Self-Driving Cars and Data Collection: Privacy Perceptions of Networked Autonomous Vehicles

Cara Bloom, Joshua Tan, Javed Ramjohn, and Lujo Bauer, ***Carnegie Mellon University***

https://www.usenix.org/conference/soups2017/technical-sessions/presentation/bloom

This paper is included in the Proceedings of the Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017).

July 12–14, 2017 • Santa Clara, CA, USA

ISBN 978-1-931971-39-3

Open access to the Proceedings of the Thirteenth Symposium on Usable Privacy and Security is sponsored by USENIX.

Self-Driving Cars and Data Collection: Privacy Perceptions of Networked Autonomous Vehicles

Cara Bloom Joshua Tan Javed Ramjohn Lujo Bauer
Carnegie Mellon University
{cbloom, jstan, lbauer, jramjohn}@andrew.cmu.edu

ABSTRACT

Self-driving vehicles and other networked autonomous robots use sophisticated sensors to capture continuous data about the surrounding environment. In the public spaces where autonomous vehicles operate there is little reasonable expectation of privacy and no notice or choice given, raising privacy questions. To improve the acceptance of networked autonomous vehicles and to facilitate the development of technological and policy mechanisms to protect privacy, public expectations and concerns must first be investigated. In a study (n=302) of residents in cities with and without Uber autonomous vehicle fleets, we explore people's conceptions of the sensing and analysis capabilities of self-driving vehicles; their comfort with the different capabilities; and the effort, if any, to which they would be willing to go to opt out of data collection. We find that 54% of participants would spend more than five minutes using an online system to opt out of identifiable data collection. In addition, secondary use scenarios such as recognition, identification, and tracking of individuals and their vehicles were associated with low likelihood ratings and high discomfort. Surprisingly, those who thought secondary use scenarios were more likely were more comfortable with those scenarios. We discuss the implications of our results for understanding the unique challenges of this new technology and recommend industry guidelines to protect privacy.

1. INTRODUCTION

Networked autonomous robots in the form of drone swarms and commercial autonomous vehicles (AVs) are being researched, tested, and deployed. This technology is set to fundamentally shift common daily practices such as the use and ownership of automobiles [52]. At the time of data collection, Uber's self-driving car fleet had been deployed in Pittsburgh, PA for five months and was planning to expand to other states. The fleet is large enough that seeing the AVs has become quotidian to residents.

Two ethical concerns with the growing prevalence of AVs have received significant attention in the media and academic discourse. Ethical decision making—especially concerns with life-or-death decisions made by AVs—has been a major focus and has influenced public willingness to accept AVs as decision makers [41]. Commercial drivers, labor economists, and corporations have focused on the market effects of robots taking human jobs, both positive and negative [53]. A third ethical concern, and the focus of this paper, is the privacy-invasive capabilities of AVs, as well as the potential security risks associated with AV data collection. This ethical concern has received very little attention relative to decision-making and labor market changes.

Commercial fleets of networked AVs have the capability to collect location and movement data about residents of an entire city simply by storing the information already captured by their many sensors and using available software to analyze it. This capability poses a new regulatory conundrum, as it combines four different aspects of privacy-invasive technologies: (1) the ubiquitous capture of data in public, (2) physical surveillance by a privately owned company, (3) the ability to scale without additional infrastructure, and (4) the difficulty of notice and choice about data practices for physical sensors that capture data about non-users. Ubiquitous data collection in public has been implemented by cities such as London [48], which has sparked public debate over the efficacy and morality of surveillance. While cities are beholden to their constituents and residents, companies are beholden to their shareholders [13]. If a city like London and a company like Uber have the same data set of geo-temporal points, the former has an obligation to use it to better its constituents and the latter has an obligation to monetize it, bettering its shareholders.

While similar issues also apply to CCTV and dashboard cameras, the scalability and potential ubiquity of a networked self-driving car fleet is remarkable. Unlike CCTV, AVs can increase the bounds of their surveillance without additional infrastructure and can cover any public roads they are legally permitted to drive on. They use public infrastructure and are not reliant on privately owned property. The networked aspect differentiates them from dashboard cameras or individual self-driving cars (such as future Fords or Teslas) due to the scale of data collection and analytic capabilities on such aggregated sensor data.

These vehicles operate in public spaces where individuals do not have a reasonable expectation of privacy, and where notice and choice would be difficult to provide. Internet of things (IoT) devices such as Alexa already suffer from the difficulty of sufficiently notifying users of data collection (no

Symposium on Usable Privacy and Security (SOUPS) 2017, July 12–14, 2017, Santa Clara, California.

screen means no terms and conditions to read on the device), a task made more difficult when devices collect data from non-users [40]. Autonomous vehicles on public roads are constantly interacting with non-users: people who have not consented to any data collection or use and, as demonstrated in our study, have not yet thought about the potential privacy impacts.

As with many powerful new technologies, the large-scale capture and analysis of data enabled by AVs could lead to both benefits to the public and concerns. Ubiquitous sensing capabilities could be used to find Amber or Silver Alert citizens [48], but the same technology could also be used for less altruistic purposes. Insurers might analyze license plate logs to find out whether a customer speeds on the highway and adjust her car insurance rates accordingly, countries could identify dissidents, or an employee of the AV company could use the technology to stalk a celebrity or an ex-girlfriend, as Uber employees were found doing [7]. AV sensors could log the physical movements of every person within the purview of the fleet, making it possible to find anyone anywhere. The chilling effects of such surveillance and related dangers of ubiquitous data collection are well documented in privacy and security literature [39].

Reasonable expectation, unfairness, and deception are central themes for privacy regulation in the United States, so one key question is: where does the public draw the line between acceptable and unacceptable practices for autonomous networked robots? New technologies such as the Internet and more recently the Internet of Things can outpace the creation of reasonable privacy standards as they are quickly integrated into people's daily lives, leading to many inventive studies of the gaps in protection and how to patch them (e.g., [28, 17]). Users of these technologies can become habituated to the lack of privacy [51], making usable, effective privacy protections more difficult to enact. Therefore, it is important to explore privacy conceptions and strategies for privacy protection during the earliest phases of a new technology's implementation, before deployment outpaces the incorporation of privacy.

Whereas other potentially privacy-invasive technologies have required users opt in, AVs cannot give all pedestrians and drivers they encounter notice and choice. Companies operating such fleets could potentially offer notice outside of the information capture environment, but it would be difficult to give people the choice to opt out of information collection in all forms. Some information would have to be collected during the opt-out process, such as a license plate number to opt out of license plate recognition. Other options such as an opt-in process, privacy policies that limit the use of collected data, or even the removal of identifiable markers from stored data, are possible approaches. To make recommendations to the few companies currently operating in the space of networked AVs, privacy conceptions about the technology and its potential uses must first be understood.

Our investigation aims to fill this gap by exploring conceptions of the sensing and analysis capabilities of AVs; people's comfort with the different capabilities; and the effort, if any, to which they would be willing to go to opt out of data collection. We ran an online study of 302 participants using scenarios of increasing privacy invasiveness to measure how likely participants thought different potential capabilities of self-driving vehicles are, and how comfortable they are with those capabilities. Scenarios were framed using the Uber self-driving car fleet as an example. We recruited in Pittsburgh where the fleet has been deployed since September 2016 in addition to four other cities to investigate whether exposure to the technology changed conceptions or sentiments.

In addition to questions about likelihood and comfort with privacy capabilities, participants answered questions about general AV technology concerns like safety, their exposure to self-driving cars, bias against Uber, and demographic information. Responses were analyzed to determine likelihood and comfort levels as well as the relationship between likelihood, comfort, and potential explanatory variables.

We found that participants consider primary uses of AV sensors such as data collection, aggregation, storage, and analysis by the cars to be likely, and that participants express moderate comfort with these scenarios. Secondary use scenarios such as the recognition, identification, and tracking of individuals or their vehicles received the lowest ratings of likelihood and highest discomfort. Surprisingly, participants who thought the technology was more likely to have a privacy-invasive capability such as tracking were more likely to be comfortable with that capability. Though participants rated many capabilities likely and expressed high levels of discomfort, only one out of three would spend more than 10 minutes using an online opt-out system.

Pittsburgh participants who had exposure to the Uber self-driving car fleet (over 60% had seen one compared to 3% for other cities) were not statistically different in their conceptions of likelihood and comfort from residents of other cities who had never seen a self-driving car. The only factor that showed a significant increase in opt-out time was whether participants had received the privacy scenario priming questions, which participants noted had raised difficult questions they had not considered before. If public attention surrounding AVs expands from safety and employment issues to privacy issues, our findings suggest that peoples' overall comfort with AVs may increase, but so might privacy-seeking behavior as well. Understanding the complex privacy concerns in this space is essential for developing industry practices and regulation.

2. BACKGROUND AND RELATED WORK

The classic work in the area of AVs and privacy discusses the privacy implications for owners and users of AVs in detail and alludes to surveillance, noting that "[networked] autonomous vehicles could enable mass surveillance in the form of comprehensive, detailed tracking of all autonomous vehicles and their users at all times and places." The work focuses solely on the passengers within an AV who have ostensibly agreed to the terms and conditions, legally relinquishing their privacy the same way consumers do when using Google Maps [16]. In this paper we assess the more complex privacy concerns of those who interact with AVs, but are not necessarily users of the system. We next review consumer perception of AVs, followed by their technological capabilities and relevant regulations.

2.1 Consumer Perception

Research into consumer perceptions of AVs has examined general interest, trust in the cars' reliability and safety, and

consumer feelings about how self-driving cars could impact the job market. Our work is one of the few that focuses on consumer privacy concerns and preferences regarding self-driving cars.

Consumer perception has been a popular area of discussion and research, given its potential impact on sales and market adoption. With AVs being deployed in test locations and viable plans to bring them to mass-market, studies have been conducted to gauge consumer interest. Schoettle and Sivak found that people are generally uninformed and had both high expectations about the benefits of autonomous vehicles and high levels of concern about riding in them. Additional concerns were changes to the job market, security against hacking, and data privacy concerns such as location and destination tracking [41]. This was one of the only studies of AVs that discussed data privacy, and it was not one of the central research questions.

In a Kelley Blue Book study (n>2000, weighted to census figures), 62% of participants did not want a world with solely AVs, with resistance decreasing with age [26], a trend corroborated by other studies of autonomy and age [1]. While these results shed insight into consumer preferences, this study was potentially biased by the extremity of its scenario, presenting a world with only autonomous cars to participants who likely live in an environment with only human-driven cars. This resistance to self-driving cars has been reinforced by other studies without extremity bias [10, 35].

