**Autonomous Vehicles and Virtue Ethics in Societal Experimentation**

STS Research Paper
Presented to the Faculty of the
School of Engineering and Applied Science
University of Virginia

By

Joseff Medina

April 14, 2020

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signed: Joseff Medina

Approved: Benjamin Laugelli Date _____
Benjamin J. Laugelli, Assistant Professor, Department of Engineering and Society

## Introduction

On the night of March 18, 2018, a robot Uber vehicle killed Elaine Herzberg. This is the first recorded event in which a vehicle designed to be completely autonomous claimed the life of a human being (Griggs & Wakabayashi, 2018). While there was a safety test driver inside the Uber vehicle, the engineered autonomous features of the vehicle made it difficult to place blame solely on the driver. As a fully autonomous vehicle, one may question whether the modified Volvo XC90 should have been capable of stopping itself before striking Mrs. Herzberg. After all, one of the main selling points of fully autonomous vehicles is that they eliminate virtually all human error on the road (National Highway Traffic Safety Administration, n.d., para. 4). With the event having occurred within recent years, the issue of what or who is to blame is still a hotly debated topic among scholars. Within the debate two camps for argument are noted. The first identifies regulation shortfalls as a leading contributor to the accident. The second line of argument focuses on the method of testing design as a shortfall contributing to the accident. While both of these camps have their merits, they fail to view Uber within the realm of engineering virtue ethics.

Through the ethical framework of virtue ethics, I will show that Uber is to be held responsible for the death of Elaine Herzberg. By failing to practice and become proficient in the cardinal principles of virtue ethics, such as prudence and temperance, Uber presents its autonomous vehicle experiment as a prime example of failed societal experimentation implementation. I will utilize Uber's practices with autonomous vehicle technology throughout my argument to demonstrate the important relationship between virtue ethics and societal experimentation.

## Background

Uber is a public transportation company, specializing in the use of an on-demand taxi service through the use of its app. Starting in 2016, Arizona granted access to Uber, allowing experimentation and testing of its fully autonomous vehicles on public roads (Harris, 2018). More than a year later, in March of 2018, Elaine Herzberg was struck and killed by one of these autonomous vehicles. Travelling at forty miles per hour, records confirmed that the car did recognize Elaine Herzberg as a pedestrian but failed to stop on its own accord. The Uber test driver was not paying attention at the time of the crash (National Transportation Safety Board [NTSB], 2018).

While other injuries and deaths involving autonomous vehicles had occurred before March 2018, it was determined that these accidents were at the fault of human operators (Vlasic & Boudette, 2016). In the case of Elaine Herzberg, both Uber's vehicle and the test driver were blamed. According to police, the test driver could have avoided the accident with manual intervention. However, it was also noted that in addition to failing to brake, Uber's autonomous vehicle had no way of warning its test driver of any impending danger (NTSB, 2018). Thus, Uber's autonomous vehicle carried a fatal flaw. Despite initially cancelling operations in several states and cities as a result of the Herzberg incident, Uber has since restarted operations in Pittsburgh as of December 2018 (Wakabayashi & Conger, 2018). Uber faced little legal punishment (Borenstein, Herker, & Miller, 2019).

## Literature Review

Because of the Herzberg incident's more recent occurrence, there does not yet seem to be a solidified consensus among scholars on how to approach the event. It is undeniable that a multitude of factors played into Uber's failure to prevent Mrs. Herzberg's death at its vehicle's

actions. Within this document, I present the assessments of other scholars (Heeven, 2018; Zhou & Sun, 2019), addressing the principle variables contributing to the incident and what we should learn from them.

The first example discussed here emphasizes the potential lack of robotic and artificial intelligence (AI) legislation as the cause of the crash. According to Douglas Heaven (2018), current roadway legislation is outdated and unfit for autonomous vehicles. The Uber vehicle involved in the Herzberg crash was intentionally designed in such a way that automatic breaking was disabled during automatic driving mode, so as to "ensure a smoother ride" (Heaven, 2018, p.38). At the same time, there was no system in place to warn the driver of an object in front of the vehicle (NTSB, 2018). This is a very serious design oversight on Uber's part, and yet it is still within regulations primarily due to slack or lacking legislation. Following Heaven's logic, it is quite possible that a lack of regulation was the cause of Elaine Herzberg's death. Without societal pressure, Uber was left unchecked to conduct its autonomous vehicle experiment on public roads without the informed consent from the Tempe, Arizona citizens (Borenstein et al, 2019). As a consequence, it took a fatal accident to draw attention to the shortfalls of policy and regulation impacting the introduction of autonomous vehicles for public transit, challenging the rolls and responsibilities of Uber in the case.

