Effects of Knowledge and Experience on Privacy Decision-Making in Connected Autonomous Vehicle Scenarios

Zekun Cai
The Pennsylvania State University
zuc204@psu.edu

Aiping Xiong
The Pennsylvania State University
axx29@psu.edu

***Abstract*—To enhance the acceptance of connected autonomous vehicles (CAVs) and facilitate designs to protect people's privacy, it is essential to evaluate how people perceive the data collection and use inside and outside the CAVs and investigate effective ways to help them make informed privacy decisions. We conducted an online survey ($N = 381$) examining participants' utility-privacy tradeoff and data-sharing decisions in different CAV scenarios. Interventions that may encourage safer data-sharing decisions were also evaluated relative to a control. Results showed that the feedback intervention was effective in enhancing participants' knowledge of possible inferences of personal information in the CAV scenarios. Consequently, it helped participants make more conservative data-sharing decisions. We also measured participants' prior experience with connectivity and driver-assistance technologies and obtained its influence on their privacy decisions. We discuss the implications of the results for usable privacy design for CAVs.**

I. Introduction

Connected Autonomous Vehicles (CAVs) are vehicles that can communicate with other systems outside of the vehicle and have self-driving capabilities [70]. CAVs are currently being developed and expected to be applied in the future [51], [52]. In order to support its autonomy without direct driver input, CAVs use wireless networks and sensors inside and outside the vehicle to obtain relevant traffic and other vital information for various use scenarios. For example, CAVs need to continuously collect 2D or 3D photos of the road scene to ensure driving safety [43]. Other use scenarios, such as authentication with face or speech recognition, require the collection and use of drivers' photos or voice data [42].

On the one hand, those data collection and use have great potential to enable CAVs to reduce traffic accidents, increase the efficiency of transportation systems, and improve driving experience. On the other hand, the ubiquitous data collection by CAVs raises unique privacy challenges from various aspects, including the technical side and the legal and policy perspective. Technical solutions have been proposed and investigated, such as obfuscation [79]. Countries and regions worldwide have also begun to update and improve their data protection regulations [3]. However, to ensure consumer acceptance of CAVs, the *human* aspects of privacy must first be understood.

With the rapid progress in data mining and machine learning, various sensor data collected by CAVs can be used and aggregated to infer potentially sensitive personal information that previously seemed impossible, such as income level or health status [48], [86], [92]. Prior work revealed that while participants were uncomfortable with secondary use scenarios of CAVs such as recognition and identification, they thought those scenarios were less unlikely [8]. Thus, people may not be fully aware of the scope of data collection and use of CAVs. Consequently, their lack of awareness of the possible inferences can result in underestimated privacy risks of CAVs. Also, in reality, service providers (e.g., automobile manufacturers) tend to focus on promoting their services (e.g., automatic auto ignition) while not being transparent about the sensor data collection and analysis required for the services (e.g., face images). Thus, a meaningful approach to inform people of both utility benefits and privacy risks of CAV services is essential.

Previous studies have shown that participants could be primed by thinking about their safety and privacy through answering privacy statements [14], [27], [36], [47], [65]. Yet, increased privacy awareness does not necessarily ensure conservative privacy decisions or actions [1], [64]. Feedback has been shown to foster participants' accurate mental models and informed decisions of privacy [72]. Users' prior experience with driver assistance technologies could also modulate their risk perception of autonomous vehicles [11].

We conducted an online survey on Amazon Mechanical Turk (MTurk) to understand people's perceived privacy risks and privacy decision-making in different CAV scenarios. We created eight CAV scenarios, in half of which the data collected were intended to improve *safety/security* and in the other half of which the data collected were intended to enhance *convenience*. For each scenario, participants were prompted for their 1) perceived balance between utility benefits and privacy risks; 2) data-sharing decisions; and 3) confidence in the privacy decisions. There were three between-subject conditions: *control*, *priming*, and *feedback*. We primed participants on the privacy implications of data collection by asking them to select the possible inferences (i.e., unintended but feasible use of

This work was funded in part by NSF award #1931441 and a seed grant from Penn State's Center for Security Research and Education (CSRE).

Usable Security and Privacy (USEC) Symposium 2022
28 April 2022, San Diego, CA, USA
ISBN 1-891562-79-7
https://dx.doi.org/10.14722/usec.2022.23036
www.ndss-symposium.org, https://www.usablesecurity.net/USEC/

collected data to infer personal information). Participants in the feedback condition were further informed of their selection performance of each possible-inference question.

Compared to the control condition, participants in the two intervention conditions increased their perceived privacy risks for the safety/security scenarios where data-sharing seems mandatory to ensure the functions. However, only participants in the feedback condition became more conservative in their data-sharing decisions for the safety/security scenarios. Participants with much experience in using connectivity and driver-assistance functions perceived more benefits and made liberal privacy decisions for the convenience scenarios in which data collection is not primarily concerned with driving.

Altogether, our results highlight the importance of equipping users the knowledge of the data collection implications of CAVs for informed privacy decision-making. In summary, the contributions of our work include:

1) We perform the first empirical study evaluating people's perceived privacy risks and their privacy decision-making in the CAV context.
2) We identify various factors, such as knowledge and experience, which can influence people's perceived privacy risks and data-sharing decisions of CAVs.
3) We investigate people's privacy decision-making as a result of considering both utility benefits and privacy risks of the collected data.
4) We suggest a scenario-based method for evaluating people's privacy decision-making of CAVs and demonstrate its feasibility based on the results of an online experiment.

II. Related Work

In this section, we first review research efforts articulating people's privacy concerns for CAVs. We next describe previous work evaluating various factors impacting people's privacy decision-making in different computing environments. We also assess two techniques, priming and feedback, which have been used to maximize people's privacy awareness and inform the consequent privacy decision-making in the contexts of smartphones and the Internet of Things (IoT). Following a discussion about the effect of driving technology experience on people's perceived risks of CAVs, we summarize how our work is different from others at the end.

A. Privacy Concerns of CAVs

While there are only few studies evaluating the privacy of CAVs from users' perspective, prior work has focused on users' privacy concerns [8], [37]. Privacy concerns refer to people's privacy-related attitudes, which are shaped by users' awareness of privacy-relevant information [55], [71], [73].

Bloom et al. [8] conducted an online survey with 302 participants, in which they investigated participants' concepts of technological capabilities and general privacy concerns of CAVs. The participants answered questions in different scenarios, in which data collection and processing were for the primary uses (e.g., image capture for navigation) or the secondary uses (e.g., aggregation and analysis of captured images for identification, recognition, and tracking of individuals and vehicles). Bloom et al. found that most participants correctly believed that CAVs have the capability to gather rich information about the environment and perform analyses of collected data for primary uses. Nevertheless, less than half of the participants rated the secondary uses of collected information were likely. The participants were more comfortable with the primary uses than with the secondary uses.

While the data collection and processing for primary uses are necessary for CAV functions and features, the secondary uses, such as improved planning of personal travel routes [93], are essential to fulfill the promise for CAVs. With the rapid progress on data mining and machine learning, various sensor data can now be used and aggregated to infer potentially sensitive personal information that previously seemed impossible such as image processing for fatigue detection [38]. Thus, a lack of awareness and knowledge of the secondary uses scenarios can result in an underestimate of the CAV's privacy risks, and uninformed privacy decisions.

B. Factors Impacting Privacy Decision-Making

People regulate their privacy through the dynamic processes of awareness, decision-making, and action selection [4]. Due to the lack of operational CAV environments, our following discussion mainly focuses on prior work about privacy decision-making in the mobile devices and IoT settings.

1) Privacy Awareness: Privacy awareness refers to people's attention, perception, and cognition of possible risks throughout the interaction with an application or service that can gather and process personal data or information [64]. An effective way to help people make better privacy decisions seems to make them aware of privacy implications of data sharing [50], especially the types of personal information that can be inferred [30], [88]. Lee and Kobsa [47] conducted an online survey on Amazon MTurk investigating people's privacy decision-making in IoT service scenarios. Each participant viewed 15 different scenarios and answered their data-sharing decision for each scenario. Along with each privacy decision, possible inferences of personal information were presented for participants to select, helping them understand some privacy implications of using various IoT services. The study results showed that participants who were more aware of the privacy implications of using IoT services made more conservative and confident decisions. Yet, it is unclear whether greater awareness of inferable personal information will impact people's privacy decisions in the CAV scenarios.