Not all research studies have corroborated these findings. A survey (n=1517) run by AlixPartners found that three-quarters of U.S. drivers would want a self-driving car during their daily commute [36], a much higher level of acceptance than other studies had found. AlixPartners claims that prior surveys injected bias by placing emphasis on worst-case scenarios and that theirs found a balance that mitigated this bias.

Existing studies focus on consumer perception within the context of AVs, rather than the general public who are impacted just by being in the vicinity of AVs. Very little work seeks to study public perception decoupled from the framing of eventual consumption of self-driving cars. One such study is the MIT Moral Machine. It presents scenarios that show moral dilemmas "where a driverless car must choose the lesser of two evils, such as killing two passengers or five pedestrians" [31]. That study concerns the potential impact on and comfort of those in close proximity to an AV, but focuses solely on ethical issues related to physical safety.

Another study by Sleeper et al. explored perceptions of and comfort with vehicle-based sensing and recording used for purposes such as automatic lane correction and adaptive braking and cruise control. That study used hypothetical scenarios to examine perceived comfort for people who indirectly interact with vehicle sensors, include bystanders and nearby drivers. The authors found that perceived comfort with vehicle sensors increased when the benefits of the vehicle sensing was clear, particularly when benefits were related to safety [42]. In contrast to that study, our study explores perceptions and comfort with networked autonomous vehicles capable of large-scale data collection and analysis.

The body of research exploring consumer perceptions of AVs does have a consensus in one area: there is reluctance among the public toward accepting self-driving cars and issues of trust need to be addressed [47]. The focus is on potential consumers, rather than the public; safety concerns, rather than privacy concerns; and on individual AVs rather than commercial fleets of networked vehicles. Our study hopes to fill these gaps in understanding, especially because deployment of a commercial fleet has preceded private ownership of fully autonomous vehicles.

2.2 Technological Capability

Autonomous vehicles require extensive data in order to operate effectively. Their sensors typically include: GPS for navigation; a wheel encoder for monitoring the movements of the car; radar on the front and rear bumpers for identifying traffic; a camera near the rear-view mirror for color identification; lane departure, read collision, and pedestrian alerts; and a spinning light detection and ranging (LiDAR) sensor on the roof used for generating a 3D map of the environment (Figure 1) [20, 38].

The cameras bring up the greatest privacy concerns, especially if captured information is aggregated and centrally stored. A conceptual analogy used by our pilot study participants is CCTV surveillance. Thirteen states forbid the use of CCTV surveillance and all states require proper notice [12]. There are two flaws in this comparison: (1) CCTV is intended for surveillance while the sensors on a car are intended for autonomous driving, and (2) unlike CCTV, which is confined to a set space, AVs could be on any public road at any time. A more apt analogy could be the dashboard camera or 'dash cam,' yet information collected by dash cams is unlikely to be stored and analyzed centrally. It is unclear whether comfort with either CCTV or dash cams would translate to comfort with information capture by commercial fleets of AVs.

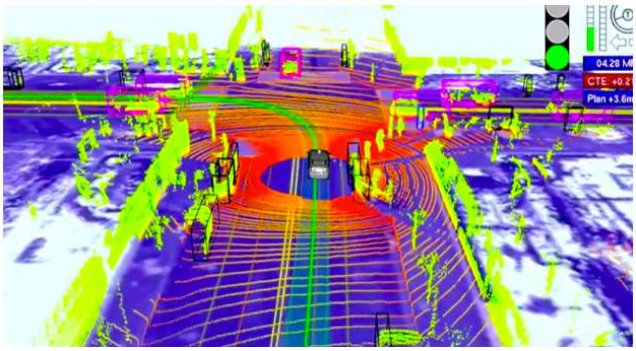

Figure 1: LiDAR Point Detection Heatmap

We suspect spinning LiDAR is the most foreign piece of sensing technology on an AV, for most people. LiDAR can be used for detecting and tracking objects; however, it is currently unable to identify individual people [38]. LiDAR data can potentially be combined with other data sources for concerning uses such as identifying how many people are at a protest.

Technological solutions aimed at mitigating AV privacy risks are not common, perhaps due to the lack of data surrounding consumer privacy preferences. Martinez-Balleste et al. describe ways to incorporate privacy-enhancing technology into the "smart city" by introducing the notion of "citizen

privacy," explained as the right of privacy for entire communities [30]. The researchers provide models aimed toward technologies that collect vast amounts of data in large, public settings; these models can be used to inform analysis of the privacy implications for AVs which can be thought of as an element of a smart city.

Self-driving cars can be seen as a new privacy invasive technology capable of always-on monitoring during operation, yet no notice is currently provided about how captured data will be used. Through our work, we gauge what the public deems acceptable, in an effort to inform the industry about what practices their potential customers and government stakeholders could want.

2.3 Regulation

Despite the tendency of U.S. law to react slowly to technological advances, the federal government has been convening stakeholders and developing regulatory principles for AVs. While there has been much discussion, there has been little movement on formal federal regulations and legislation. The Government Accountability Office (GAO) analyzed ten auto companies with regard to their location-data services and found that each had moved towards recommended privacy practices [49]. A year after the GAO report, the Alliance of Automobile Manufacturers submitted a statement to the Federal Trade Commission (FTC) with a commitment from the member companies to uphold privacy principles, specifically the traditional Fair Information Privacy Practices (FIPPs) [3]. In 2016, a National Highway Traffic Safety Administration report reiterated existing privacy stances by the government, such as notice and choice, desire to encourage competition in the realm of privacy, and the need to secure data [50].

Recently, state policymakers have taken steps to address AV privacy concerns [18]. The State of California passed a law that requires manufacturers of AVs to provide written disclosure of what data is collected [44], prompting backlash from the automotive industry [15]. As of 2016, 20 states have introduced autonomous car regulation, and since 2012, 34 states and D.C. have considered autonomous car legislation [33]. Eleven of these states have passed legislation, with two states using executive orders to mandate policy. While the California law is generally cited by the media, Michigan was highlighted as the first state to pass comprehensive AV regulations [4]. The legislation focused less on privacy constraints and instead legalized self-driving ride-sharing services, allowing for truck platoons, autonomous cars without drivers, and testing and usage on public roads [4]. The only major restriction, which states like Georgia, Maryland, Illinois, and Tennessee, have also introduced, is that the deployment of autonomous vehicles on public property is limited to automakers, requiring companies like Uber, Lyft, and Google to work with automakers in order to deploy vehicles [5].

At the federal level there is no binding legislation that addresses the privacy concerns associated with AVs. The Center for Internet and Society at Stanford maintains a wiki with current legislative and regulatory actions in the space of cyber law [43]; as of March 2017, the only enacted legislation with reference to AVs is the Fixing America Surface Transportation (FAST) Act. This legislation only instructs the GAO to "assess the status of autonomous transportation technology policy developed by U.S. public entities" [24]. Interestingly, the only other federal bill listed was the Autonomous Vehicle Privacy Protection Act of 2015. Unfortunately, this bill is still in committee deliberations by the House Transportation and Infrastructure Subcommittee on Highways and Transits [25]. The bill is not yet fully developed, only stating that the GAO needs to provide a public report assessing the ability of the Department of Transportation to address autonomous vehicle issues like consumer privacy—almost the same provision as in the FAST Act [25].

The closest regulations to data collection by the many sensors and cameras on an AV are those for dash cams. Legal authors Stitilis and Laurinaitis recognize that privacy is a huge concern with dash cams and hold that the benefits do not necessarily outweigh the harms. They relate back to the traditional view of privacy as the right to be left alone and cite EU laws that guarantee the right to privacy. Even with simple dash cam footage, it is difficult to balance priorities—cams are difficult to ban and people in public spaces do not have a reasonable expectation of privacy [45]. Deleting dash cam footage can be considered evidence tampering, which raises the question of if self-driving cars record information, would retention be necessary for legal compliance [23]? Despite the lack of uniform regulations, dash cams appear to have more privacy regulations than AVs at the state level, where some states prohibit recording when the owner is not driving and prohibit using them to surreptitiously record audio while being hidden from plain sight [23]. Publication of the collected data involves separate regulation and public perception, though cases generally involve simple uses such as determining the cause of an accident [21].

Another precedent regarding pedestrian privacy is the controversy surrounding Google Street View. The Street View technology was met with substantial scrutiny, with accusations about failing to properly blur faces and collecting excessive data, such as Wi-Fi signals [46]. Despite its use in public spaces the use of automated technology to collect data about people and their behaviors prompted considerable anxiety and response from the company [11].

Greenblatt asserts that the law has not prepared for the emergence of self-driving cars and will not be ready [19]. Given the deployment of AVs and lack of federal legislation, along with a mixed response from the states, Greenblatt appears to be right. Much like the rest of the Internet of Things, technology has outpaced the law, which, especially in the realm of privacy and security, has led to deficiencies that have damaged consumer trust in IoT devices [14]. If AVs follow the same direction as IoT devices have, the trust the public has in self-driving cars could be damaged by a major privacy or security breach—hampering the adoption of the technology and potentially inviting unwanted regulation. Our work hopes to provide the industry with guidance for crafting privacy protections that build trust rather than break it.

Journalists have investigated the extent of data collection and tracking features in high-tech cars, with mixed results. Articles have speculated car companies collect more than they say [6]. Companies are quick to respond, but often do not assuage privacy concerns or disclose data collection pro-

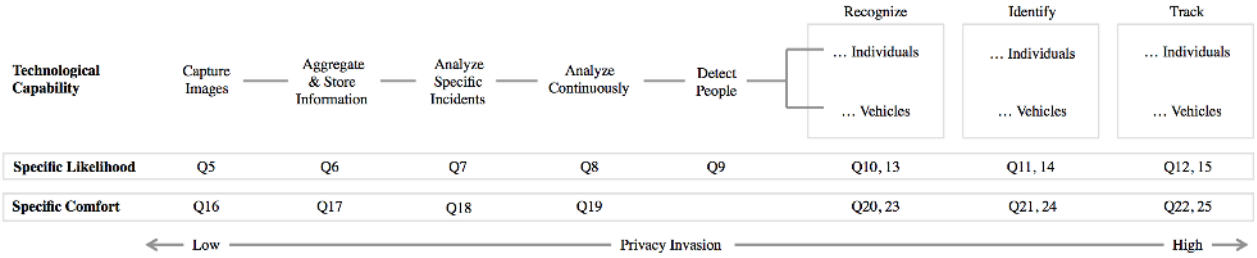

Figure 2: Privacy Scenarios

cedures [32, 22]. Policies enacted by ride-sharing companies have become the standard for the industry, a self-regulatory approach that currently defines most of U.S. privacy law [22]. Automotive data collection, in AVs or otherwise, presents a new set of privacy challenges for the industry. Companies such as BMW "[have] been inundated with requests from advertisers and technology companies to get their hands on vehicle data" [37]. The potential uses of and inferences from vehicle data by advertisers are extensive, even more so for AVs which necessarily collect a greater amount of data.