Heaven (2018) continues, predicting on how society's future may be affected if more advanced autonomous technologies, such as AI controlled vehicles, are further developed without proper societal guidance. He states that it is essential that society shape autonomous systems, imparting the ethics of its operators on to it. By doing so, manufacturers and regulators must strive to anticipate the positive and negative consequences of those systems, with the goal to minimize the damage caused by AI to society. By understanding why an AI behaves a certain

way and imparting the ethics of its operators on to it, Heaven theorizes that harmony can be achieved between humanity and autonomous systems. Arizona's decision to accept Uber's self-driving cars onto its roads without understanding the technology could have been a factor leading up to Herzberg's death.

In contrast to the regulatory shortfall argument, other scholars place blame for the Herzberg incident on Uber's methods of testing its autonomous vehicle technology. According to Zhi Quan Zhou and Liqun Sun (2019), Uber's use of human test drivers as moderators and oracles for its autonomous vehicles was not well vetted. An oracle is a mechanism used in software development to determine whether the outcomes of test-case executions are correct. According to police and Uber records of the incident, all onboard hardware of the modified Volvo XC90 functioned properly. Zhou and Sun (2019) claim that the fatal fault lay in the integration between the test driver and the software. An example of highly complex software, it would be expensive and difficult for Uber to develop a proper, computerized oracle to test its self-driving cars. As a compromise, Uber deemed that human observation would be a sufficient means of monitoring and reporting the results of its self-driving car experiments (Zhou & Sun, 2019).

Zhou and Sun (2019) propose humans are insufficient oracles for the purposes of testing such advanced, critical, and sensitive software. Humans are not perfect, and can be particularly error prone. For example, "humans monitoring an automated system are likely to become bored and disengaged" (Levin, 2018, para. 5), resulting in dangerous situations and faulty testing. According to dashcam footage, the Uber test driver in the vehicle that killed Mrs. Herzberg appeared to lack due care at the time of the incident (Smith, as cited in Levin, 2018). This suggest the test driver was not attentive and prepared to attempt a safety response. In order to

rectify this fault in Uber's testing methods, Zhou and Sun (2019) suggest a new method of autonomous testing be put in place. They propose metamorphic, autonomous testing would not only further eliminate human error, but would also provide more meaningful and quantifiable test data based on data points directly from the autonomous driving software.

The examples of scholarly analyses I have just discussed put forth quite disparate claims as to what caused the death of Elaine Herzberg. Both articles provide different calls to action for different parties involved in the incident. In addition, both scholarly analyses do not attempt to analyze the incident from an ethical standpoint. Morality plays a key role in what occurred on that night in March of 2018. The lack of virtue ethics on Uber's part should not be overlooked.

## Conceptual Framework

I will draw upon the ethical framework of virtue ethics in order to argue my case of Uber's moral fault. Virtue ethics will allow me to demonstrate the essential character traits that Uber should have practiced in dealing with a new and untested technology. In addition, I will also demonstrate the importance of risk analysis in the correct application of a societal experiment.

According to Van de Poel and Royakkers (2011), virtue ethics is an "ethical theory that focuses on the nature of the acting person" (p. 95). The nature of a person can further be defined as reflecting particular character traits. Some examples of virtuous character traits include humility, reliability, self-control, and courage. An entity is not inherently virtuous. Achieving a virtuous state is only possible through extended practice, implementation, or mentorship. According to Aristotle, the growth and development of an entity's character into a mature and virtuous character is essential so that its virtuous traits can be applied correctly (Van de Poel & Royakkers, 2011). It is important to be able to use a trait within moderation. For example, the

virtue of courage can become rashness if used excessively, and cowardice if used too little.Thus, the correct application of a virtue can only be learned through experience and practice.