2) Context-dependent privacy decisions: In previous work on IoT privacy decision-making, contextual factors (e.g., service ownership and location) have been systematically varied in the hypothetical scenarios [5], [19]. The results showed that participants' privacy decisions were context dependent. Howard and Dai [31] examined people's attitudes toward self-driving cars in a group-administrated survey setting. Results from 107 participants showed that individuals were most attracted to potential safety benefits and the convenience. In this study, we, therefore, investigated the CAV services for the

purposes of safety/security and convenience. For each service purpose, we generated different CAV scenarios by varying *what* was collected [19].

3) Utility-privacy tradeoff: In real-world scenarios, people make decisions by evaluating more than one attribute that influences their final decision [39], [53]. Given a scenario of CAV, data-sharing decisions are reached by considering factors from different aspects. For example, users could decide to share their location data, knowing that the data will benefit others suffering from traffic jams, even though there might be a chance of being identified. Likewise, people could decide not to share personal information such as photos inside the driving cockpit, in case a company will utilize the captured behavior for usage-based automotive insurance [17]. In such cases, the data-sharing decision represents a result of a weighting process. Thus, the prior work in which utility benefits or privacy risks were examined separately only revealed to what extent the single level of each relevant factor was accepted [8], but not the tradeoffs in between.

4) Effects of priming and feedback on informed privacy decisions: Priming refers to the phenomenon that when a stimulus (e.g., one word or a picture) makes associated information from the long-term memory (e.g., a concept) more available to people in the short-term memory. Consequently, people tend to consider that information into their behavioral responses. Previous studies showed that participants could be primed by thinking about their safety and privacy through answering privacy statements [14], [27], [65].

In the hypothetical IoT scenarios, when participants were prompted to select possible inference of collected data in IoT scenarios, a positive correlation was evident between the level of privacy awareness and the probability to make more conservative privacy decisions [47]. The results suggested that participants' who were knowledgeable about possible inferences of personal information in IoT scenarios tended to make conservative decisions about the data sharing decisions. However, due to lack of a control condition without possible-inference selection, it is hard to infer the causal relation between the priming effect of possible-inference selection on participants' privacy decisions in hypothetical scenarios. Moreover, increased privacy awareness does not necessarily ensure conservative privacy decisions or actions [1], [64].

Feedback has been shown to encourage participants to develop accurate mental models and make informed privacy decisions [61], [72]. Tsai et al. [77] studied the impact of giving feedback on helping people manage their privacy on a location sharing application. In the study, participants were informed of whom their data was shared with and when the data was shared. Their results showed that when participants got adequate feedback, they were more willing to share data, and were also more comfortable with sharing their locations. Considering the difficulty for everyday users to make possible inferences of collected data of CAVs accurately, we hypothesized that informing user selection performance through feedback could facilitate their data-sharing decisions.

5) Decision confidence: Confidence in judgment or decision making refers to an individual's beliefs about the goodness of his or her judgments or choices [63]. A higher level of privacy awareness was positively correlated to making more confident decision in the IoT settings [47]. However, it is unclear whether people are confident about their data-sharing decisions in various CAV scenarios and what factors may impact their confidence.

C. Influence of Technology Experience

Users' prior experience with advanced driver assistance systems (e.g., adaptive cruise control) has a positive effect on their acceptance of autonomous vehicles. Rödel et al. conducted an online survey, in which participants viewed five different scenarios from non-autonomous to full autonomous [68]. For each scenario, participants answered 11 questions about autonomous vehicles, including acceptance, perceived ease of use, and attitude. The highly experienced participants showed a tendency of higher acceptance of autonomous vehicles than the inexperienced participants. Previous studies have also revealed that users' familiarity and experience with driver assistance systems could modulate their privacy perception [11], [45], [78].

D. The Present Study: Informed Privacy Decisions of CAVs

To the best of our knowledge, no prior works have tried to evaluate the privacy awareness of CAVs and the impact of increased privacy awareness on people's privacy decisions. It is also unclear whether people will make informed privacy decisions of CAVs if they are equipped with knowledge of privacy implications. Thus, in this work, we examined the effects of priming and feedback in facilitating privacy awareness of different CAV data-collection scenarios, and how the privacy awareness impacted people's consequent benefit-risk tradeoff, privacy decision, and decision confidence. We also investigated the possible moderation due to people's prior experience with driver assistance and connectivity technologies.

III. Method

We conducted an online survey on Amazon Mechanical Turk (MTurk). There were three between-subject conditions: *control*, *priming*, and *feedback*. Due to the current unavailability of the cyber-physical operating environment of CAVs, we conducted the vignette (scenario) survey [21]. Such a method was found to well approximate real-world behaviors [26] and has been used in other similar settings, such as IoT privacy [20], [47]. We created a set of data collection and use scenarios related to CAV features that enhance driving safety/security or convenience [58] and presented the same set in a randomized order at each condition. Each scenario described the data flow and information usage of a specific CAV function or feature (e.g., object detection). After viewing each scenario, participants were prompted for 1) perceived tradeoff between utility benefits and privacy risks of the data collection and use; 2) data-sharing decision; and 3) confidence rating of the data-sharing decision.

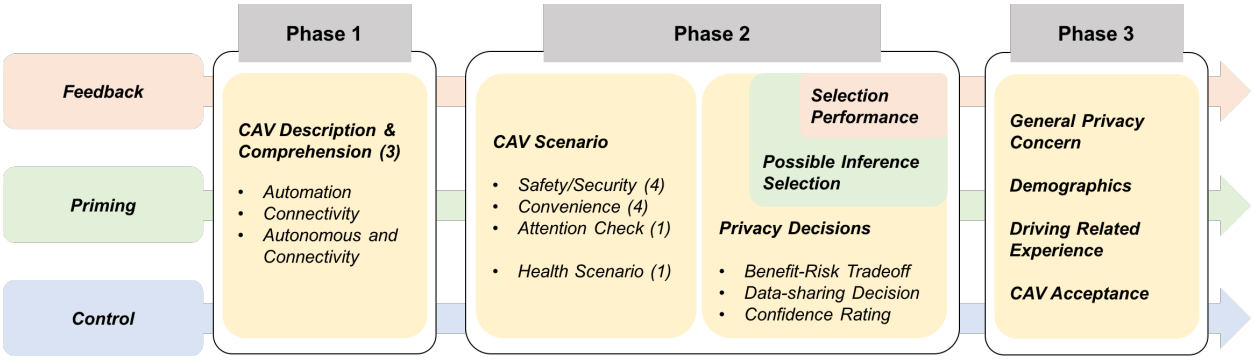


Fig. 1: Experiment flow chart. "Phase 1", "Phase 2", and "Phase 3" boxes show the details in each phase. All three conditions were the same except that participants answered possible inference questions in the top two conditions (i.e., feedback and priming) at Phase 2. Moreover, participants in the feedback condition received extra information about their inference-selection performance.

To increase participants' privacy awareness, we also asked them to select possible inferences of collected data (i.e., unintended but feasible use of the collected data [41]) before answering the three questions of each scenario in the priming and feedback conditions. Participants in the feedback condition were further informed of their inference-selection performance in the form of the correct answer to the question. In the end, participants answered a few post-session questions, including their general privacy attitudes and demographics.

A. Participants

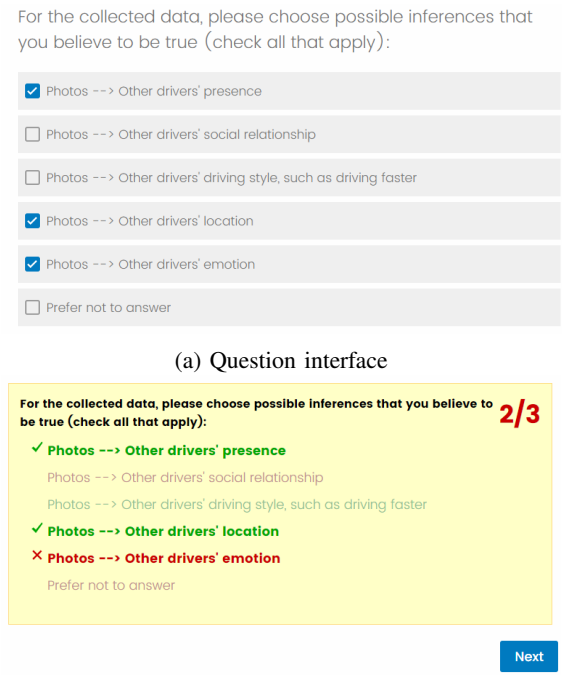
We recruited 600 Amazon MTurk workers in August 2020. We did not run an a priori power analysis due to a lack of empirical data on the effect size. Yet, the sample size was comparable to a similar study in the IoT context (488) [47]. The human intelligent task (HIT) was posted with restrictions to workers who (1) are at least 18 years old; (2) completed more than 100 HITs and with a HIT approval rate of at least 95%; (3) are located in the United States; and (4) are vehicle owners. This experiment complied with the American Psychological Association Code of Ethics and was approved by the institutional review board at the Pennsylvania State University. Informed consent was obtained from each participant. The experiment data that were stored and analyzed are anonymized.