There are few legal protections for pedestrians and drivers against the capture and use of images taken by AVs. Even when their images are taken without their knowledge and consent, the current legal protection of tort law is limited when the likenesses of pedestrians are captured in photographs in a public environment, such as a city street [23].

3. METHODOLOGY

Our work seeks to understand public perceptions surrounding autonomous vehicles from a privacy standpoint, focusing on potentially privacy invasive capabilities of commercial AV fleets. We designed an online survey to explore people's conceptions of the sensing and analysis capabilities of self-driving vehicles; their comfort with the different capabilities; and the effort, if any, to which they would be willing to go to opt out of data collection.

3.1 Survey Questionnaire

As the first study explicitly investigating privacy conceptions surrounding networked fleets of AVs, an exploratory survey was chosen as the research method with the goal of identifying what consumers think is reasonable for AVs to do.

Participants were asked to focus on a fleet of self-driving cars operated by a single company that shares information with each other as well as the company, rather than single individually owned cars which have different capabilities and associated concerns. Only sensors on the outside of the car were to be considered, not any within the vehicle or any corresponding mobile application, to limit unknown effects. As a quality check, participants chose either "Yes, I understand" or "No, I did not read the short text" to move on to the next section.

The survey structure split participants into two groups to control for the priming effect of privacy questions. The Primed group received the set of scenario questions represented visually in Figure 2 to investigate conceptions of the sensing and analysis capabilities of self-driving vehicles (Specific Likelihood questions, Q16-25) as well as a set to gauge comfort level with the technological capabilities (Specific Comfort questions, Q5-15). The Unprimed group skipped these two sections and began with a set of General Comfort questions (Q26-35), which are represented visually in Figure 3. Two scenarios in this set concerned privacy. Eight other scenarios were included to both obfuscate the privacy questions and to facilitate comparison of discomfort due to privacy reasons with discomfort due to other aspects of the technology (e.g. safety or job market concerns). Both the Primed group and Unprimed group answered the General Comfort questions, the latter responding after answering the two specific question sets.

All scenario questions were piloted iteratively and discussed with pilot participants, who fell into one of four groups: non-technical, university students, security and/or privacy students, and robotics students. Pilots with the latter two groups developed the content and validity of scenarios to accurately fit the technology and accomplish research goals. Additional pilots were done to increase understanding of the scenarios. A small-scale pilot (n=41) using online recruiting was run to gather preliminary data, then final minor edits were made using data from these responses.

Specific Likelihood Questions

Participants were asked to answer questions about their conceptions of the current technological capabilities of AV fleets. These questions were designed to identify what people thought AV fleets were already doing. Participants rated different scenarios on a five point Likert scale from "Strongly Disagree" to "Strongly Agree." The scenarios began with those we assessed as least privacy invasive (i.e. image capture) and increased in invasiveness to scenarios involving recognition, identification, and tracking of people and vehicles. To help participants understand complex privacy concepts, examples were provided using the context of the Uber self-driving car fleet. Scenarios and examples can be found in Appendix A Q5-15 and are demonstrated visually in Figure 2.

Specific Comfort Questions

After the likelihood questions, participants in the Primed group indicated their comfort level with the same technological capabilities on a five point Likert scale from "Very Uncomfortable" to "Very Comfortable." While the Specific Likelihood questions measured what participants thought was realistically occurring, the Specific Comfort questions

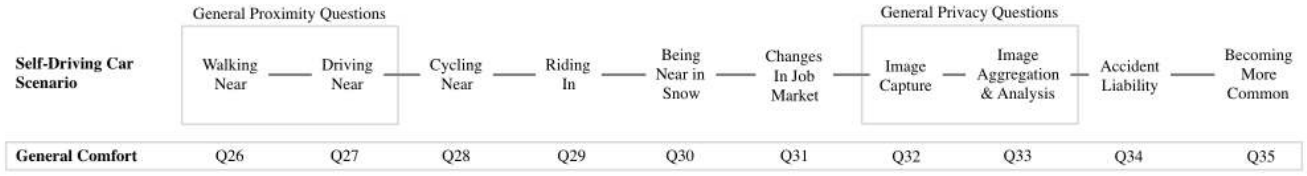

Figure 3: General Scenarios

asked participants how comfortable they would be if the technological capability was realized. By using the same capability scenarios, the relationship between conceptions of likelihood and comfort could be explored, as well as the difference in comfort at different levels of technological capability. Explanatory examples for these questions encouraged participants to imagine the technological capability in an every-day scenario and can be found in Appendix A Q16-25.

General Comfort with Self-Driving Cars

The General Comfort questions (Q26-34) covered general concerns with AVs identified by Schoettle and Sivak [41], including walking and driving near AVs, changes in the job market, and legal liability in an accident (Figure 3). Within the General Comfort set were two privacy questions (Q32-33), one concerning image capture and the other concerning aggregation and analysis of captured images. Using this set of scenarios, the effect of being primed with the likelihood and comfort scenarios could be measured for both general comfort with fleets of AVs and privacy concerns.

Quantification of Effort to Opt-out

The scenario-based questions measured conceptions and attitudes about AV technology, not behavior. To investigate whether discomfort would lead to action, we asked the following question:

> Q36. Suppose the company operating the fleet of self-driving cars has implemented a system so pedestrians and drivers can opt out of data collection by the cars. By going through their online system, people can have images of them blurred so their identity is protected and they cannot be tracked. How many minutes would you spend in the system to successfully opt out?

Response options were grouped into five-minute buckets for times between one and thirty minutes with options for "0 minutes" and "More than 30 minutes."

Exposure and Bias

We investigated the effects of using Uber as the example for our study, fearing that Uber as an example would bias results due to the many news stories circulating about the company during data collection and their strong effect on public opinion [7]. However, feedback from pilot participants indicated that even if Uber had not been used as the example, many participants would have still used the Uber self-driving car fleet as a mental model. Using Uber consistently kept some participants from using it as a mental model when others did not, which made the biasing effects easier to detect and measure. Additionally, using AVs that were already deployed in public spaces and familiar to many of our participants' made it more likely that they would be able to accurately envision and have developed opinions about the scenarios that we cover in our survey.

To measure the bias created by the use of Uber as an example, we asked participants to express their agreement with five statements on a five point Likert scale from "Strongly Disagree" to "Strongly Agree." The questions (Q36-40) assessed topics such as whether they would have answered the questions differently if Uber had not been used as an example. In addition to the bias questions, exposure to the technology and interaction with Uber were measured (Q45). Exposure questions included whether participants had read an article about Uber self-driving cars or ridden in one; interaction questions included whether participants used the Uber app or had protested against Uber.

Participant Characteristics

To further understand participants and the role characteristics play in their conceptions of networked AVs, demographic information was collected including gender, age, educational experience, and industry. Technical experience and general privacy attitudes were also recorded, the latter using the IUIPC question framework [29]. Email addresses were only collected to distribute compensation.

3.2 Recruitment

Participants were recruited from five cities of similar size and demographics: Pittsburgh, PA; Cleveland, OH; Cincinnati, OH; Rochester, NY; and Buffalo, NY. Participants were recruited in all five cities using local Craigslist ads and posts on city-specific Reddit forums. Posters were also used to recruit in six major central neighborhoods of Pittsburgh. Multiple methods were used to avoid bias from any one type of respondent and participants outside the specified cities were disqualified. Tracking of recruitment method was done via unique survey links. Participants who finished the survey could choose to give their email address to be entered into a random drawing for one of six $50 Amazon Gift Cards. The survey was run for two weeks beginning February 16 and closing March 3, 2017.

3.3 Analysis

We performed hypothesis tests to understand the relationship between participants' perceptions of likelihood and comfort with AV technological capabilities. We test the correla-

tion between participants' perceived likelihood and comfort with specific self-driving car capabilities using Spearman's ρ. To understand whether perceived likelihood ratings differed between person- and vehicle-specific capabilities, as well as how these ratings differed between different groups of participants, we binned likelihood ratings into {likely, very likely} and {very unlikely, unlikely, neither likely nor unlikely} and use Fisher's exact test. Comfort ratings were similarly tested using {uncomfortable, very uncomfortable} and {very uncomfortable, uncomfortable, neither uncomfortable nor comfortable} bins. In addition, we tested whether participants' specified opt-out minutes differed between participant segments using the Mann-Whitney U test.

All hypothesis tests used a significance level of $\alpha = 0.05$. For general self-driving comfort ratings, opt-out minutes, and comfort with specific AV capabilities, we performed exploratory testing with respect to many variables. To account for this, we applied the Holm-Bonferonni method within each family of tests and report corrected p-values.

4. RESULTS

Of the 312 survey responses, 248 gave complete responses and ten were excluded. Participants were excluded for failing the attention-check question (two participants), entering a location outside the scope of the study (one), or because they were Uber employees (seven). These last were excluded due to concerns about the lack of generalizability from their data to other populations. Additionally, multiple Uber employees seemed to be taking the survey only to see what the questions were, as they chose the neutral option for every Likert question and did not enter an email address for the gift card raffle.

Our sample was slightly skewed by the recruitment methods. Over half of participants (55%) were recruited via Reddit, which led to the sample being more male, technically experienced, and younger than the general population, due to the demographics of Reddit users [9]. Of the participants who answered demographic questions, 61% identified as male. The average age was 34 years, ranging from 18 to 79, and 24% were majoring in or had a degree or job in computer science, computer engineering, information technology, or a related field. The sample was more well-educated than the population with 13 with professional or doctoral degrees (5%), 45 with masters degrees (18%), 108 with bachelors degrees (43%), 16 with an associates degree, 49 with some college experience (19%), and 21 participants who had no college experience (8%). Based on the IUIPC privacy questions, the overwhelming majority of participants had strong beliefs concerning their own privacy. It should be noted though, that these questions were given at the end of the survey which had already raised many privacy concerns and could have increased participants' privacy sentiments.

Participants were randomly assigned to either the Primed or Unprimed group. The Primed group had 158 (52%) participants and the Unprimed group had 144 (48%). Of the five recruitment locations, the largest sample came from Pittsburgh (200, 68%), followed by Cleveland (63, 21%).

