicle to assign values to objects in a crash.

An automated vehicle may legally position itself laterally within a lane to be closer to a small vehicle and away from a large vehicle, as a collision with the small vehicle would be less severe. In this case, the automated vehicle may assign equal values to the small and large vehicles, but because of its lateral positioning, it is increasing the risk of collision for the small vehicle without the driver's consent. This scenario requires AV developers to consider whether other drivers have a right to consent to small but systemic increases in crash risk. Automated vehicle developers may also have to determine whether an AV is justified in prioritizing its own passenger over others if, perhaps, the small car in the example has four passengers while the large vehicle is driving empty; the risk of colliding with the small car may be lower for the AV passenger, but the cumulative risk for all parties may be higher.

Finally, automated vehicle developers must defend decisions to violate the law in order to preserve safety, as has already occurred with speed limits and may again happen during emergency maneuvers such as crossing pavement markings. In each case, the developers of the vehicle should first acknowledge the ethical dimensions of the decision. They must explain to regulators and potential customers how these seemingly mundane decisions involve complex tradeoffs between safety and efficiency, or require assigning relative values to different objects or user groups, or violate norms of legality or consent. Developers should acknowledge that even small shifts in risk, when multiplied over trillions of miles, can result in measurable differences in crash rates. Finally, developers should be clear about what their vehicles will do in these morally complex situations, and defend their decisions using the language of ethics.

Noah Goodall is a Senior Research Scientist with the Virginia Transportation Research Council. He has a PhD in Civil Engineering from the University of Virginia and is a licensed engineer in the Commonwealth of Virginia. He published the first peer-reviewed article on the ethics of vehicle automation and is a member of the Transportation Research Board Standing Committee on Vehicle-Highway Automation. Email: njg2q@virginia.edu

Notes

This material is based on work supported by the National Science Foundation under grant no. 1522240. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the NSF.

1. Steven Loveday, "The 15 Cheapest Cars with Adaptive Cruise Control," *U.S. News & World Report*, 24 October 2018, https://cars.usnews.com/cars-trucks/cheapest-cars-with-adaptive-cruise-control.

---

2. “40+ Corporations Working On Autonomous Vehicles,” *CB Insights—Blog*, 28 August 2019, https://www.cbinsights.com/blog/autonomous-driverless-vehicles-corporations-list.

3. Kathleen Frazer Oswald, “A Brief History of Smart Transportation Infrastructure,” *Transfers* 6, no. 3 (2016): 123–129, https://doi.org/10.3167/TRANS.2016.060310.

4. Julia M. Hildebrand, “Media Ecologies of Autonomous Automobility: Gendered and Racial Dimensions of Future Concept Cars,” *Transfers* 8, no. 1 (2018), 64–85, https://doi.org/10.3167/TRANS.2018.080106; Dag Balkar, “Masculinity and Autonomous Vehicles: A Degendered or Resegregated Future System of Automobility?” *Transfers* 8, no. 1 (2018), 44–63, https://doi.org/10.3167/TRANS.2018.080105.

5. John Markoff, “Google Cars Drive Themselves, in Traffic,” *New York Times*, 9 October 2010.

6. “Morals and the Machine,” *Economist*, 2 June 2012.

7. Michael Anderson, Susan Leigh Anderson, and Chris Armen, “Towards Machine Ethics: Implementing Two Action-Based Ethical Theories,” paper presented at the AAAI Fall Symposium on Machine Ethics, Arlington, VA, 3–6 November 2005; Patrick Lin, George Bekey, and Keith Abney, “Autonomous Military Robotics: Risk, Ethics, and Design” (San Luis Obispo: Ethics + Emerging Sciences Group at California Polytechnic State University, 20 December 2008); Luke Muehlhauser and Louie Helm, “Intelligence Explosion and Machine Ethics,” in *Singularity Hypotheses: A Scientific and Philosophical Assessment*, ed. Amnon H. Eden, Johnny Søraker, James H. Moor, and Eric Steinhart (Berlin: Springer, 2012), 101–126.