B. Apparatus and Stimuli

The study was performed with participants' own laptops or computers. To ensure the readability of the stimuli's content, we did not allow participants to continue the study if they were using any mobile device. To situate participants into the hypothetical scenarios for the CAVs, we presented a brief description of CAVs and examples of the CAV features [70] at the beginning of the study. We also proposed three True/False questions (i.e., 1. the role of human driver; 2. CAVs' data collection and use; 3. the CAV's definition) evaluating participants' understanding of CAVs based on the description (see details in Appendix B1).

We defined two scenario categories focusing on two different aspects of CAV features: enhancing driving safety/security and convenience [25]. We referred to the use cases in the IoT and CAV literature [8], [10], [31], [47], [59] and created four scenarios in each category describing the collection and use of various types of data (e.g., audio [69], visual [32], and biometric [75]). Each safety/security-related scenario focused on a specific CAV function or feature and listed the relevant data flow and information usage. At the end of the scenario description, we also highlighted the intended purpose of data collection and the core data inference to support the function or feature. For example, in one safety/security-related scenario, we described that cameras and sensors outside the CAVs collect images of other vehicles on the road to predict other vehicles' trajectory and plan the CAV's next action. In the end of the scenario description, we also made it clear that such data collection and use were for the purpose of safety and the trajectory prediction was inferred from the collected photos (see Appendix A for the scenario descriptions). The four convenience-related scenarios were constructed in the same way except that the collected data were mainly for the convenience purpose, for example, collecting playlists to infer driver's music preference for better music recommendation). To understand the possible learning from priming and feedback interventions, we proposed the ninth scenario that was about collecting biometric data for the driver's state of health [86].

Besides the core inference of the collected data in each of the safety/security- and convenience-related scenarios, we specified three to four possible inferences that are unintended but feasible from the collected data based on the relevant literature (see Table I). For example, user's mood can be inferred from analyzing data such as mood labeled musical tracks (see S03 in Appendix A), although the core inference was to provide an appropriate playlist of songs recommended for the driver [13]. One to two impossible inferences were also identified, which were infeasible based on the stated data collection in each scenario (e.g., the use of road scene photos to infer the driver's mood, see S02 in Appendix A). Possible and impossible inferences for all scenarios were reviewed by two outside experts, one in the field of human factors and the other in data privacy. Each of them decided individually if each inference is possible or not for the scenario. They were instructed to make intuitive decisions based on their



(a) Question interface

(b) Feedback interface

Fig. 2: Top panel shows the interface of possible-inference selection. Bottom panel shows the feedback interface after participants submitted an incorrect possible-inference selection.

TABLE I: The possible inferences are generated based on the current available algorithms and techniques.

Data Type	Possible Inference	References
Photo of Drivers	Mood; Demographics; Non-Driving Activity; Personality Types; Identity	[6], [34], [62], [83], [85]
Voice	Mood; Demographics; Physical Status; Personality Types; Identity	[56], [74], [82], [84]
Photo of Road Scene	Scene Understanding; Frequency of Visit; Whereabouts; Income Level; Preference	[29], [32], [49]
Photo of Other Vehicles	Object Detection; Frequency of Visit; Whereabouts; Income Level; Preference	[29], [33], [54], [60], [80]
Photo of the Environment	Object Detection; Presence; Demographics; Identity; Social Relationship	[15], [34], [85]
OBD Data	Vehicle Condition; Use Pattern; Driving Pattern; Frequency of Driving; Personal Settings/Preference	[16], [22], [57]
Playlist	Music Preference; Personality Type; Mood; Demographics	[13], [40], [67], [90]

knowledge and expertise, rather than referring to literature or other materials. The average ratio of agreement (Cohen's Kappa) of the coding was 0.68, indicating a moderate agreement level [44]. We resolved the disagreement by replacing ambiguous inferences with clear ones.

All possible and impossible inferences were listed as options in the inference-selection questions in a randomized manner (see Appendix B2 for the details). Participants' correct answer rate was served as the measure to gauge their privacy awareness. Considering the difficulty of understanding the privacy implications of CAV data collection, we proposed the feedback intervention. After participants selected the possible inferences (see Figure 2a), we informed them of their selection performance with a feedback interface (Figure 2b). Participants' correctly selected options were highlighted in green, and incorrectly selected items were in red. Unselected correct and incorrect options were also highlighted using the same color coding but in a transparent manner. The correct selection rate of possible inferences (e.g., 2/3 in Figure 2b) was also presented on the top right corner of the interface.

C. Procedure

Participants who accepted the HIT on MTurk were directed to our survey on Qualtrics. After the informed consent, participants were randomly assigned to one of the three conditions, feedback (i.e., inference-selection performance was provided), priming (i.e., inference selection without feedback), and control (i.e., without inference selection).

Each condition consisted of three phases (see Figure 1). In Phase 1, we presented the brief description of the CAVs. All participants were asked to read the description carefully and then answered the three True/False questions, which were presented in a randomized order. In Phase 2, participants answered privacy-decision questions across the four safety/security-related scenarios and the four convenience-related scenarios. After viewing each scenario description, participants first selected the utility-privacy tradeoff with a 5-point scale ("1" means "Benefits are much less than risks", "5" means "Benefits are much greater than risks"). Then, they made the data-sharing decision ("Yes", "No"), and evaluated their confidence rating of the decision with another 5-point scale ("1" means "Very unconfident", "5" means "Very confident"). Following each scenario description, participants in the feedback and priming conditions were also asked to select the possible inferences of the collected data before answering the three privacy-decision questions. Moreover, participants in the feedback condition were informed of their selection performance (see Figure 2b).

In Phase 2, we also included one scenario to check participants' attention [28]. The attention-check scenario was presented in the same way as the eight scenarios except that participants were asked to select specified correct options for the questions. The attention-check scenario and the eight scenarios were presented in a randomized order. To understand the possible learning from priming and feedback, we presented the health scenario at the end of Phase 2. Participants in all conditions answered the three privacy-decision questions as in the control condition.

At Phase 3, we measured participants' general privacy attitudes using a subset of Internet users' information privacy concerns (IUIPC) questions [55] and asked participants to indicate their agreement on the descriptions using a 5-point Likert scale ("1" means "Strongly disagree", "5" means "Strongly agree"). The general privacy attitude was measured after Phase

TABLE II: Summary of survey responses in the experiment.

Measurement	Response	Description	Data Type
CAV Comprehension (Phase 1)	Understanding Level	Correct answer rate of the three True/False questions about CAV	Numerical (0.0∼1.0)
Privacy Awareness (Phase 2)	Understanding Level	Correct answer rate to the possible inferences question in a given CAV scenario	Numerical (0.0∼1.0)
Privacy Decisions (Phase 2)	Utility-Privacy Tradeoff	Perceived balance between utility benefits and privacy risks in a given CAV scenario	Ordinal (5pt-scale)
	Privacy Decision	Intention to use or not to use a given CAV scenario	Categorical (binary)
	Decision Confidence	Perceived level of confidence in making a privacy decision for a given CAV scenario	Ordinal (5pt-scale)
Privacy Attitude (Phase 3)	Agreement Level	Level of agreement with presented statements	Ordinal (5pt-scale)
Technical Experience (Phase 3)	Experience Level	Level of experience in driver assistance and connectivity functions	Categorical (3-class)
CAV Acceptance (Phase 3)	Willingness to Use CAVs	Intention to use CAVs	Ordinal (4pt-scale)

2 to avoid the possible priming effect on the privacy decisions. Following that, participants' demographic information such as gender and age range was collected. We also asked about participants' experience in using driver assistance functions and connectivity functions, respectively ["No, not at all" (1), "No, rarely" (2), "Yes, sometimes" (3), "Yes, quite often" (4)]. They also indicated their willingness to use CAVs in the future ["No, never" (1), "No, rarely" (2), "Yes, for some cases" (3), "Yes, always" (4)].