4.1 Exposure and Bias

Participants indicated their experience with Uber's AV technology in the survey by checking any of the fourteen statements that applied to them, seen in Figure 4. Statements

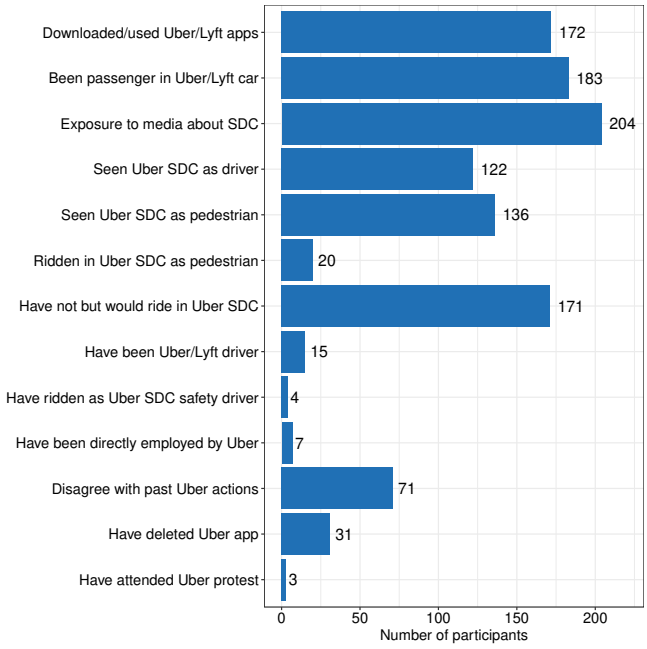

Figure 4: Participant exposure to Uber and self-driving cars (SDC).

covered not only exposure to self-driving cars, but also exposure to ride sharing technology, attitudes towards self-driving cars, attitudes towards Uber, behaviors indicative of negative opinions of Uber, and employment status by Uber or another ride-sharing service. Participants who self-identified as Uber employees (excluding drivers) are included in Figure 4 but were excluded from all other analyses; there were seven Uber employees in the study including four who had ridden as "safety drivers" in Uber self-driving cars.

Participants recruited from Pittsburgh had higher rates of exposure to ride sharing and self-driving technology in all areas. Notably, 78% of Pittsburgh participants and 42% of non-Pittsburgh participants had read an article, viewed a program, or learned online about Uber self-driving vehicles, indicating a high level of exposure to information about Uber's AV technology prior to this study. Seven percent of Pittsburgh participants had already ridden in a self-driving car and 64% had seen one as a pedestrian, compared to <1% and 3% of non-Pittsburgh participants, respectively. These results suggest that Pittsburgh residents generally have high exposure to the technology itself while residents of cities without the Uber self-driving car fleet have little to no exposure, though there may be some response bias where Pittsburghers with greater exposure were more likely to take our survey.

Negative attitudes towards Uber and associated behaviors were also prevalent in our sample, which we tested because of the many public controversies associated with the company. Twenty-three percent of all participants disagreed with actions Uber had taken and 10% had deleted the Uber mobile app. Three participants (<1%) had participated in protests against Uber. Most importantly, due to biases or preconceived notions, 14% of participants agreed and 3% strongly agreed that if Uber had not been used as an example they

would have answered the survey questions differently. Furthermore, 18% would trust another self-driving car company over Uber to have their best interests in mind, indicating that some of the distrust is company-related, not directly related to AV technology.

Since Uber is the most visible company currently operating networked fleets of autonomous robots in public spaces and despite its controversies, it was logical to use Uber as an example in scenarios. We decided that the ecological validity and use of a single mental model outweighed incurred bias. As one participant added in the free-text response, "I would have automatically used Uber in my own mind as an example." If some participants had used Uber as their mental model, while others used Google or Tesla, interpretation of our results would be more difficult. Consistent use of Uber as an example standardized the context for all participants and allowed us to ask participants about a technology that was already deployed in their city or cities like theirs.

4.2 Conceptions of Technological Capabilities

The trend in ratings of likelihood was inversely related to how privacy invasive the Specific Likelihood Question was, as ranked by researchers and shown in Figure 5.

Participants overwhelmingly rated basic capabilities such as image capture and aggregated storage as likely to be occurring, 87% and 91% respectively. Detection of humans was rated as likely by a similar proportion, at 88%. Under the assumption that images were already captured and stored, 94% of participants thought analysis for specific incidents, such as traffic accidents, was likely and 88% thought it was likely information was analyzed continuously for general tasks such as navigation. These are primary uses that directly impact the function of AVs. We found a clear division in ratings of capability between primary and secondary uses, where secondary uses are uses not necessary for the primary function of the AV. The secondary uses we explored are identification, recognition, and tracking of individuals and vehicles. Participants found primary uses to be highly likely, yet no more than half of participants rated each secondary use scenario as likely. Likelihood ratings for secondary uses are summarized in Table 1. Notably, the scenario that received the lowest likelihood rating by participants was also one of the most privacy invasive as ranked by coauthors: identification of individuals at 22%.

Overall there was a clear delineation in ratings of likelihood between primary and secondary use scenarios. Due to lack of information about the capabilities Uber self-driving cars actually have, only two scenarios are known to be occurring: image capture and detection of people. Almost 9 out of 10 participants accurately thought these verifiable scenarios were likely, as expected. A substantial minority of participants, no fewer than 1 out of 5, believed that even the most privacy invasive scenarios were likely to be occurring. While most participants held that primary uses were likely and secondary uses were not, many thought that the AV technology was being used to the extent of its capability in extremely privacy invasive ways, such as identifying pedestrians.

4.3 Comfort and Privacy Preferences

Discomfort level with each of the Specific Comfort questions (Q16-25) was quantified using the proportion of participants who chose "Uncomfortable" or "Very Uncomfortable." Participants were generally more comfortable with primary uses than with secondary uses. Discomfort was lowest for the least privacy invasive scenario (image capture, 16%) and highest for one of the most privacy invasive scenarios (tracking of vehicles, 85%). Generally high levels of discomfort were seen with: image storage (42%), analysis of specific incidents (36%), and continuous analysis (43%). The example used for the incident analysis scenario was Uber reviewing images captured of an accident, which could have explained why the associated discomfort was lower; as P95 noted in her free response, "If I have an accident with a driverless car, the recording is something useful, but that in my opinion should be the only reason the recordings/information should be released." Participants could have viewed this scenario as similar to dash cameras, which have known benefits and accepted norms of behavior. Of the secondary use scenarios, more than half of participants were uncomfortable with every scenario except vehicle recognition (43%), which was notably also the scenario rated most likely.

Scenario	Individuals	Vehicles
Recognition	38% (53)	46% (64)
Identification	22% (31)	28% (38)
Tracking	42% (58)	34% (47)

Table 1: Perceived likelihood of secondary use scenarios. The percentage (count) of participants that saw a scenario as likely or very likely are shown.

Comfort levels tended to decrease as questions increased in privacy invasiveness. The proportion of participants uncomfortable with aggregated storage was statistically significantly greater than with just image collection (Fisher's Exact Test, 42% vs. 15%, $p<0.001$). For secondary use scenarios—recognition, identification, and tracking—participants were more comfortable with recognition than identification or tracking. In particular, participants expressed higher discomfort with tracking of vehicles than identification of vehicles (85% vs. 71%, $p=0.040$) and higher discomfort with identification than recognition for both vehicles and individuals (71% vs. 43% for vehicles, 76% vs. 54% for individuals, $p<0.002$ for both). Notably, we did not observe statistically significant differences in comfort between continuous analysis and analysis of specific events (43% vs. 36%) nor between identification and tracking of individuals (both 76%). We also did not observe statistically significant differences in comfort for the three secondary use scenarios between individuals and vehicles.

4.3.1 Relationship Between Likelihood and Comfort

We also investigated whether rating a given capability scenario as likely was correlated with comfort with that same scenario. We found that there was a statistically significant positive correlation between likelihood and comfort ratings for identification (Spearman's $\rho=0.28$, $p=0.001$) and tracking ($\rho=0.17$, $p=0.049$) of individuals; and recognition ($\rho=0.19$, $p=0.028$), identification ($\rho=0.30$, $p<0.001$), and tracking ($\rho=0.22$, $p=0.019$) of cars. Likelihood and comfort ratings correlated most strongly for secondary use scenarios involving identification.

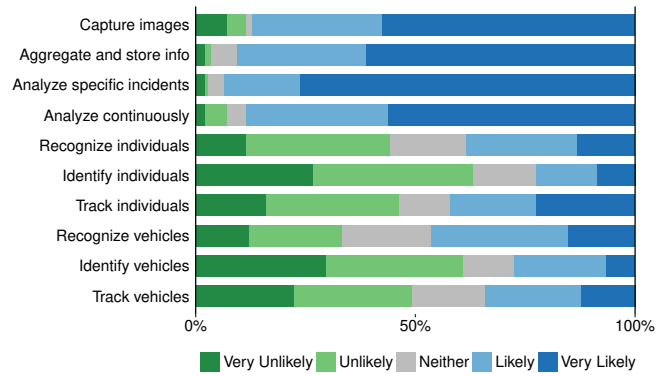

Figure 5: Likelihood ratings.

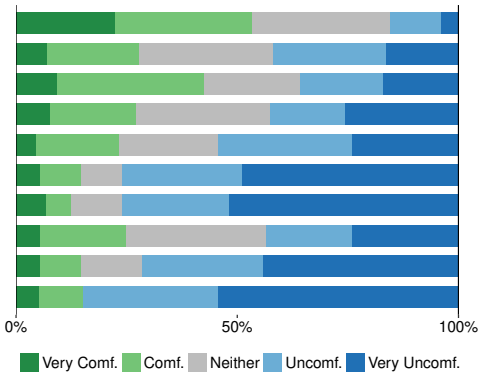

Figure 6: Comfort ratings.

Scenario	Overall	PGH	Non-PGH
Capture images	16% (20)	14% (13)	19% (7)
Aggregate and store info	42% (54)	43% (40)	38% (14)
Analyze specific incidents	36% (46)	36% (33)	35% (13)
Analyze continuously	43% (55)	39% (36)	51% (19)
Recognize individuals	54% (70)	57% (52)	49% (18)
Identify individuals	76% (98)	75% (69)	78% (29)
Track individuals	76% (98)	78% (72)	70% (26)
Recognize vehicles	43% (56)	46% (42)	38% (14)
Identify vehicles	71% (92)	68% (63)	78% (29)
Track vehicles	85% (95)	84% (67)	88% (28)

Table 2: Discomfort with technological capabilities in different scenarios, overall and by whether participants lived in Pittsburgh. The percentage (count) of participants that were uncomfortable or very uncomfortable with a scenario are shown.

Based on Cohen's guidelines we can interpret the strength of these observed relationships between likelihood and comfort ratings [8]. We observed moderate and near-moderate positive correlations between likelihood and comfort for vehicle identification and individual identification, respectively. Among participants that thought vehicle identification was likely, 59% were uncomfortable with this capability. In contrast, among those that thought vehicle identification was not likely, 76% were uncomfortable.