Risk is defined as a measurement of a particular hazard. According to van de Poel and Royakkers (2011), risk most commonly "is the product of the probability of an undesirable event and the effect of that event" (p. 221). Ideally, risks can be quantified and eliminated in a new technology before it is introduced into society. However, it is not always possible to know how a risk will play out or if there is an unknown risk when the technology is used in non-simulated, real world conditions. The precautionary principle is often applied in such cases, where specific protocols are put in place to deal with uncertain risks. An example of the precautionary principle in action would be a speed limiter on a vehicle. Even though the vehicle may be capable of much greater speeds, the speed limiter is used as a precaution in order to prevent the possibility of unknown risks causing harm at high speeds. In contrast to the precautionary principle, societal experiments are a way for engineers to test how their new technologies behave with less precautions. In a societal experiment, a technology is introduced to society at large in order to evaluate its interactions and effects on society in a less strict and unsimulated environment. Societal experiments are a compromise between risk mitigation and the further development of technology, with the potential for disastrous consequences if poorly implemented.

In the next section, I will demonstrate Uber's moral failure through the ethical framework of virtue ethics by demonstrating the necessity of virtue ethics in establishing a morally just societal experiment. In particular, I will focus on the virtuous character traits of prudence and temperance. Prudence is the virtue of foresight and care, whereas temperance is the virtue of restraint.

## Evidence

### *Prudence*

The death of Elaine Herzberg was the result of Uber's failure to practice the virtuous traits of prudence and temperance. Uber first failed, in the virtue of prudence. Prudence is the virtue of foresight and wisdom, in which actions are disciplined and based off of careful reasoning. For any societal experiment to be considered morally just, all known risks and associated technological faults should be minimized before societal adoption of the technology.

The New York Times address the reestablished autonomous vehicle experimentation efforts of Uber after the death of Elaine Herzberg (Wakabayashi & Conger, 2018). In December of 2018, in contrast to its previous operations in Arizona, Uber resolved to limit its experiments to only occurring within shorter routes, lower speeds, and in sunny, clear weather. Despite its scaled down efforts, quite a number of persistent faults in its technology surfaced. "The cars have reacted more slowly than human drivers and struggled to pass so-called track validation tests, the last step before returning to city streets, according to a dozen Uber documents and emails as well as interviews with seven current and former employees, who spoke on the condition of anonymity because they were not allowed to talk publicly about the company," (Wakabayashi & Conger, 2018, para. 3). In addition, many test drivers and engineers within Uber have voiced their concern over the company's tendency to test software updates and hardware on the road without prior controlled testing. This evidence from The New York Times suggests that Uber did not have the foresight to fully test for faults within its self-driving technology before attempting real-world street implementation (Wakabayashi & Conger, 2018).

It can be assumed that these technological issues were present in Uber's autonomous vehicles at the time of Elaine Herzberg's death. Applying this assumption, Uber had either

implemented a new technology without sufficient controlled testing, or had knowingly introduced a risky technology into society. In either case, it was an action lacking in the virtue of prudence. The testimonies of Uber's own employees as to the use of untested technology serve to further prove Uber's lack of prudence. With such a new technology, the wise decision would have been to thoroughly test and minimize all known risks before introducing it into a societal experiment. Thus, Uber acted without prudence in order to advance its own agenda, eventually at the cost of Elaine Herzberg's life.

Uber acted immorally with respect to virtue ethics, as its lack of prudence resulted in the company disregarding discussions and observations of faults in its technology. A possible counterargument to this would be that Uber behaved morally under utilitarian ethics parameters. In contrast to virtue ethics, utilitarian ethics bases moral actions upon objective and quantifiable logic and facts. Under utilitarian ethics, it is possible that Uber believed that it had sufficiently solved the technology issues impacting its self-driving cars' safety parameters to the best of its ability at the time the autonomous vehicle's were introduced in Arizona, based on decisions and conclusions of senior engineers. According to the New York Times, "Eric Meyhofer, who heads the unit, declared that Uber was going back to 25 m.p.h. The faster speed would prove that the cars were 'unequivocally worthy of being back on the road,'" (Wakabayashi & Conger, 2018, para. 12). Under Utilitarian ethics, it could be argued that an action is morally just if it is based on the logic, experience, and guidance of senior engineers. If this is the case, then Uber would have implemented a societal experiment correctly. In addition, utilitarian ethics, as an ethical framework based on facts and logic, would most likely disregard the comments of Uber's anonymous, concerned employees as rumors not in favor with the goals of the company.