A pilot study (N = 20) was conducted on Amazon MTurk to ensure all scenarios, survey questions, and procedures were understandable to the participants. Based on the results, we made minor edits to the scenarios and survey questions.

D. Data Exclusion

For the obtained results, we first removed 14 duplicated responses and 67 participants who completed the survey either shorter than 5 min or longer than 30 min (T_{Mean}: 12.5 min, T_{Median}: 11 min). Correct answer rate of the three questions about CAV descriptions at Phase 1 was 88.5%. Results were similar across the three conditions (feedback: 86.9%, priming: 91.0%, control: 87.9%), $\chi^2_{(2)}$ = 1.58, p = .45, indicating that most participants in all conditions got a basic understanding of CAVs. To ensure that the results at Phase 2 were based on participants' understanding of CAV, we excluded 60 participants who answered two or more questions incorrectly at Phase 1. Another 60 participants failed the attention check and an extra 22 participants chose "prefer not to answer" to at least one question in Phases 2 and 3, and thus were excluded. The removal of those data is warranted since the current study is based on the CAV scenarios with which the public are not familiar. Thus, participants' comprehension of CAV and their attention to the scenario descriptions are critical to guarantee the data quality. For the remaining participants, there was an approximately equal number in each condition, feedback (125), priming (128), and control (128).

E. Analysis Plan

We summarize the survey responses collected through the abovementioned procedures in Table II. Our statistical analysis focused on the measures related to privacy decisions at Phase 2. We first measured the correct answer rate of the possible-inference selection for the priming and feedback conditions. The correct answer rate was coded as the percentage of correctly selected options in each scenario, i.e., correctly chose the possible inferences but not selected the impossible inferences. Correct answer rate for possible-inference selection was determined for each participant and grouped as a function of 2 (condition: priming, feedback) $\times$ 2 (scenario: safety/security, convenience) for mixed analyses of variance (ANOVA). We also measured participants' utility-privacy tradeoff rating, data-sharing decision, and decision confidence rating of each scenario across all conditions at Phase 2. The three measures were determined for each participant and grouped as a function of 3 (condition: control, priming, feedback) $\times$ 2 (scenario: safety/security, convenience) for mixed ANOVAs.

At Phase 3, we evaluated participants' general privacy attitudes, their experience in using connectivity and driver assistance functions, and their intention to use CAV. Based on the participants' responses to the two questions, their experience was categorized into three levels: little (never or rarely used either function), some (sometimes or often use at least one of the functions), much (sometimes or often use both functions). We conducted ANOVAs as Phase 2 but added experience (little, some, much) as another measured between-subject factor for possible-inference selection and three privacy-decision measures, respectively.

IUIPC is typically used to understand people's general privacy attitude toward online information [24]. However, due to the intervention manipulation at Phase 1 and potential impacts of people's prior experience on connectivity and driver assistance functions, we chose to examine results of the IUIPC questions as dependent variable (see similar examination by [18]). Thus, general privacy attitude and willingness to use CAVs were also determined for each participant and grouped as a function of 3 (condition: control, priming, feedback) $\times$ 2 (scenario: safety/security, convenience) $\times$ 3 (experience: little, some, much) for ANOVAs.

We conducted null-hypothesis testing ($\alpha = 0.05$) for those measures. The null hypothesis was rejected when the obtained results among the conditions were significantly different from each other. Post-hoc tests with Bonferroni correction were performed, testing pairwise comparisons with corrected p values for possible inflation. We report the critical findings in the following section. Complete descriptive and inferential statistics are shown in Tables VIII and IV in Appendix C.

IV. Results

Participants' demographic information is shown in Table III, and the distributions are similar across conditions. Each participant who completed the study was paid $2.00. To decide the payment, we considered the US federal minimum wage

TABLE III: Demographic information of participants.

Gender		Age		Ethnicity		Degree		Related Major or Experience on CS or IT		Mileage (Miles/Year)	
Male	54.6%	18-24	3.4%	African / African American	8.7%	High	20.7%	No	71.4%	<2,000	7.3%
Female	45.1%	25-34	36.5%	American Indian / Alaska Native	0.5%	College	4.2%	Yes	28.1%	2,000 - 5,000	19.2%
Other	0.3%	35-44	31.5%	Asian	6.0%	Associate	11.5%	Unknown	0.5%	5,000 - 10,000	38.3%
		45-54	15.0%	Caucasian	78.7%	Bachelor	43.6%			10,000 - 20,000	27.8%
		>55	13.6%	Hispanic / Latino	3.4%	Professional degree (Masters / Ph.D.)	18.9%			>20,000	5.0%
				Native Hawaii / Pacific Islander	0.3%	Medical	0.8%			Unknown	2.4%
				More than one race	2.4%	Unknown	0.3%				

($7.25/hour) and the average survey time ascertained in the pilot testing (15 min).

A. Possible Inference Selection

Participants' correct rate for the possible inference selection was higher for the safety/security scenarios (72.9%) than for the convenience scenarios (67.2%), $F_{(1,251)} = 52.44$, $p < .001$, $\eta_p^2 = .173$, and was higher for the feedback condition (74.5%) than for the priming condition (65.5%), $F_{(1,251)} = 32.33$, $p < .001$, $\eta_p^2 = .114$, resulting in significant main effects of scenario and condition. Yet, the effect of feedback was similar across scenarios, $F < 1.0$, resulting in a non-significant two-way interaction. Thus, participants showed more privacy awareness of data collected for the safety/security purpose than for the convenience purpose. The extra feedback informed participants' inference selection, which was effective in enhancing people's privacy awareness.

B. Privacy Decisions

Figure 3 shows the average results of three privacy-decision measures at Phase 2 (utility-privacy tradeoff, data-sharing decision, and confidence rating) for different scenarios across the three conditions.

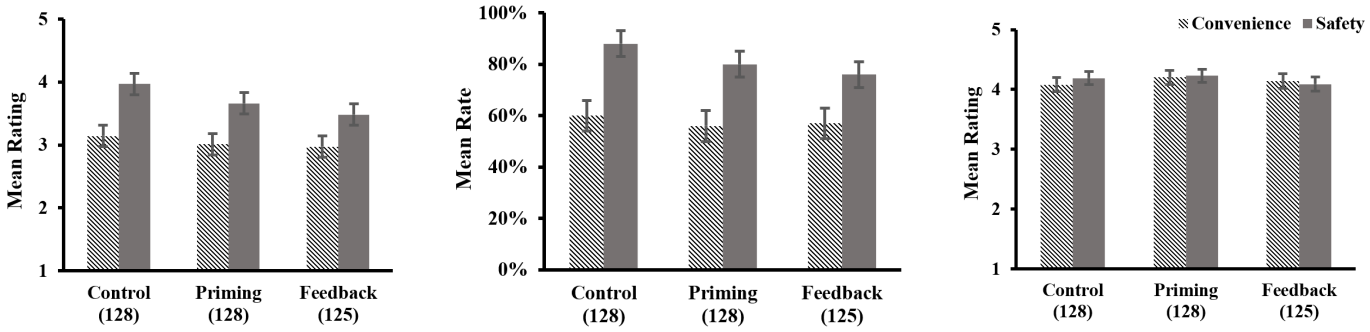
1) Utility-Privacy Tradeoff: As shown in Figure 3a, participants weighted the safety/security scenarios as more beneficial (3.70) than the convenience scenarios (3.04), $F_{(1,378)} = 264.49$, $p < .001$, $\eta_p^2 = .412$. The effect of condition was mainly revealed by participants in the feedback condition (3.22) perceiving more risks than those in the control condition (3.56), $p_{adj} = .007$. Results of the priming condition (3.33) showed no significant differences with the other two conditions, $p_{adjs} \geq .11$. Moreover, the difference across conditions was only evident in the safety/security scenarios, $F_{(2,378)} = 8.67$, $p < .001$, $\eta_p^2 = .042$, but not in the convenience scenarios, $F_{(2,378)} = 1.19$, $p = .61$, $\eta_p^2 = .006$, resulting in a significant two-way interaction. Specifically, participants in the control condition gave higher rating in the benefit-risk tradeoff than those in the feedback and the priming conditions, $p_{adjs} \leq .026$, while there were no significant differences between the latter two conditions, $p_{adj} = .415$. While informing possible-selection performance (i.e., feedback) was sufficient to increase participants' privacy awareness in general, their perceived privacy risks were largely dependent on the data-collection contexts.