For the remaining scenarios, we observed small to moderate effect sizes (from $\rho = 0.17$ for individual tracking, to $\rho = 0.22$ for vehicle tracking). In all cases the direction of the correlation is positive; increased likelihood ratings were associated with increased comfort ratings.

We were surprised to find positive correlations between likelihood conceptions and comfort. Participants were more comfortable with a capability if they thought it was likely happening. It was expected that participants who thought a particular capability was already occurring would be more uncomfortable with it because they would feel more pressing concern with a technology that is already in use. Instead the opposite was observed, with higher ratings of likelihood related to higher levels of comfort.

4.3.2 Other Factors Related to Comfort

A number of additional factors were tested for effect on discomfort. Rather than use the Specific Comfort questions presented only to the Primed group, this analysis used the General Comfort questions shown to all participants as the measure of comfort with self-driving cars. In this exploratory analysis the intention was to uncover variables that could explain what made participants more or less comfortable with basic privacy invasive capabilities of AVs, namely image capture and analysis (Figure 3, Q32-33).

To obfuscate and compare these privacy questions, other concerns with AVs were also measured. Four scenarios in the General Comfort questions (Figure 3, Q26-35) concerned discomfort with proximity to a self-driving car: walking near (24% expressed discomfort), driving near (25%), cycling near (49%), and being near one in the snow (61%). Other causes of discomfort identified by Schoettle and Sivak's large-sample survey [41] were also explored. Forty-four percent of participants were uncomfortable with changes in the job market due to self-driving cars and 72% were uncomfortable with legal liability resulting from an accident with a self-driving car. We also asked one general question about how comfortable participants felt about self-driving cars becoming more common, to which 30% indicated discomfort.

The privacy scenarios (Q32-33) in the General Comfort questions made more participants feel uncomfortable than any other scenario: 85% were uncomfortable with image capture of people and license plates, and 77% were uncomfortable with that data being aggregated and analyzed. It was surprising that more participants felt uncomfortable with image capture than data aggregation and analysis, but a possible explanation is that participants are more uncomfortable with the fact that data is collected than potential uses of those data. Participants in pilot studies had difficulty articulating negative outcomes from image analysis, which could also explain the observations.

Using the Mann-Whitney U test, the following variables were investigated for a difference in discomfort: priming with privacy scenarios, Pittsburgh residence, gender, technical experience, and bias against Uber. As an example, we tested whether comfort with image capture statistically significantly differed between Pittsburgh and non-Pittsburgh participants. The Kruskal-Wallis test was used to compare discomfort between overall conceptions of likelihood and age range. Overall conceptions of likelihood were quantified as a numerical score (0-11) representing the number of scenar-

ios from the Specific Likelihood questions a participant had found likely. Surprisingly, none of the explanatory variables we explored had a statistically significant impact on discomfort.

To determine if these variables did not explain differences in privacy comfort, or if they did not explain any difference in comfort, the same tests were run on comfort with proximity to self-driving cars and comfort with these cars becoming more common in general. Discomfort with proximity as a driver and proximity as a pedestrian were quantified for each participant as whether they were uncomfortable with both, one, or neither of the scenarios (0-2).

We found statistically significant differences in discomfort with proximity to self-driving cars between participants with different discomfort levels for the Specific Comfort scenarios (Kruskal-Wallis, $\chi^2(10) = 42.28$, $p<0.001$). Similarly, we found statistically significant differences in discomfort with self-driving car technology becoming more common between discomfort levels with specific scenarios ($\chi^2(10) = 35.32$, $p<0.001$). In both cases, however, we did not observe a clear trend relating discomfort (with proximity or with the technology becoming more common) and overall discomfort with specific scenarios.

Whether participants had technical experience explained statistically significant differences in comfort with self-driving car technology becoming more common in general (Mann-Whitney $U = 4506.5$, $p=0.049$). Technical experience—studying or employed in computer science, computer engineering, information technology, or related—was related to increased comfort with the technology becoming more common (technical: 17% uncomfortable, non-technical: 34% uncomfortable), but it did not explain comfort with either privacy-related scenario (image capture or analysis).

The survey did not ask participants why they were uncomfortable with any specific scenario, so it is possible that the reason participants expressed discomfort with proximity is in fact because of privacy invasion and not for safety reasons. In this case having higher concern with proximity could be explained by discomfort with the sensors, not the possibility of being endangered, which is not corroborated by the dominance of safety in public discourse surrounding the technology. It is also possible that the lack of statistical significance for the two privacy questions within the General Comfort questions set could be due to a high baseline discomfort level.

4.3.3 Indications of Opt-Out Behavior

The set of explanatory variables described in the previous section were investigated for their effect on how long participants were willing to spend in an online system in order to opt out of identifiable data collection. Nine percent of participants would not use the online system, 37% would spend 5 minutes or fewer, 22% would spend 6-10 minutes, 20% would spend 11-30 minutes, and 12% would spend more than half an hour. Priming with the specific scenario questions was the only variable for which we observed statistically significantly different opt-out times (Mann-Whitney $U = 9847.5$, $p=0.022$), with opt out times higher for the Primed group (primed median: 6-10 minutes, non-primed median: 1-5 minutes).

This difference can be partially explained by the open text responses participants chose to give at the end of the survey. Four thoughtful responses discussed the opt out question specifically, three of whom disagreed with the idea of opting out, arguing instead that people should opt in or simply not have identifiable information captured. These responses showed nuanced thought about the nature of the technology and privacy implications which another participant (P91) noted had "raised issues [she] had never even considered." The nature of the scenario questions given to the Primed group presented scenarios and privacy implications that pilot study participants said they had not thought of before the study. Simply posing questions about potential privacy invasive scenarios increased the amount of time participants would spend to mitigate such invasions. It also shows that when the public is made aware of potential privacy invasions without accurate information about actual data collection and use practices, there is an increase in privacy-seeking behavior.

5. DISCUSSION

This study explored a previously unknown space: technological and privacy perceptions surrounding networked AVs, specifically the Uber commercial fleet of self-driving vehicles. We identified what technological capabilities the public ascribed to fleets of self-driving cars, how comfortable they were with those capabilities, and the effort to which they would go to protect themselves from privacy invasion. What we found was a complex space where perceived likelihood correlated with higher comfort, attributes that we thought would predict attitude and behavior had no observed effect, and simply asking questions about potential privacy scenarios increased participants' predictions of the time they would spend to opt out. Nevertheless, findings gleaned from this study can be used to recommend industry strategy and practices to assuage discomfort, protect privacy, and increase acceptance of this new technology.

5.1 Limitations

Sampling and recruitment bias could have played a role in our results. Participants came only from mid-sized cities in the Midwestern and Mid-Eastern regions of the United States, which limits the generalizability to more urban or rural populations as well as other nations. This limitation was the result of a conscious design choice: we specifically wanted to focus on people who had experience with fleets of AVs, which meant recruiting in Pittsburgh; then to compare opinions of people who were significantly less exposed to self-driving vehicles, we chose cities geographically near and demographically similar to Pittsburgh so as to avoid additional confounds. Future studies should diversify to more urban and rural areas, as well as to other cultures. Comparisons between exposed and unexposed populations should be available soon, as Uber deploys fleets in cities like San Francisco and more rural areas such as Michigan [4].

Another limitation of this study is the format used to conduct it. An online survey allowed us to reach over 300 people and learn about their conceptions of AV technology, but it was limited in depth. Many variables that could explain comfort and inform policy are as of yet unidentified and unexplored. More in-depth research could also assess what costs and benefits people think can come from the surveillance capabilities of networked fleets of autonomous robots.

5.2 Privacy Conceptions

Using the scenarios concerning technological capability (Figure 2) we learned what the public thinks self-driving cars currently do and how they feel about it. As expected, participants overwhelmingly (and correctly) believed that AVs have the capability to gather rich information about their environment and detect humans, as well as that AV fleets can perform off-line analyses of the collected information. The majority of participants generally thought secondary uses of collected information such as identification and tracking were not likely, though these scenarios still had a substantial minority of participants (22% to 46%) rating them as likely. As expected, comfort significantly decreased as scenarios became more invasive and a division was found between primary and secondary uses. Secondary uses were differentiated by participants in pilots and free-text responses by their degree of necessity and invasion: the invasion was often found to be needlessly 'too far', whereas primary uses could be rationalized.

Surprisingly, for the secondary use scenarios, rather than higher conceptions of likelihood correlating with higher discomfort, we observed the opposite. Participants who rated a potentially privacy-invasive scenario as likely were more likely to be comfortable with that scenario; this might be explained by learned helplessness or resignment to perceived inevitability. Learned helplessness is when in negative situations where an individual has no ability to change the circumstances, such as the invasion of privacy by autonomous vehicles, people increasingly accept the situation as a coping mechanism. With no power to change the environmental factors that cause a negative response, the negative response itself is changed [54].

Similarly, if participants had perceived the technological capability as not only likely, but as normal or inevitable, this could have led to increased comfort. These findings support the need for research and privacy enhancing technologies and policies early in the technology's life cycle. As people become resigned over time, the deployment of AV technology may outpace restrictions, as previously mentioned in reference to IoT technology, making it harder to integrate privacy protections.

Causes of Discomfort

Though a participant's perceived likelihood of a particular scenario explained her comfort with that scenario, other expected explanatory variables did not. None of the explanatory variables tested explained any difference in comfort with AV image capture and analysis (Q32-33). In contrast, greater technical experience was associated with increased comfort with self-driving cars becoming more common in general. We expected that technical experience would have one of two potential effects: greater knowledge leading to a better understanding of potential negative impacts and consequences and hence more concern; or, alternatively, better understanding of the benefits and hence less concern. Support for the latter was found, but only for comfort with AV technology in general, not for comfort with privacy scenarios, where technical experience had no observable effect. A possible explanation is that comfort with AV technology in general is derived mainly from safety and employment concerns, rather than privacy concerns.

We expect that proximity concern is a combination of privacy concerns and safety concerns, with significantly greater weight given to safety than privacy based on the narrative of public discourse, open-text responses of participants, and the phrasing of the questions. In this case, privacy discomfort could be indicative of safety discomfort for other reasons, such as that they are both caused by an innate distrust of the technology. More nuanced exploration would be needed to answer these questions, perhaps via interview studies.