This utilitarian viewpoint is faulty. Even if it were the case that Uber's senior engineers were giving final approval for the autonomous vehicle technology, the prudent and wise decision would have been to consider the amount of concern being raised by numerous employees as a worthy point of investigation. According to The New York Times, some "engineers thought there was another reason: Mr. Meyhofer wanted to demonstrate progress to his boss, Mr. Khosrowshahi. And they worried that Uber was taking shortcuts to hit internal milestones, according to two current employees," (Wakabayashi & Conger, 2018). This quote suggests that, despite his senior status as an engineer, Mr. Meyhofer did not act in a manner reflective of the prudence expected in his position. This potential for error would only be mitigated if Uber had the prudence to investigate Meyhofer's actions based on the concern of its employees, rather than rely on the utilitarian procedure of accepting senior decisions as is. In the event that the concerns are just rumors, then no harm is done and a morally just action was still performed. With an acceptable amount of prudence, a societal experiment's chance of success can be increased through the mitigation of risk.

***Temperance***

Uber also failed to exercise the virtue of temperance, or self-control. An action is morally just with regards to temperance if it is done in a manner that is not impulsive or apathetic. Temperance is important with regards to the implementation of a societal experiment so that it is done in a timely, yet not hasty, manner.

According to evidence from The Guardian, Uber acted in haste and selfishness to take advantage of a behind closed doors deal with Arizona's governor (Harris, 2018). The Guardian obtained emails between Governor Ducey and Uber. Uber and Governor Ducey were in close communication from the start of his term. Only a month after he was sworn in, Governor Ducey

began meetings with Uber executives. Many of the new Governor's first acts consisted of putting in place new legislation to appease Uber and other on demand taxi app companies. In June of 2015, only 6 months after assuming office, Governor Ducey simultaneously opened a customer support center for Uber in Phoenix, AZ, and issued an executive order allowing public testing of autonomous vehicles. This new legislation also put in place a new Autonomous Vehicle Oversight Committee. All members of this committee were loyal to Governor Ducey, and the committee only met once, with no action being taken between the time of its creation and the writing of The Guardian article (Harris, 2018). Ducey and Uber worked together to obtain permission from Phoenix's City Council for Uber to operate at Sky Harbor Airport. According to a member of Phoenix's city council, Ducey's office threatened the council in the event that the council did not pass legislation favorable to Uber (Harris, 2018). In August 2016, Uber publicly introduced its autonomous vehicle fleet to Pittsburgh, despite many experts raising concern over an extremely new and untested technology being used in society (Harris, 2018). That same month, it also secretly introduced the same program in Phoenix. Uber emailed Governor Ducey's office to let them know of its operations, while also asking for a "discreet" point of contact within Phoenix's Police Department with whom Uber could establish communications (Harris, 2018). According to The Guardian, because "of Arizona's regulatory vacuum, neither Uber nor Ducey were obliged to inform the public that Uber's cars would now be driving themselves on public roads. Neither, it seems, did they believe they had an ethical duty to do so" (Harris, 2018, para. 29). In December of 2016, Uber was caught testing autonomous vehicles in San Francisco without applying for the proper permits. After several sightings of Uber's autonomous vehicles running red lights, the program was banned from California (Harris, 2018). Rather than reworking the autonomous vehicle technology, Uber immediately reintroduced the same vehicles

into Arizona's fleet of vehicles. Governor Ducey is quoted as saying "Arizona welcomes Uber self-driving cars with open arms and wide open roads" (Harris, 2018, para. 34). Rather than waiting until its technology was more mature and safer, Uber acted rashly and without temperance to use Arizona as a playground for still dangerous and untested technology.

From the evidence above, it is made clear that Uber took shortcuts in order to put its autonomous vehicles on public roads as soon as possible. Its eagerness to implement societal experimentation, regardless of any warning signs or proper communication channels, shows a lack of temperance. Starting with Uber's initial dealings with Governor Ducey, it is made evident that Uber attempted to form a relationship with the Arizona Government that would allow the company to act without constraint on Arizona's streets. Following initial talks with Arizona's governor, Uber's attainment of operability within Sky Harbor airport created an example of Uber testing its newfound relationship. Uber then gained confirmation that it could do anything it wanted, without temperance, within the state of Arizona once the Autonomous Vehicle Oversight Committee was filled with inactive figureheads who failed to regulate Uber's actions. The same month of the Autonomous Vehicle Oversight Committee's founding, Uber publicly introduced its autonomous vehicles onto public roads in Pittsburgh. Uber also introduced a secret autonomous vehicle program into Phoenix at the same time. Both Uber and the Arizona government were aware of the program in Phoenix, but neither party made an effort to inform the public. This allowed the company to act without any particular restraints because the technology was introduced in such a manner that the public was unaware, and yet also under the protection of the Arizona governor. Uber then illegally implemented autonomous vehicle tests in San Francisco. The technology was proven to be defective when the vehicles were sighted running red lights and were banned from California. Despite these points of evidence, Uber still