2) Data-sharing Decisions: Participants revealed more intentions to share personal information for the safety/security scenarios (81.3%) than for the convenience scenarios (57.7%), resulting a main effect of scenario, $F_{(1,378)} = 283.99$, $p < .001$, $\eta_p^2 = .429$. The main effect of condition was not significant, but there was a significant two-way interaction of condition $\times$ scenario. Participants' data-sharing decisions differed across conditions for the safety/security scenarios, $F_{(2,378)} = 6.05$, $p = .005$, $\eta_p^2 = .031$, but not for the convenience scenarios, $F < 1$. Specifically, for the safety/security scenarios, participants in the feedback condition showed less intention to share data than those in the control, $p_{adj} = .002$ (see Figure 3b). All the other pairwise comparisons were not significant, $p_{adjs} \geq .116$. Thus, although participants in the feedback condition increased the awareness of privacy risks in general, they only reduced data-sharing decisions for the safety/security scenarios.

We averaged the binary data-sharing decision to obtain the average decision rates of scenario type for each participant in ANOVA. Considering the possible distribution among the four scenarios for the same purpose (i.e., safety/security or convenience), we further performed a Generalized Linear Mixed-Effects Regression (GLMER) with the package lme4 [7] in R. GLMER allows controlling for the random effect for participants without data aggregation[9]. The model was the same with the one used in ANOVA except that we added the random intercepts of participants and scenarios. The GLMER result (See Table X in Appendix C) showed the same results as ANOVA. The main effect of scenario, $\chi^2_{(1)} = 5.58$, $p = .018$, and the interaction effect, $\chi^2_{(2)} = 6.60$, $p = .037$, are significant, but no main effect of conditions, $\chi^2_{(2)} = 2.40$, $p = .302$.

3) Confidence Rating: Participants showed similar confidence on their data-sharing decisions regardless of scenarios (convenience: 4.14; safety/security: 4.17) or conditions (feedback: 4.12; priming: 4.21; control: 4.13). The two-way interaction of condition $\times$ scenario was not significant either.

4) The Health Scenario: At the end of Phase 2 for each condition, participants viewed the same health scenario and answered the three privacy-decision questions as in the control condition. Neither the utility-privacy tradeoff nor the decision confidence rating showed any significant differences across the three conditions (see Table IV in Appendix C). However, the data-sharing decision rates differed across conditions, $\chi^2_{(2)} = 8.83$, $p = .012$. Specifically, more participants in the feedback



(a) **Utility-Privacy Tradeoff** (b) **Data-Sharing Decision** (c) **Confidence Rating**

Fig. 3: Results of three privacy-decision measures at Phase 2 as a function of condition (Control, Priming, Feedback) and scenario (Convenience, Safety/Security). Error bars represent 2 standard errors. **Left panel (a)** shows that compared to participants in the control condition, those in the priming and feedback conditions perceived more privacy risk for the safety/security scenarios only. **Center panel (b)** shows that only participants in the feedback condition revealed less intention to share the data than those in the control condition for the safety/security scenarios. **Right panel (c)** shows no statistically significant difference for the confidence rating measure in general.

condition (59%) showed willingness to share their information on the health scenario than participants in the control condition (41%), $p_{adj} = .009$. Pairwise comparisons involved the priming condition (52%) were not significant, $p_{adjs} \geq .238$. Instead of being more conservative on their privacy decisions, participants in the feedback condition increased their data-sharing decisions for their biometric information.

C. Effect of Experience on Privacy Decisions

We asked participants' experience of using connectivity and driving-assistance functions. We conducted ANOVAs by adding experience (little, some, much) as the measured between-subject factor for possible-inference selection, privacy-decision measures, general privacy attitude, and willingness to use CAVs, respectively. Figure 4 shows the results of the three privacy-decision measures at Phase 2 (see Table VII in Appendix C for descriptive statistics and Table V for ANOVA results). For the possible-inference selection and confidence rating on privacy decisions, no term involved experience was significant, $Fs \leq 2.06$. The analyses below focus on the effect of experience on the other measures.

1) Privacy-utility Tradeoff: Only the two-way interaction of experience × scenario was significant, $F_{(2,372)} = 7.49$, $p = .001$, $\eta_p^2 = .039$. Participants with much experience (3.25) viewed the convenience scenarios as more beneficial than those with some (2.96) or little (2.88) experience, $F_{(2,372)} = 5.37$, $p = .010$, $\eta_p^2 = .028$, but no difference was evident for the safety/security scenarios, $F < 1$ (see Figure 4a).

2) Data-sharing decisions: The main effect of experience was significant, $F_{(2,372)} = 8.44$, $p = .001$, $\eta_p^2 = .043$. Post-hoc pairwise comparisons revealed larger data-sharing decision rate for participants with much experience (76.2%) than those with little experience (61.4%), $p_{adj} < .001$. Yet, there was no significant difference between participants with some experience and the others, $p_{adjs} \geq .078$. The two-way interaction of experience × scenario was also significant, $F_{(2,372)} = 3.94$, $p = .020$, $\eta_p^2 = .021$. For the convenience scenarios, participants with much experience showed more data-sharing intention (67.0%) than those with some experience (55.7%), $p_{adj} = .017$, or little experience (47.8%), $p_{adj} < .001$. For the safety/security scenarios, the difference was also evident between participants with much experience (85.4%) and those with little experience (75.1%), $p_{adj} = .014$, but not other pairwise comparisons, $p_{adjs} \geq .192$ (see Figure 4b). The three-way interaction of experience × scenario × condition was not significant.

3) The Health Scenario: There was only a main effect of experience on participants' utility-privacy tradeoff, $F_{(2,372)} = 3.25$, $p = .040$, $\eta_p^2 = .017$. Specifically, participants with much experience (3.20) gave higher rating than those with little experience (2.70), $p_{adj} = .011$, while ratings from those with some experience (2.85) showed no difference with the other two groups, $p_{adjs} \geq .089$.

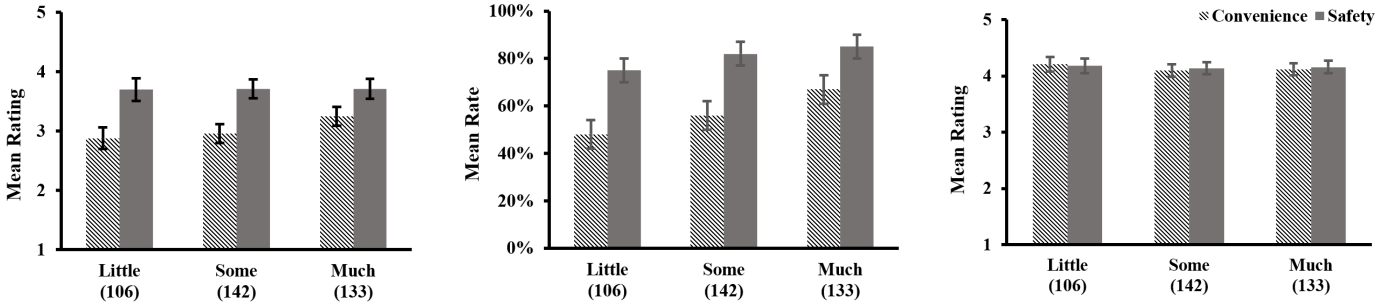
The main effect of experience was also significant for the data-sharing decision in the health scenario, $\chi^2_{(2)}$ = 29.95, p = .001. Post-hoc analysis showed larger data-sharing rate for participants with much experience (68.4%) than those with some experience (45.8%), $p_{adj} = .001$, or little experience (34.0%), $p_{adj} < .001$, respectively. Yet, there was no significant difference between the latter two groups, $p_{adj} = .205$.