Time to Opt Out

Though high levels of discomfort with the different technological capabilities were found, half of participants would spend only five or fewer minutes using an online system to opt out of identifiable data collection by commercial autonomous vehicles. The only factor that explained a difference in opt-out time was whether the participant had been primed with specific privacy scenarios. Presenting people with scenarios that suggested the possibility of privacy invasion made people predict that they would spend more effort mitigating the privacy invasion. No other variables, including exposure to self-driving cars or bias against Uber, explained a difference in time to opt out.

Should the public be exposed to questions regarding privacy invasive capabilities, there could be an increased move towards privacy-seeking behavior such as opting out or perhaps protesting. Research and media attention is currently focused on safety and employment, but more of our participants were uncomfortable with privacy invasive capabilities than with either of these popular concerns. Even participants in the Unprimed group, who did not see questions regarding recognition, identification, and tracking, were more likely to be uncomfortable with privacy scenarios than with proximity scenarios. If public attention were to shift towards the third ethical concern—privacy—findings in this study indicate that discussions would reveal great discomfort and the act of discussing such concerns could cause a change in behavior concerning commercial self-driving vehicles.

5.3 Recommendations for Industry Practice

One of the central questions investigated by this study was where the public draws the line on acceptable and unacceptable privacy practices by companies operating networked autonomous vehicles in public spaces. The sentiments of participants tended toward acceptance of technologies they thought were being implemented as necessary components, but toward discomfort with secondary analysis of information such as recognition, identification, and tracking of people or vehicles. Additionally, participants would overwhelmingly use a system to opt out of identifiable information capture, though some expressed that an opt-out tool is unsuited to the technology.

The synthesis of these findings shows that people, regardless of their exposure to AV technologies, are uncomfortable with privacy-invasive secondary uses and, to a lesser extent, with primary uses such as continuous analysis of data captured by networked AVs. The only secondary use that could potentially be considered useful and acceptable was recognition of vehicles, which participants rationalized could be useful for taking extra precautions against erratic drivers. With other

new technologies, the argument can be made that if the privacy intrusions conflict with individuals' preferences, they need not use that technology; but with sophisticated sensors operating in public places people have no practical ability to avoid information capture. It is then necessary that companies operating such fleets of AVs and other robots like drones either implement industry self-regulation or be regulated to protect the public. Our findings suggest that such regulation should focus on secondary data uses, with which the public is overwhelmingly uncomfortable and would actively avoid if given the opportunity.

Currently this regulation could take three forms: industry self-regulation, federal regulation, or state and local restrictions. The Alliance of Automotive Manufacturers has jurisdiction due to the necessity of autonomous vehicle companies partnering with traditional automotive companies [5] and this organization is committed to the Fair Information Privacy Practices (FIPPs) [2]. All of the foundational necessities are in place, but this organization has not yet applied them directly to AVs, or in particular to concerns raised by their external sensors. Federal regulation could take multiple forms; traditionally roads and cars are under the jurisdiction of the National Highway Transportation Safety Administration [34], though the FTC frequently crosses into other jurisdictions to enact privacy regulation. Both agencies support notice and choice, the first two FIPPs. Local and state governments are interfacing directly with these AV companies already though, and do require knowledge of their practices before allowing them access to public roads. These cities and states could set precedent for broader practice by working with the companies to create practices that balance the need for information with citizens' privacy. The companies themselves could create or adapt other privacy enhancing technologies such as face and license plate blurring, such as that done by Google Maps cars [16].

Additionally, it is in the best interest of companies operating AV fleets to be more transparent about their data collection and use practices. While the public has not yet considered the privacy implications of AV technology the way it has safety implications, this study found that bringing up privacy concerns causes people to be less comfortable with being near and utilizing self-driving car technology and to express intentions of actively mitigating privacy invasion. Such attitudes could cause increased backlash not only from the public, which has already been vocal about reservations about safety and employment, but from the city and state governments that are currently debating whether to allow autonomous vehicles to operate within their jurisdictions.

Safety concerns can be rebutted with the argument that the new technology (AVs) is less concerning than the current environment (human drivers), but companies like Uber cannot argue that data capture by networked autonomous vehicles is less concerning than the current environment where there are no networked vehicles capable of city-scale surveillance. Standard arguments for the technology are more difficult to apply and companies have yet to make a case for—or provide public services that—demonstrate data collection is net positive for the populations of the cities they operate in.

6. CONCLUSION

Our study investigated the largely unexplored space of privacy concerns surrounding autonomous vehicles. We found that participants generally thought networked fleets of autonomous vehicles were collecting and analyzing data about them, and that more than 40% thought this technology was already being used to track people's movements. Scenarios such as tracking and identification caused overwhelming discomfort, while participants expressed moderate discomfort with primary uses of data such as continuous analysis for navigation. If a participant thought a particular capability was likely to be occurring, she was more comfortable with that capability, perhaps because she thought it was normal or because she was resigned to it.

Surprisingly, privacy concerns caused higher proportions of participants to express discomfort than either of the more common concerns—physical proximity or changes in the job market. These feelings of discomfort with privacy-invasive capabilities were not explained by any of the variables we examined, indicating that attitudes were either too nuanced for detection by this study, were resistant to the effects of other variables, or were explained by unexplored additional factors. Interestingly, the amount of time participants predicted they would spend on privacy-protective behaviors was not as resistant: simply asking priming questions about autonomous vehicle capabilities increased participants' predictions of how long they would spend in an online system to opt out of identifiable data collection. Future studies can further investigate the relationship between priming, attitudes, and behaviors, and increase the understanding of privacy concern in this technological context.

Autonomous vehicle technology is set to become increasingly prevalent in the next decade and permanently alter daily life for millions of people [27]. Privacy research early in the development life cycle of this unique technology can be used to shape industry practices and regulation before intentional or unintentional privacy invasions become a part of the technology. It is important to investigate privacy implications of networked autonomous vehicles before deployment outpaces understanding of potential ramifications. We recommend policies differentiate between primary and secondary uses of sensor data, restricting secondary uses to preserve public privacy.

7. ACKNOWLEDGMENTS

This work was supported in part with a gift from Google. The authors would like to thank Matthew Bajzek and Aki Hitomi for help with early versions of this work.

8. REFERENCES

[1] H. Abraham, C. Lee, S. Brady, C. Fitzgerald, B. Mehler, B. Reimer, and J. F. Coughlin. Autonomous vehicles, trust, and driving alternatives: A survey of consumer preferences. Technical report, Massachusetts Institute of Technology, Cambridge, MA, 2016. http://agelab.mit.edu/files/publications/2016_6_Autonomous_Vehicles_Consumer_Preferences.pdf. Accessed March 2017.

[2] Alliance of Automobile Manufacturers. Automotive privacy, 2017. https://autoalliance.org/connected-cars/automotive-privacy-2/.

[3] Alliance of Automotive Manufacturers Inc. Letter to FTC, 2014. https://autoalliance.org/connected-

cars/automotive-privacy-2/letter-to-ftc/. Accessed October 2016.

[4] J. Bhuiyan. Michigan just became the first state to pass comprehensive self-driving regulations. *Recode*, Dec. 2016. https://www.recode.net/2016/12/9/13890080/michigan-dot-self-driving-cars-laws-automakers. Accessed March 2017.

[5] J. Bhuiyan. A series of U.S. state laws could prevent Uber or Google from operating self-driving cars. *Recode*, Feb. 2017. http://www.recode.net/2017/2/25/14738966/self-driving-laws-states-gm-car-makers. Accessed March 2017.

[6] J. M. Broder. That Tesla data: What it says and what it doesn't. *The New York Times*, 2013. http://wheels.blogs.nytimes.com/2013/02/14/that-tesla-data-what-it-says-and-what-it-doesnt. Accessed October 2016.

[7] B. Carson. Uber's unraveling: The stunning, 2 week string of blows that has upended the world's most valuable startup. *Business Insider*, Mar 2017. http://www.businessinsider.com/uber-scandal-recap-2017-3.

[8] J. Cohen. *Statistical Power Analysis for the Behavioral Sciences*. L. Erlbaum Associates, 1988.

[9] M. Duggan and A. Smith. 6% of online adults are Reddit users. *Pew Internet & American Life Project*, 2013. http://www.pewinternet.org/2013/07/03/6-of-online-adults-are-reddit-users/.

[10] Ernst & Young. Autonomous vehicles: How much do we need?, 2016. http://www.ey.com/gl/en/industries/automotive/ey-autonomous-vehicles-how-much-human-do-we-need. Accessed March 2017.

[11] A. Eustace. WiFi data collection: An update, 2010. https://googleblog.blogspot.com/2010/05/wifi-data-collection-update.html.

[12] R. Fogel. CCTV and video surveillance laws in US, Dec. 2011. http://www.smartsign.com/blog/cctv-laws-in-us/. Accessed March 2017.

[13] R. E. Freeman. *Strategic management: A stakeholder approach*. Cambridge University Press, 2010.

[14] FTC Staff Report. Internet of Things privacy & security in a connected world. Technical report, The Federal Trade Commission, Jan. 2015. https://www.ftc.gov/system/files/documents/reports/federal-trade-commission-staff-report-november-2013-workshop-entitled-internet-things-privacy/150127iotrpt.pdf. Accessed March 2017.

[15] M. Geuss. Automakers balk at California's proposed self-driving car rules. *Ars Technica*, Oct. 2016. http://arstechnica.com/cars/2016/10/automakers-balk-at-californias-proposed-self-driving-car-rules/. Accessed October 2016.

[16] D. J. Glancy. Privacy in autonomous vehicles. *Santa Clara L. Rev.*, 52:1171, 2012.

[17] J. Gluck, F. Schaub, A. Friedman, H. Habib, N. Sadeh, L. F. Cranor, and Y. Agarwal. How short is too short? Implications of length and framing on the effectiveness of privacy notices. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*, pages 321–340, Denver, CO, 2016. USENIX Association.

[18] E. Goodman. Self-driving cars: Overlooking data privacy is a car crash waiting to happen. *The Guardian*, Aug. 2016. https://www.theguardian.com/technology/2016/jun/08/self-driving-car-legislation-drones-data-security. Accessed October 2016.

[19] N. Greenblatt. Self-driving cars will be ready before our laws are. *IEEE Spectrum*, Jan. 2016. http://spectrum.ieee.org/transportation/advanced-cars/selfdriving-cars-will-be-ready-before-our-laws-are. Accessed March 2017.

[20] E. Guizzo. How Google's self-driving car works. *IEEE Spectrum*, Oct. 2011. http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/how-google-self-driving-car-works. Accessed March 2017.

[21] C. J. Junior, S. Muse, and C. Jung. Crowd analysis using computer vision techniques. *IEEE Signal Processing Magazine*, 27(5):66–67, 2010. http://ieeexplore.ieee.org/document/5562657/. Accessed October 2016.