8. Gary Marcus, “Moral Machines,” *New Yorker*, 27 November 2012, http://www.newyorker.com/online/blogs/newsdesk/2012/11/google-driverless-car-morality.html (school bus vs. railing); Noah J. Goodall, “Ethical Decision Making during Automated Vehicle Crashes,” *Transportation Research Record: Journal of the Transportation Research Board* 2424, no. 1 (2014): 58–65, https://doi.org/10.3141/2424-07 (bus vs. railing; motorcyclists); Burkhard Bilger, “Auto Correct: Has the Self-Driving Car at Last Arrived?” *New Yorker*, 25 November 2013 (animal vs. child); Patrick Lin, “Why Ethics Matters for Autonomous Cars,” in *Autonomes Fahren*, ed. Markus Maurer, J. Christian Gerdes, Barbara Lenz, and Hermann Winner (Berlin: Springer, 2015), 69–85 (child vs. grandmother); Jean-François Bonnefon, Azim Shariff, and Iyad Rahwan, “The Social Dilemma of Autonomous Vehicles,” *Science* 352, no. 6293 (2016): 1573–76, https://doi.org/10.1126/science.aaf2654 (pedestrians).

9. Philippa Foot, “The Problem of Abortion and the Doctrine of Double Effect,” *Oxford Review* 5 (1967): 5–15; Judith Jarvis Thomson, “The Trolley Problem,” *Yale Law Journal* 94, no. 6 (1985): 1395–1415; Judith Jarvis Thomson, “Turning the Trolley,” *Philosophy & Public Affairs* 36, no. 4 (2008): 359–374, https://doi.org/10.1111/j.1088-4963.2008.00144.x.

10. Peter Danielson, “Surprising Judgments about Robot Drivers: Experiments on Rising Expectations and Blaming Humans,” *Etikk i Praksis—Nordic Journal of Applied Ethics* 9, no. 1 (2015): 73–86, here 78.

---

11. Keeling, in a defense of the trolley problem in automated vehicle ethics, outlines four arguments against it: the Not Going to Happen Argument, the Moral Difference Argument, the Impossible Deliberation Argument, and the Wrong Question Argument. See Geoff Keeling, “Why Trolley Problems Matter for the Ethics of Automated Vehicles,” *Science and Engineering Ethics* (first online 4 March 2019).

12. Alex Roy, “Autonomous Cars Don’t Have a ‘Trolley Problem’ Problem,” *The Drive*, 19 October 2016, http://www.thedrive.com/tech/5620/autonomous-cars-dont-have-a-trolley-problem-problem.

13. Rodney Brooks, “Unexpected Consequences of Self Driving Cars,” *Robots, AI, and Other Stuff*, 12 January 2017, https://rodneybrooks.com/unexpected-consequences-of-self-driving-cars; Matt McFarland, “Google’s Chief of Self-Driving Cars Downplays ‘the Trolley Problem,’” *Washington Post*, 1 December 2015.

14. Noah J. Goodall, “Machine Ethics and Automated Vehicles,” in *Road Vehicle Automation*, ed. Gereon Meyer and Sven Beiker (Berlin: Springer, 2014), 93–102; Keeling, “Why Trolley Problems Matter.”

15. USDOT and NHTSA, “Federal Automated Vehicle Policy” (Washington, DC: US Department of Transportation and the National Highway Traffic Safety Administration, September 2016).

16. NHTSA’s revised guidance has removed Ethical Considerations as a standalone safety assessment area. Aspects of vehicle automation with ethical implications, such as risk assessment and the challenge of balancing the objectives of legality and mobility, are discussed elsewhere in the guidance. See USDOT and NHTSA, “Automated Driving Systems (ADS): A Vision for Safety 2.0.” (Washington, DC: United States Department of Transportation and the National Highway Traffic Safety Administration, September 2017), 15.