4) General Privacy Attitude: The main effect of experience was significant (see Table VI in Appendix C). Post-hoc comparisons only revealed less privacy concern for participants with much experience (4.38) than those with little experience (4.52), $p_{adj} = .046$. Although the main effect of condition was significant, post-hoc pairwise comparisons revealed non-significant differences, $p_{adjs} \geq .078$.

5) Willingness to Use CAVs: Participants' willingness to use CAVs did not differ across conditions (see Table VI in Appendix C), but increased with their experience (little: 2.63; some: 3.05; much: 3.38). Post-hoc analysis revealed that all pairwise comparisons were significant, $p_{adjs} \leq .004$. Yet, the interaction of condition × experience was not significant.

V. General Discussion

In an online study with 381 participants, we obtained the effect of scenario in primary privacy measurements. Increasing



(a) **Utility-Privacy Tradeoff** (b) **Data-Sharing Decision** (c) **Confidence Rating**

Fig. 4: Results of three privacy-decision measures at Phase 2 as a function of scenario (Convenience, Safety/Security) and experience (Little, Some, Much). Error bars represent 2 standard errors. **Left panel (a)** shows that participants with much experience perceived less privacy risk than the others mainly for the convenience scenarios. **Center panel (b)** depicts that participants with much experience not only revealed more intention to share the data than the others for the convenience scenarios, but also were more willing to share the data than those with little experience for the safety/security scenarios. **Right panel (c)** shows no statistically significant difference for the confidence rating.

participants' knowledge of possible inferences from collected data (i.e., feedback) was effective in increasing their privacy awareness and helping them make more prudent privacy decisions for safety/security scenarios where data sharing seems mandatory for CAVs. Participants with much experience in connectivity and driver-assistance functions perceived more benefits and made liberal privacy decisions for convenience scenarios in which data sharing is not critical to driving.

A. Service Scenarios Influence Privacy Decisions

Generally, participants were clear about the possible risks for data collection in the safety/security scenarios and reckoned that the benefits overweighted the risks. Contrary to the larger data-sharing rate for the safety/security scenarios, only about half of the participants decided to share data for the convenience scenarios. Such a contrast implies that participants had privacy concerns about CAV data sharing, but their privacy concerns might give way to the utility when safety/security is the primary concern. When primed by possible inferences of collected data (i.e., the priming and feedback conditions), participants' perceived privacy risks became more evident for the safety/security scenarios, indicating that without extra aid users may underestimate the privacy risks of data sharing or have misconceptions [2]. These results highlight the importance of incorporating the privacy protection in mandatory data collection scenarios for privacy design of CAVs.

Our analysis focused on the purpose of scenarios. In the IoT settings, prior work identified additional factors that impacted users' privacy decisions, including the eternal entities that requested and processed the data [91], [46], as well as the data subjects [66]. For the four safety/security scenarios in our study, half (S06 & S07 in Appendix A) collected data of the pedestrians, bicyclists, and other vehicles, and the other half (S05 & S08) collected data of the driver. We did an exploratory analysis to understand the effects of data subjects (self vs. others). Data subjects showed the main effect only for all three privacy decision measures. Participants perceived higher risk when the collected data were about themselves (3.53 in benefit-risk tradeoff) than when the data were about others (3.87), $F_{(1,379)} = 63.14, p < .001, \eta_p^2 = 0.143$. They also became less likely to share the data, $F_{(1,379)} = 31.79, p < .001, \eta_p^2 = 0.077$, and were less confident in their decisions, $F_{(1,379)} = 26.35, p < .001, \eta_p^2 = 0.065$. Thus, the effects of data subjects seem to be orthogonal to the effects of condition, scenario, and experience obtained in the main analysis. Future work should consider more systematic evaluations of those factors in the CAV context, such as with whom the data is shared (e.g., automobile manufacturers or third parties).

B. Priming Is Not Sufficient for Informed Privacy Decision-Making

We included possible-inference selection as the way to prime participants on the risks of data collection in different CAV scenarios. After answering possible-inferences questions, participants gave more weight on privacy risks for the safety/security scenarios. Nevertheless, such a priming effect did not show significant impacts on participants' data-sharing decisions. The additional feedback on possible-inference selection performance led to a higher accuracy rate of the possible-inference questions, suggesting a higher privacy awareness from the participants. Moreover, participants' data-sharing decisions for the safety/security scenarios became more prudent, suggesting that the obtained liberal data-sharing decision might have been due to their lack of knowledge about the privacy implications [87], [89]. The lack of knowledge about privacy implications could also make the possible inferences open to misinterpretation [47]. Thus, investigations of better ways to define and present privacy implications of personal information in the dynamic situations of CAVs are needed.

With feedback, participants selected more correct answers to possible-inference questions and made informed privacy-related decisions. Yet, the absence of feedback in the health scenario, instead, showed an opposite pattern. One possible explanation is that participants seemed to expect feedback to remind them of the potential privacy risks and assume the absence of feedback indicated minimal risks (i.e., false negative) [87]. Due to the rapid advancements in data mining and machine learning, such feedback may not be available in

newly created service scenarios. Thus, it is critical to propose effective mechanisms for people to maintain what they have learned from the feedback when it is absent.

Alternatively, the increased data sharing in the health scenario could imply that the feedback was effective in *calibrating* participants' privacy decisions. Compared to the convenience scenario (60% in the control condition), the baseline data-sharing rate of the health scenario (41% in the control condition) was lower, indicating that participants might have concerns about sharing sensitive personal information (i.e., biometric or health information) initially. Participants in the feedback condition increased their sharing of the health scenario, implying that the knowledge participants learned through feedback may alleviate some of their concerns on sharing sensitive personal information. Thus, the feedback was not only effective in helping people make more conservative privacy decisions for the safety/security scenarios but also able to help them relieve some concerns of sensitive information sharing. Of course, this is a post-hoc explanation. Future researches need to replicate and investigate the possible calibration effect of feedback more thoroughly.

C. Individual Difference in Privacy Decisions of CAVs

Compared to participants with little experience in using connectivity and driver-assistance functions, participants with much experience perceived more benefits for the convenience scenarios and revealed more willingness to share data in general. Instead of priming privacy risks, the possible-inference questions seemed to gauge participants with much experience think more about utility, indicating the influence of experience on individuals' mental models of privacy decisions [12]. The same pattern was also reported in Brell et al.'s survey [11]. They found that participants' prior experience had a significant influence on the perceived benefits but not that of barriers (e.g., risks), suggesting that participants with more relevant experience were likely to be attracted by the obvious utility but neglect the potential privacy risks.

Such influence of prior experience indicates that it is necessary to understand users' privacy decisions of CAVs based on individual differences [81]. We mainly rely on quantitative measures investigating people's privacy decision-making in the CAV context. Future work could consider examining individuals' mental models [12] of CAV privacy by using qualitative methods such as open-ended questions [23] or interviews.

D. Uniqueness of the CAV Environment

According to the conclusion from the IoT domain, people with more knowledge about possible inferences (score higher in possible inference selection questions) tended to be more conservative when making privacy decisions. While the average possible-inference selection accuracy rate in our experiment (i.e., CAVs: 69.98%) was lower than that measured in the IoT contexts (e.g., 72% in [47]), the overall data-sharing rate in the CAVs (69.49%) was much higher than that in the IoT (e.g., 39.84% in [47]). Even if considering the convenience scenarios of CAVs only, the data-sharing rate was still 57.7%. A possible explanation is that the prior driving experience makes the users tend to perceive CAVs from the perspective of vehicles instead of cyber-physical systems, resulting in an underestimation of potential privacy risks within the CAV scenarios. Such an effect of experience reveals a unique aspect of human privacy decision-making in the CAV environment.

E. Limitations

Our participants were MTurk workers who tend to be young, more educated, and more aware of privacy issues [35]. Although our participants were diverse with regard to demographics, it may not represent the U.S. population. Our survey was conducted in the U.S., indicating the obtained results may not necessarily represent privacy awareness and decision-making of CAVs in other regions (e.g., EU or East Asia).

Despite the introduction to CAVs at the beginning of our questionnaire, it was still possible that the concept was obscure to the participants. The services described were somewhat far from real application, which could have led to difficulties for the participants to precisely evaluate and compare the benefit and potential risks. Our scenarios focused on purposes of safety/security and convenience, which only covered a small portion of CAV data-collection contexts. Due to our main interests in the data collected and processed by CAVs, we focused on the sensor data, video recording (interior and exterior), biometric or health data, in current work. Future studies could consider other service purposes, as well as the data collection and use beyond the sensor data, such as data exchange with other vehicles or the transport infrastructure [42]. We examined participants' overall privacy attitudes using IUIPC. Other instruments such as the online privacy literacy scale (OPLIS) [76] have been validated to understand inconsistency between people's privacy perception and privacy decisions. Thus, future work could apply those instruments to better understand the possible gap in the CAV context.