[22] A. LaFrance. Driverless-car makers on privacy: Just trust us. *The Atlantic*, Mar. 2016. http://www.theatlantic.com/technology/archive/2016/03/self-driving-car-makers-on-privacy-just-trust-us/474903/. Accessed October 2016.

[23] S. Lehto. The surprising legal ramifications of having a dashcam in your car. *Road & Track*, Jan. 2017. http://www.roadandtrack.com/car-culture/a32124/the-surprising-legal-ramifications-of-having-a-dash-cam-in-your-car/. Accessed March 2017.

[24] Library of Congress. H.R 22 - FAST Act, Dec. 2015. https://www.congress.gov/bill/114th-congress/house-bill/22. Accessed March 2017.

[25] Library of Congress. H.R 3876 - Autonomous Vehicle Privacy Protection Act of 2015, Nov. 2015. https://www.congress.gov/bill/114th-congress/house-bill/3876. Accessed March 2017.

[26] T. Lien. Consumers aren't as excited as the auto industry about self-driving cars. *Los Angeles Times*, 2016. http://www.latimes.com/business/technology/la-fi-tn-kbb-self-driving-car-survey-20160927-snap-story.html. Accessed March 2017.

[27] T. Litman. Autonomous vehicle implementation predictions. *Victoria Transport Policy Institute*, 28, 2014. http://www.vtpi.org/avip.pdf.

[28] B. Liu, M. S. Andersen, F. Schaub, H. Almuhimedi, S. A. Zhang, N. Sadeh, Y. Agarwal, and A. Acquisti. Follow my recommendations: A personalized privacy assistant for mobile app permissions. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*, pages 27–41, Denver, CO, 2016. USENIX Association.

[29] N. K. Malhotra, S. S. Kim, and J. Agarwal. Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information systems research*, 15(4):336–355, 2004.

[30] A. Martinez-Balleste, P. Perez-martinez, and A. Solanas. The pursuit of citizens' privacy: A privacy-aware smart city is possible. *IEEE Communications Magazine*, 51(6):136–141, 2013. http://ieeexplore.ieee.org/document/6525606/. Accessed October 2016.

[31] Massachusetts Institute of Technology. MIT Moral Machine, July 2016. http://moralmachine.mit.edu. Accessed March 2017.

[32] E. Musk. A most peculiar test drive, 2013. https://www.tesla.com/blog/most-peculiar-test-drive. Accessed October 2016.

[33] National Conference of State Legislature. Autonomous vehicles - self-driving vehicles enacted legislation, Feb. 2017. http://www.ncsl.org/research/transportation/autonomous-vehicles-self-driving-vehicles-enacted-legislation.aspx. Accessed March 2017.

[34] National Highway Traffic Safety Administration. Laws administered by NHTSA, 2017. https://www.nhtsa.gov/laws-regulations/statutory-authorities.

[35] K. Naughton. Billions are being invested in a robot that Americans don't want. *Bloomberg*, May 2016. https://www.bloomberg.com/news/articles/2016-05-04/billions-are-being-invested-in-a-robot-that-americans-don-t-want. Accessed March 2017.

[36] K. Naughton. Three-quarters of U.S. drivers say they'd cede wheel to robot. *Bloomberg*, June 2016. https://www.bloomberg.com/news/articles/2016-06-30/three-quarters-of-u-s-drivers-say-they-d-cede-wheel-to-robot. Accessed March 2017.

[37] C. Neiger. Advertisers are begging car companies for your data. *The Motley Fool*, Jan. 2015. http://www.fool.com/investing/general/2015/01/25/advertisers-are-begging-car-companies-for-your-dat.aspx. Accessed October 2016.

[38] J. Petit. Self-driving and connected cars: Fooling sensors and tracking drivers, 2015. https://www.blackhat.com/docs/eu-15/materials/eu-15-Petit-Self-Driving-And-Connected-Cars-Fooling-Sensors-And-Tracking-Drivers.pdf. Accessed March 2017.

[39] N. M. Richards. The dangers of surveillance. *Harvard Law Review*, 126(7):1934–1965, 2013.

[40] F. Schaub, R. Balebako, A. L. Durity, and L. F. Cranor. A design space for effective privacy notices. In *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*, pages 1–17, Ottawa, 2015. USENIX Association.

[41] B. Schoettle and M. Sivak. A survey of public opinion about autonomous and self-driving vehicles in the US, the UK, and Australia. 2014.

[42] M. Sleeper, S. Schnorf, B. Kemler, and S. Consolvo. Attitudes toward vehicle-based sensing and recording. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '15, pages 1017–1028, New York, NY, USA, 2015. ACM.

[43] Stanford Center for Internet and Society. Automated driving: Legislative and regulatory action. http://cyberlaw.stanford.edu/wiki/index.php/. Last Updated Feb. 2015. Accessed March 2017.

[44] State of California. Autonomous vehicles in California, 2016. https://www.dmv.ca.gov/portal/dmv/detail/vr/autonomous/bkgd. Accessed October 2016.

[45] D. Stitilis and M. Laurinaitis. Legal regulation of the use of dashboard cameras: Aspects of privacy protection. *Computer Law and Security Review*, 32(4):316–326, Apr. 2016. http://www.sciencedirect.com/science/article/pii/S0267364916300267. Accessed October 2016.

[46] D. Streitfeld. Google concedes that drive-by prying violated privacy. *The New York Times*, Mar. 2013. http://www.nytimes.com/2013/03/13/technology/google-pays-fine-over-street-view-privacy-breach.html. Accessed October 2016.

[47] C. Tennant, S. Howard, B. Franks, and M. Bauer. Autonomous vehicles: Negotiating a place on the road. Technical report, London School of Economics, 2016. http://www.lse.ac.uk/website-archive/newsAndMedia/PDF/AVs-negociating-a-place-on-the-road-1110.pdf. Accessed March 2017.

[48] M. J. Thomas. *Combining facial recognition, automatic license plate readers and closed-circuit television to create an interstate identification system for wanted subjects*. PhD thesis, Monterey, California: Naval Postgraduate School, 2015.

[49] United States Government Accountability Office. In-car location-based services, Dec. 2013. http://www.gao.gov/assets/660/659509.pdf. Accessed October 2016.

[50] U.S. Department of Transportation National Highway Traffic Safety Administration. Federated automated vehicles policy, 2016. http://www.nhtsa.gov/nhtsa/av/index.html. Accessed October 2016.

[51] Y. Wang, P. G. Leon, A. Acquisti, L. F. Cranor, A. Forget, and N. Sadeh. A field trial of privacy nudges for Facebook. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 2367–2376. ACM, 2014.

[52] Wilson and Mitchell. Beyond Uber, Volvo and Ford: Other automakers' plans for self-driving vehicles. *Los Angeles Times*, Aug 2016. http://www.latimes.com/business/autos/la-fi-automakers-self-driving-20160819-snap-htmlstory.html.

[53] M. Wisniewski. Driverless cars could improve safety, but impact on jobs, transit questioned. *Chicago Tribune*, Jul. 2016. http://www.chicagotribune.com/news/ct-driverless-cars-getting-around-20160703-story.html.

[54] C. B. Wortman and J. W. Brehm. Responses to uncontrollable outcomes: An integration of reactance theory and the learned helplessness model. *Advances in experimental social psychology*, 8:277–336, 1975.

APPENDIX

A. SURVEY

1. What city do you live in?

◯ Pittsburgh, PA
◯ Rochester, NY

◯ Buffalo, NY
◯ Cincinnati, OH
◯ Cleveland, OH
◯ Other - Write In (Required)

2. How did you learn about this survey?
◯ Poster
◯ Craigslist
◯ Reddit
◯ Word of mouth
◯ Other - Write In (Required)

Please read survey information carefully. This survey explores opinions about fleets of self-driving cars. It is NOT intended to test or judge your knowledge of self-driving car technology.

For this survey suppose: (1) A fleet of self-driving cars is operated in your city (2) The cars are owned and operated by a private company (3) The cars are networked to share information with each other and the company

One example of this is the Uber self-driving car fleet currently operated in Pittsburgh, PA.

This survey is NOT about: (1) Individually owned self-driving cars (2) Any sensors on the inside of the car

3. Do you understand what this survey is and is not about?
◯ No, I didn't read the short text (please read)
◯ Yes, I understand

4. Likelihood of Self-Driving Car Scenarios [Primed group only]

You will be presented with scenarios about a networked fleet of self-driving cars. Choose how likely you think the scenarios are to be happening now from 'very unlikely' to 'very likely.' Please read each question carefully.

5. Self-driving cars capture images of their surroundings
◯ Very Unlikely
◯ Unlikely
◯ Neither Unlikely nor Likely
◯ Likely
◯ Very Likely

6. Information captured by the self-driving cars is aggregated and stored
For example, Uber stores data collected by all of its self-driving cars in a central location
◯ Very Unlikely
◯ Unlikely
◯ Neither Unlikely nor Likely
◯ Likely
◯ Very Likely

7. Information captured by a self-driving car during a specific incident is analyzed by the operating company
For example: Images captured by an Uber self-driving car during a car accident are used by Uber to determine the cause
◯ Very Unlikely
◯ Unlikely
◯ Neither Unlikely nor Likely
◯ Likely
◯ Very Likely

8. Information captured continuously by the self-driving cars is analyzed
For example: Data collected by all Uber self-driving cars is used by Uber to understand weather conditions
◯ Very Unlikely
◯ Unlikely
◯ Neither Unlikely nor Likely
◯ Likely
◯ Very Likely

9. Self-driving cars detect humans
◯ Very Unlikely
◯ Unlikely
◯ Neither Unlikely nor Likely
◯ Likely
◯ Very Likely

10. A self-driving car recognizes a person that has been encountered before by a different self-driving car in the fleet
◯ Very Unlikely
◯ Unlikely
◯ Neither Unlikely nor Likely
◯ Likely
◯ Very Likely

11. Individuals are identified by name when they encounter one of the self-driving cars in the fleet
For example: Uber knows that the pedestrian next to one of its self-driving cars is Alice
◯ Very Unlikely
◯ Unlikely
◯ Neither Unlikely nor Likely
◯ Likely
◯ Very Likely

12. Individuals are tracked using each time they encounter one of its self-driving cars in the fleet
◯ Very Unlikely
◯ Unlikely
◯ Neither Unlikely nor Likely
◯ Likely
◯ Very Likely

13. A self-driving car recognizes a vehicle that has been seen by another self-driving car in the fleet
For example: Uber knows that different self-driving cars encountered the same vehicle on different days, but does not know who owns the vehicle
◯ Very Unlikely
◯ Unlikely
◯ Neither Unlikely nor Likely
◯ Likely
◯ Very Likely

14. Vehicle owners are identified by name when a vehicle encounters one of the self-driving cars in the fleet
For example: Uber knows that the minivan in front of one of its self-driving cars is owned by Alice ◯ Very Unlikely
◯ Unlikely
◯ Neither Unlikely nor Likely
◯ Likely

◯ Very Likely

15. Vehicles are tracked using each time they encounter one of the self-driving cars in the fleet
For example: Uber assembles a list with location, date, and time of each time self-driving cars encountered Alice's minivan
◯ Very Unlikely
◯ Unlikely
◯ Neither Unlikely nor Likely
◯ Likely
◯ Very Likely

Comfort with Self-Driving Cars [Primed group only]

Choose how comfortable you are with the scenarios from 'very uncomfortable' to 'very comfortable.' Please read each question carefully.