17. USDOT and NHTSA, “Federal Automated Vehicle Policy,” 26.

18. Goodall, “Ethical Decision Making”; Goodall, “Machine Ethics and Automated Vehicles”; Keeling, “Why Trolley Problems Matter”; Lin, “Why Ethics Matters for Autonomous Cars.”

19. Scott Le Vine, You Kong, Xiaobo Liu, and John Polak, “Vehicle Automation and Freeway ‘Pipeline’ Capacity in the Context of Legal Standards of Care,” *Transportation* 46, no. 4 (2017): 1215–1244, https://doi.org/10.1007/s11116-017-9825-8.

20. Mark Michaelian and Fred Browand, “Field Experiments Demonstrate Fuel Savings for Close-Following,” PATH Research Report, September 2000.

21. J. O. Pearson, “Automobiles: Sudden Emergency as Exception to Rule Requiring Motorist to Maintain Ability to Stop within Assured Clear Distance Ahead,” in *American Law Reports 75*, Third, 2005, 333.

22. Vine et al., “Vehicle Automation,” 8 (block quote), 18.

---

23. Alex Hern, “Self-Driving Cars Don’t Care about Your Moral Dilemmas,” *The Guardian*, 22 August 2016.

24. Myra Blanco, Jon Atwood, Sheldon M. Russell, Tammy Trimble, Julie A. McClafferty, and Miguel A Perez, “Automated Vehicle Crash Rate Comparison Using Naturalistic Data” (Virginia Tech Transportation Institute, January 2016).

25. Bilger, “Auto Correct.”

26. David N. Lee, “A Theory of Visual Control of Braking Based on Information about Time-to-Collision,” *Perception* 5, no. 4 (1976): 437–459, https://doi.org/10.1068/p050437.

27. “Required Position and Method of Turning at Intersections; Local Regulations,” § 46.2–846 Code of Virginia, accessed 29 October 2019, http://law.lis.virginia.gov/vacode/title46.2/chapter8/section46.2-846.

28. Dmitri Dolgov and Christopher Urmson, “Controlling Vehicle Lateral Lane Positioning,” US Patent and Trademark Office no. 8781670, filed 28 May 2013, issued 15 July 2014, http://patft.uspto.gov/netacgi/nph-Parser?Sect2=PTO1&Sect2=HITOFF&p=1&u=/netahtml/PTO/search-bool.html&r=1&f=G&l=50&d=PALL&RefSrch=yes&Query=PN/8781670.

29. This, of course, depends on the specifics of the situation: it could be that preferring to crash into the bus is safer, as the bus may be more likely to stay on the road, whereas crashing into the small car may be more likely to result in the small car coming off the road, perhaps leading to a worse crash.

30. For an example model of crash injury severity as a function of vehicle mass and speed, see Reza Tolouei, Mike Maher, and Helena Titheridge, “Vehicle Mass and Injury Risk in Two-car Crashes: A Novel Methodology,” *Accident Analysis and Prevention* 50 (January 2013): 155–166, https://doi.org/10.1016/j.aap.2012.04.005; Bonnefon et al., “Social Dilemma of Autonomous Vehicles.”

31. David Z. Morris, “Mercedes-Benz’s Self-Driving Cars Would Choose Passenger Lives Over Bystanders,” *Fortune,* 15 October 2016, http://fortune.com/2016/10/15/mercedes-self-driving-car-ethics.

32. Geoff Keeling, Katherine Evans, Sarah M. Thornton, Giulio Mecacci, and Filippo Santoni de Sio, “Four Perspectives on What Matters for the Ethics of Automated Vehicles,” in *Road Vehicle Automation 6,* ed. Gereon Meyer and Sven Beiker (Berlin: Springer, 2019).

33. Nicholas J. Garber, “Speed Variance and Its Influence on Accidents” (Washington, DC: AAA Foundation for Traffic Safety, July 1988).

34. Paul Ingrassia, “Look, No Hands! Test Driving a Google Car,” *Reuters*, 17 August 2014, http://uk.reuters.com/article/us-google-driverless-idUKKBN0GH02P20140817.