F. Practical Implications

Our results imply a few design recommendations for informed privacy decisions in the CAV setting. We found that participants assessed both benefits and risks in their privacy evaluation, but their privacy concerns might give way to the utility depending on the purpose of data collection. Thus, it is essential to *provide cues of data-collection purpose* (e.g., safety vs. convenience) and *explicitly communicate utility benefits and privacy risks* such that users can evaluate tradeoffs in the specific data collection scenario. Informing users of privacy implications is critical for them to reach a comprehensive decision about using a service or not in the emerging CAV context. Thus, we also recommend *explaining data collection implications to afford knowledge acquisition*. Users' prior experience with driving assistance and connectivity functions led to increased utility perception and liberal privacy decisions. *Individually tailored control strategies* should be considered during the privacy design of CAVs. The above design recommendations could serve as a baseline for the privacy design of CAVs, which have to be validated in future work.

REFERENCES

[1] A. Acquisti and J. Grossklags, "Privacy attitudes and privacy behavior," in *Economics of information security*. Springer, 2004, pp. 165–178.

[2] A. Adams, "The implications of users' multimedia privacy perceptions on communication and information privacy policies," in *Proceedings of Telecommunications Policy Research Conference*, 1999, pp. 1–23.

[3] C. B. ALA'A AL-MOMANI and F. KARGL, "A comparison of data protection regulations for automotive systems," *Data Protection and Privacy: Data Protection and Artificial Intelligence*, p. 57, 2021.

[4] I. Altman, *The environment and social behavior: privacy, personal space, territory, and crowding.* Brooks/Cole Publishing Co., 1975.

[5] P. Bahirat, Y. He, A. Menon, and B. Knijnenburg, "A data-driven approach to developing iot privacy-setting interfaces," in *23rd International Conference on Intelligent User Interfaces*, 2018, pp. 165–176.

[6] M. S. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan, "Recognizing facial expression: machine learning and application to spontaneous behavior," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 2. IEEE, 2005, pp. 568–573.

[7] D. Bates, M. Maechler, B. Bolker, S. Walker, R. H. B. Christensen, H. Singmann, B. Dai, F. Scheipl, and G. Grothendieck, "Package 'lme4'," *Linear mixed-effects models using S4 classes. R package version*, vol. 1, no. 6, 2011.

[8] C. Bloom, J. Tan, J. Ramjohn, and L. Bauer, "Self-driving cars and data collection: Privacy perceptions of networked autonomous vehicles," in *Thirteenth Symposium on Usable Privacy and Security (SOUPS)*, 2017, pp. 357–375.

[9] M. Brauer and J. J. Curtin, "Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items." *Psychological Methods*, vol. 23, no. 3, p. 389, 2018.

[10] T. Brell, H. Biermann, R. Philipsen, and M. Ziefle, "Conditional privacy: Users' perception of data privacy in autonomous driving." in *VEHITS*, 2019, pp. 352–359.

[11] T. Brell, R. Philipsen, and M. Ziefle, "scary! risk perceptions in autonomous driving: The influence of experience on perceived benefits and barriers," *Risk Analysis*, vol. 39, no. 2, pp. 342–357, 2019.

[12] L. J. Camp, "Mental models of privacy and security," *IEEE Technology and Society Magazine*, vol. 28, no. 3, pp. 37–46, 2009.

[13] E. Çano, R. Coppola, E. Gargiulo, M. Marengo, and M. Morisio, "Mood-based on-car music recommendations," in *International Conference on Industrial Networks and Intelligent Systems*. Springer, 2016, pp. 154–163.

[14] I. Chong, H. Ge, N. Li, and R. W. Proctor, "Influence of privacy priming and security framing on mobile app selection," *Computers & Security*, vol. 78, pp. 143–154, 2018.

[15] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1. Ieee, 2005, pp. 886–893.

[16] C. Deng, C. Wu, N. Lyu, and Z. Huang, "Driving style recognition method using braking characteristics based on hidden markov model," *PloS One*, vol. 12, no. 8, p. e0182419, 2017.

[17] S. Derikx, M. De Reuver, and M. Kroesen, "Can privacy concerns for insurance of connected cars be compensated?" *Electronic Markets*, vol. 26, no. 1, pp. 73–81, 2016.

[18] S. Egelman and E. Peer, "Predicting privacy and security attitudes," *ACM SIGCAS Computers and Society*, vol. 45, no. 1, pp. 22–28, 2015.

[19] P. Emami-Naeini, S. Bhagavatula, H. Habib, M. Degeling, L. Bauer, L. F. Cranor, and N. Sadeh, "Privacy expectations and preferences in an iot world," in *Thirteenth Symposium on Usable Privacy and Security (SOUPS)*, 2017, pp. 399–412.

[20] P. Emami-Naeini, H. Dixon, Y. Agarwal, and L. F. Cranor, "Exploring how privacy and security factor into iot device purchase behavior," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–12.

[21] J. Finch, "The vignette technique in survey research," *Sociology*, vol. 21, no. 1, pp. 105–114, 1987.

[22] D. Frassinelli, S. Park, and S. Nürnberger, "I know where you parked last summer: Automated reverse engineering and privacy analysis of modern cars," in *2020 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2020, pp. 1401–1415.

[23] L. T. Gröber, M. Fassl, A. Gupta, and K. Krombholz, "Investigating car drivers' information demand after safety and security critical incidents," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–17.

[24] T. Groß, "Validity and reliability of the scale internet users' information privacy concerns (iuipc)," *Proceedings on Privacy Enhancing Technologies*, 2021.

[25] J. Guanetti, Y. Kim, and F. Borrelli, "Control of connected and automated vehicles: State of the art and future challenges," *Annual Reviews in Control*, vol. 45, pp. 18–40, 2018.

[26] J. Hainmueller, D. Hangartner, and T. Yamamoto, "Validating vignette and conjoint survey experiments against real-world behavior," *Proceedings of the National Academy of Sciences*, vol. 112, no. 8, pp. 2395–2400, 2015.

[27] M. Harbach, M. Hettig, S. Weber, and M. Smith, "Using personal examples to improve risk communication for security & privacy decisions," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014, pp. 2647–2656.

[28] D. J. Hauser and N. Schwarz, "Attentive turkers: Mturk participants perform better on online attention checks than do subject pool participants," *Behavior Research Methods*, vol. 48, no. 1, pp. 400–407, 2016.

[29] J. Hays and A. A. Efros, "Im2gps: estimating geographic information from a single image," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2008, pp. 1–8.

[30] J. Hong, "The privacy landscape of pervasive computing," *IEEE Pervasive Computing*, vol. 16, no. 3, pp. 40–48, 2017.

[31] D. Howard and D. Dai, "Public perceptions of self-driving cars: The case of berkeley, california," in *Transportation Research Board 93rd Annual Meeting*, vol. 14, no. 4502, 2014, pp. 1–16.

[32] D. Huber, H. Herman, A. Kelly, P. Rander, and J. Ziglar, "Real-time photo-realistic visualization of 3d environments for enhanced tele-operation of vehicles," in *2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops*. IEEE, 2009, pp. 1518–1525.

[33] S. Johnsen and A. Tews, "Real-time object tracking and classification using a static camera," in *Proceedings of IEEE International Conference on Robotics and Automation, workshop on People Detection and Tracking*. CiteSeer, 2009.

[34] S.-G. Jung, J. An, H. Kwak, J. Salminen, and B. J. Jansen, "Assessing the accuracy of four popular face recognition tools for inferring gender, age, and race," in *Twelfth International AAAI Conference on Web and Social Media*, 2018.

[35] R. Kang, S. Brown, L. Dabbish, and S. Kiesler, "Privacy attitudes of mechanical turk workers and the us public," in *10th Symposium On Usable Privacy and Security (SOUPS)*, 2014, pp. 37–49.

[36] P. G. Kelley, L. F. Cranor, and N. Sadeh, "Privacy as part of the app decision-making process," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 3393–3402.