16. I would feel _____ if self-driving cars captured images of me (but did not store or analyze those images.)
For example: An Uber self-driving car captures an image of you in a crosswalk, then discards the image after it leaves the intersection.
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

17. I would feel _____ if self-driving cars captured and stored images of me (but did not analyze those images)
For example: An Uber self-driving car captures an image of you in a crosswalk and it is stored on a computer with many similar images, but Uber does not use the images.
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

18. I would feel _____ if self-driving cars captured images of me and analyzed images of specific events
For example: Uber analyzes specific images captured by a self-driving car (including images of you) to determine the cause of a traffic incident.
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

19. I would feel _____ if self-driving cars captured images of me and analyzed images continuously
For example: Uber continuously analyzes images captured by all self-driving cars (including images of you) to gauge traffic conditions.
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

20. I would feel _____ if each time I encountered a self-driving car, I was recognized from past encounters with other self-driving cars (but not by name).
For example: Uber knows that different self-driving cars encountered you in different locations on different days, but does not know who you are
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

21. I would feel _____ if I was identified by images captured by a self-driving car
For example: An Uber self-driving car captures an image of your face as you cross the street and Uber links the image to your name
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

22. I would feel _____ if I was tracked each time I encountered a self-driving car.
For example: Uber assembles a list with location, date, and time of each time you encounter a self-driving car.
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

23. I would feel _____ if each time my car encountered a self-driving car, it was recognized from past encounters with other self-driving cars (but not by owner's name).
For example: Uber knows that different self-driving cars encountered your car in different locations on different days, but does not know who owns the car.
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

24. I would feel _____ if my car was identified by images captured by a self-driving car
For example: An Uber self-driving car captures an image of your license plate as you drive and Uber uses the links the license plate to your name
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

25. I would feel _____ if my car was tracked each time it encountered a self-driving car.
For example: Uber assembles a list with location, date, and time of each time your car encounters a self-driving car.
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

General Self-Driving Car Questions

You will be presented with scenarios about a networked fleet of self-driving cars. Choose how comfortable you are with the scenarios from 'very unlikely' to 'very likely.' Please read each question carefully.

26. I would feel _____ walking near a self-driving car.
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

27. I would feel _____ driving near a self-driving car.
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

28. I would feel _____ cycling near a self-driving car.
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

29. I would feel _____ riding in a self-driving car.
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

30. I would feel _____ being near a self-driving car in the snow.
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

31. I would feel _____ about the changes in the job market due to self-driving cars.
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

32. I would feel _____ if a self-driving car captured pictures of me and my license plate.
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

33. I would feel _____ if images captured by self-driving cars were aggregated and analyzed
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

34. I would feel _____ about legal liability in an accident with a self-driving car.
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

34. I would feel _____ about networked fleets of self-driving cars becoming more common in general.
◯ Very Uncomfortable
◯ Uncomfortable
◯ Neither Uncomfortable nor Comfortable
◯ Comfortable
◯ Very Comfortable

Opting Out of Information Capture

Suppose the company operating the fleet of self-driving cars has implemented a system so pedestrians and drivers can opt out of data collection by the cars. By going through their online system, people can have images of them blurred so their identity is protected and they cannot be tracked.

35. How many minutes would you spend in the system to successfully opt out?
◯ 0
◯ 1-5
◯ 6-10
◯ 11-15
◯ 16-20
◯ 21-25
◯ 26-30
◯ More than 30

Questions about Uber

36. I feel that companies operating networked fleets of self-driving cars have my best interests in mind
◯ Strongly disagree
◯ Disagree
◯ Neither agree nor disagree
◯ Agree
◯ Strongly agree

37. I feel that Uber's self-driving car division has my best interests in mind
◯ Strongly disagree
◯ Disagree
◯ Neither agree nor disagree
◯ Agree
◯ Strongly agree

38. I feel that Uber has my best interests in mind
◯ Strongly disagree
◯ Disagree
◯ Neither agree nor disagree
◯ Agree
◯ Strongly agree

39. I would have answered the survey questions differently had Uber not been used as the example
◯ Strongly disagree
◯ Disagree
◯ Neither agree nor disagree
◯ Agree
◯ Strongly agree

40. I would trust a different networked self-driving car fleet over Uber's to have my best interests in mind
◯ Strongly disagree
◯ Disagree
◯ Neither agree nor disagree
◯ Agree
◯ Strongly agree

Demographic Questions

41. Please specify your gender
◯ Man
◯ Woman
◯ Other (please specify):
◯ Prefer not to answer

42. Please indicate your age

[textbox]

43. Select the highest education level you have achieved
◯ No high school
◯ Some high school
◯ High school graduate, diploma, or the equivalent
◯ Trade, technical, or vocational training
◯ Some college
◯ Associate degree
◯ Bachelor's degree
◯ Master's degree
◯ Professional or doctoral degree
◯ Prefer not to answer

44. Select the industry in which you work
◯ Accounting
◯ Advertising
◯ Aerospace / Aviation / Automotive
◯ Agriculture / Forestry / Fishing
◯ Biotechnology
◯ Business / Professional Services
◯ Business Services (Hotels, Lodging Places)
◯ Computers (Hardware, Desktop Software)
◯ Communications
◯ Construction / Home Improvement
◯ Consulting
◯ Education
◯ Engineering / Architecture
◯ Entertainment / Recreation
◯ Finance / Banking / Insurance
◯ Food Service
◯ Government / Military
◯ Healthcare / Medical
◯ Internet
◯ Legal
◯ Manufacturing
◯ Marketing / Market Research / Public Relations
◯ Media / Printing / Publishing
◯ Mining
◯ Non-Profit
◯ Pharmaceutical / Chemical
◯ Research / Science
◯ Real Estate
◯ Retail
◯ Telecommunications
◯ Transportation / Distribution
◯ Utilities
◯ Wholesale
◯ Other - Write In
◯ Not applicable

45. Check all that apply:
□ I have downloaded and used the Uber and/or Lyft mobile apps
□ I have been a passenger in an Uber and/or Lyft car
□ I have read an article, viewed a program, or learned online about Uber self-driving cars
□ I have seen an Uber self-driving car while I was a driver
□ I have seen an Uber self-driving car while I was a pedestrian
□ I have ridden in an Uber self-driving car as a passenger
□ I have not yet ridden, but would ride as a passenger in an Uber selfdriving car
□ I am or have been an Uber and/or Lyft driver
□ I have ridden as a safety driver in an Uber self-driving car
□ I am currently or have previously been employed by Uber directly (not as a driver)
□ I disagree with actions Uber has taken
□ I have deleted the Uber app
□ I have attended a protest against Uber
□ None of the above

Privacy and Technology Questions

46. Are you majoring in or have a degree or job in computer science, computer engineering, information technology, or a related field?
◯ Yes
◯ No

47. Privacy is really a matter of people's right to exercise control and autonomy over decisions about how their information is collected, used, and shared.
◯ Strongly agree
◯ Agree
◯ Neither agree nor disagree
◯ Disagree
◯ Strongly disagree

48. Control of personal information lies at the heart of privacy.
◯ Strongly agree
◯ Agree
◯ Neither agree nor disagree
◯ Disagree
◯ Strongly disagree

49. I believe that privacy is invaded when control is lost or unwillingly reduced as a result of a marketing transaction.
◯ Strongly agree
◯ Agree
◯ Neither agree nor disagree
◯ Disagree
◯ Strongly disagree

50. Companies seeking information should disclose the way the data are collected, processed, and used.
◯ Strongly agree
◯ Agree
◯ Neither agree nor disagree
◯ Disagree
◯ Strongly disagree

51. It is very important to me that I am aware and knowledgeable about how my personal information will be used.
◯ Strongly agree
◯ Agree
◯ Neither agree nor disagree
◯ Disagree
◯ Strongly disagree

52. It usually bothers me when companies ask me for personal information.
◯ Strongly agree
◯ Agree
◯ Neither agree nor disagree
◯ Disagree
◯ Strongly disagree

53. When companies ask me for personal information, I sometimes think twice before providing it.
◯ Strongly agree
◯ Agree
◯ Neither agree nor disagree
◯ Disagree
◯ Strongly disagree

54. It bothers me to give personal information to so many companies.
◯ Strongly agree
◯ Agree
◯ Neither agree nor disagree
◯ Disagree
◯ Strongly disagree

55. I'm concerned that companies are collecting too much personal information about me.
◯ Strongly agree
◯ Agree
◯ Neither agree nor disagree
◯ Disagree
◯ Strongly disagree

56. To be entered into the raffle for Amazon gift cards, please provide your email address: (We will never use this for purposes out of this research)

[textbox]

57. Is there anything else you would like to add about networked self-driving cars or this survey in general?

[textbox]

Hello,

We would like to thank you again for participating in our study. If you are selected for the raffle, an Amazon gift card code will be sent to the email you provided.

This study is aimed at determining people's awareness and preferences toward the privacy considerations surrounding Uber's self-driving cars. The data you provided will be used to help determine future areas of study and help craft recommendations for the industry in addressing consumer privacy needs and concerns.

Deployed fleets of autonomous vehicles like Uber's self-driving cars are a new phenomenon, and researching these cars in ordinary, real-world scenarios has just begun. From what we know, Uber self-driving cars have three different types of sensors:

1. Radar sensors that map the physical world around the car. They do not collect video and do not store any information; they are just used for navigational purposes.
2. The large camera lens on the roof is used to detect colors, such as those on a traffic light or a stop sign. It does not collect photo or video.
3. Twenty other cameras are used to detect braking vehicles, pedestrians, and other obstacles. Some cameras store video that can be reviewed later manually by people, or via automated computer algorithms.

Some participants in this study were exposed to this information during the study, while others were not. This was done to gauge how people perceive the privacy concerns surrounding Uber's cars with and without context.

Thanks again for your time and ongoing participation in our study. For any further feedback on the study, feel free to email at: selfdrivingcarresearch@cmu.edu