lacked temperance and simply moved these test vehicles into Arizona's fleet of autonomous vehicles, where Governor Ducey encouraged its operations to continue. Uber's behavior in Arizona, the secretive implementation of unreliable vehicles in San Francisco, and then the transfer of those cars to the Arizona fleet all suggest Uber acted with a lack of self-control and restraint. As a result of Uber's lack of temperance, its societal experiment was rushed and implemented without a strong legal foundation or without regard to safety.

## Conclusion

As a result of Uber's failure to act in a moral manner, its societal experiment involving autonomous vehicles killed Elaine Herzberg. Uber acted in a manner lacking of the virtuous characters of prudence and temperance. With new technology, it is essential that any known risks and possible risks be minimized before attempting to implement a societal experiment. Uber knowingly failed to do this, and chose to place the rest of society at risk in order to further advance its own agenda. In addition, Uber's close relationship with Arizona's governor allowed the company free reign within the state, and Uber took advantage of this relationship without temperance. Uber rushed its technology into service, disregarding consideration for public safety and concern. Consequently, Elaine Herzberg lost her life to this new and untested technology. Had Uber practiced proper temperance and prudence, it is possible that its societal experiment may have occurred without incident. Prudence minimizes the chance of risks coming to fruition within a societal experiment. Temperance allows a technology to be processed through proper channels and reach a maturation level that support a successful societal experiment. The outcome of a societal experiment is heavily dependent on an engineer's practice and correct implementation of virtue ethics, such as prudence and temperance.

## References


Borenstein, J., Herkert, J., & Miller, K. W. (2019). Autonomous vehicles and the ethical tension between occupant and non-occupant safety. In D. Wittkower (Ed.), *2019 Computer Ethics - Philosophical Enquiry (CEPE) Proceedings*, (14 pp.). Retrieved from https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety#nhtsa-action

Griggs, T., & Wakabayashi, D. (2018, March 21) How a self-driving Uber killed a pedestrian in Arizona. *The New York Times*. Retrieved from https://www.nytimes.com/interactive/2018/03/20/us/self-driving-uber-pedestrian-killed.html

Harris, M. (2018, March 28). Exclusive: Arizona governor and Uber kept self-driving program secret, emails reveal. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2018/mar/28/uber-arizona-secret-self-driving-program-governor-doug-ducey

Heaven, D. (2018). Robot laws. *New Scientific, 239*(3184), 38-41. Retrieved from https://www.sciencedirect.com/science/article/abs/pii/S0262407918313940

Levin, S. (2018, March 22). Uber crash shows 'catastrophic failure' of self-driving technology, experts say. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2018/mar/22/self-driving-car-uber-death-woman-failure-fatal-crash-arizona

National Highway Traffic Safety Administration (n.d.) *Automated vehicles for safety* [brochure]. Retrieved from https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety#nhtsa-action

National Transportation Safety Board (2018). *Collision between vehicle controlled by developmental automated driving system and pedestrian* (Accident report HWY18MH010). Retrieved from https://www.ntsb.gov/investigations/AccidentReports/Reports/HWY18MH010-prelim.pdf

Van de Poel, I., & Royakkers, L. (2011). *Ethics, technology, and engineering: an introduction*. Chichester, West Sussex, UK: Wiley-Blackwell.

Vlasic, B., & Boudette, N. (2016, June 30). Self-driving Tesla was involved in fatal crash, U.S. says. *The New York Times*. Retrieved from https://www.nytimes.com/2016/07/01/business/self-driving-tesla-fatal-crash-investigation.html

Wakabayashi, D., & Conger, K. (2018, December 5). Uber's self-driving cars are set to return in a downsized test. *New York Times*. Retrieved from https://www.nytimes.com/2018/12/05/technology/uber-self-driving-cars.html

Zhou, Z., & Sun, L. (2019). Metamorphic testing of driverless cars. *Communications of the ACM, 62*(3), 61-67. Retrieved from https://dl.acm.org/doi/abs/10.1145/3241979