[37] M.-K. Kim, J.-H. Park, J. Oh, W.-S. Lee, and D. Chung, "Identifying and prioritizing the benefits and concerns of connected and autonomous vehicles: A comparison of individual and expert perceptions," *Research in Transportation Business & Management*, vol. 32, p. 100438, 2019.

[38] W. Kong, L. Zhou, Y. Wang, J. Zhang, J. Liu, and S. Gao, "A system of driving fatigue detection based on machine vision and its application on smart device," *Journal of Sensors*, vol. 2015, 2015.

[39] A. Krause and E. Horvitz, "A utility-theoretic approach to privacy and personalization." in *AAAI*, vol. 8, 2008, pp. 1181–1188.

[40] T. Krismayer, M. Schedl, P. Knees, and R. Rabiser, "Predicting user demographics from music listening information," *Multimedia Tools and Applications*, vol. 78, no. 3, pp. 2897–2920, 2019.

[41] J. Kröger, "Unexpected inferences from sensor data: a hidden privacy threat in the internet of things," in *IFIP International Internet of Things Conference*. Springer, 2018, pp. 147–159.

[42] I. Krontiris, T. Giannetsos, P. Schoo, and F. Kargl, *Buckle-up: autonomous vehicles could face privacy bumps in the road ahead*. Ruhr-Universität Bochum, 2020.

[43] J. Ku, M. Mozifian, J. Lee, A. Harakeh, and S. L. Waslander, "Joint 3d proposal generation and object detection from view aggregation," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 1–8.

[44] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, pp. 159–174, 1977.

[45] C. Lee, C. Ward, M. Raue, L. D'Ambrosio, and J. F. Coughlin, "Age differences in acceptance of self-driving cars: A survey of perceptions

and attitudes," in *International Conference on Human Aspects of IT for the Aged Population*. Springer, 2017, pp. 3–13.
[46] H. Lee and A. Kobsa, "Understanding user privacy in internet of things environments," in *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)*. IEEE, 2016, pp. 407–412.
[47] ——, "Confident privacy decision-making in iot environments," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 27, no. 1, pp. 1–39, 2019.
[48] J.-F. J. Lee and J. Williams, "New way to utilize remote sensing data: Automated road travel survey," *Transportation Research Record*, vol. 2460, no. 1, pp. 15–21, 2014.
[49] H. Li, H. Zhu, S. Du, X. Liang, and X. Shen, "Privacy leakage of location sharing in mobile social networks: Attacks and defense," *IEEE Transactions on Dependable and Secure Computing*, vol. 15, no. 4, pp. 646–660, 2016.
[50] J. Lin, S. Amini, J. I. Hong, N. Sadeh, J. Lindqvist, and J. Zhang, "Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, 2012, pp. 501–510.
[51] T. Litman, "Autonomous vehicle implementation predictions: Implications for transport planning," 2020.
[52] ——, "Autonomous vehicle implementation predictions: Implications for transport planning," 2021.
[53] R. D. Luce and J. W. Tukey, "Simultaneous conjoint measurement: A new type of fundamental measurement," *Journal of Mathematical Psychology*, vol. 1, no. 1, pp. 1–27, 1964.
[54] Y. Ma, Z. Li, Y. Li, H. Li, and R. Malekian, "Driving style estimation by fusing multiple driving behaviors: a case study of freeway in china," *Cluster Computing*, vol. 22, no. 4, pp. 8259–8269, 2019.
[55] 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*, vol. 15, no. 4, pp. 336–355, 2004.
[56] G. Muhammad, S. M. M. Rahman, A. Alelaiwi, and A. Alamri, "Smart health solution integrating iot and cloud: A case study of voice pathology monitoring," *IEEE Communications Magazine*, vol. 55, no. 1, pp. 69–73, 2017.
[57] Y. L. Murphey, R. Milton, and L. Kiliaris, "Driver's style classification using jerk analysis," in *2009 IEEE Workshop on Computational Intelligence in Vehicles and Vehicular Systems*. IEEE, 2009, pp. 23–28.
[58] S. R. Narla, "The evolution of connected vehicle technology: From smart drivers to smart cars to... self-driving cars," *Ite Journal*, vol. 83, no. 7, pp. 22–26, 2013.
[59] N. Ostern, A. Eßer, and P. Buxmann, "Capturing users' privacy expectations to design better smart car applications." in *PACIS*, 2018, p. 97.
[60] S. Parkinson, P. Ward, K. Wilson, and J. Miller, "Cyber threats facing autonomous and connected vehicles: Future challenges," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 11, pp. 2898–2915, 2017.
[61] S. Patil, R. Hoyle, R. Schlegel, A. Kapadia, and A. J. Lee, "Interrupt now or inform later? comparing immediate and delayed privacy feedback," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015, pp. 1415–1418.
[62] I. S. Penton-Voak, N. Pound, A. C. Little, and D. I. Perrett, "Personality judgments from natural and composite facial images: More evidence for a "kernel of truth" in social perception," *Social Cognition*, vol. 24, no. 5, pp. 607–640, 2006.
[63] W. M. Petrusic and J. V. Baranski, "Judging confidence influences decision processing in comparative judgments," *Psychonomic Bulletin & Review*, vol. 10, no. 1, pp. 177–183, 2003.
[64] S. Pötzsch, "Privacy awareness: A means to solve the privacy paradox?" in *IFIP Summer School on the Future of Identity in the Information Society*. Springer, 2008, pp. 226–236.
[65] P. Rajivan and J. Camp, "Influence of privacy attitude and privacy cue framing on android app choices," in *Twelfth Symposium on Usable Privacy and Security (SOUPS)*, 2016.
[66] P. A. Rauschnabel, J. He, and Y. K. Ro, "Antecedents to the adoption of augmented reality smart glasses: A closer look at privacy risks," *Journal of Business Research*, vol. 92, pp. 374–384, 2018.
[67] P. J. Rentfrow and S. D. Gosling, "The do re mi's of everyday life: the structure and personality correlates of music preferences." *Journal of Personality and Social Psychology*, vol. 84, no. 6, p. 1236, 2003.
[68] C. Rödel, S. Stadler, A. Meschtscherjakov, and M. Tscheligi, "Towards autonomous cars: The effect of autonomy levels on acceptance and user experience," in *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 2014, pp. 1–8.
[69] S. Sachdev, J. Macwan, C. Patel, and N. Doshi, "Voice-controlled autonomous vehicle using iot," *Procedia Computer Science*, vol. 160, pp. 712–717, 2019.
[70] L. Sandt and J. M. Owens, "Discussion guide for automated and connected vehicles, pedestrians, and bicyclists." 2017.
[71] F. Schaub, B. Könings, and M. Weber, "Context-adaptive privacy: Leveraging context awareness to support privacy decision making," *IEEE Pervasive Computing*, vol. 14, no. 1, pp. 34–43, 2015.
[72] R. Schlegel, A. Kapadia, and A. J. Lee, "Eyeing your exposure: quantifying and controlling information sharing for improved privacy," in *Proceedings of the Seventh Symposium on Usable Privacy and Security*, 2011, pp. 1–14.
[73] H. J. Smith, T. Dinev, and H. Xu, "Information privacy research: an interdisciplinary review," *MIS Quarterly*, pp. 989–1015, 2011.
[74] M.-H. Su, C.-H. Wu, K.-Y. Huang, Q.-B. Hong, and H.-M. Wang, "Personality trait perception from speech signals using multiresolution analysis and convolutional neural networks," in *2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. IEEE, 2017, pp. 1532–1536.
[75] M. Swan, "Connected car: quantified self becomes quantified car," *Journal of Sensor and Actuator Networks*, vol. 4, no. 1, pp. 2–29, 2015.
[76] S. Trepte, D. Teutsch, P. K. Masur, C. Eicher, M. Fischer, A. Hennhöfer, and F. Lind, "Do people know about privacy and data protection strategies? towards the "online privacy literacy scale"(oplis)," in *Reforming European data protection law*. Springer, 2015, pp. 333–365.
[77] J. Y. Tsai, P. Kelley, P. Drielsma, L. F. Cranor, J. Hong, and N. Sadeh, "Who's viewed you? the impact of feedback in a mobile location-sharing application," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2009, pp. 2003–2012.
[78] I. P. Tussyadiah, F. J. Zach, and J. Wang, "Attitudes toward autonomous on demand mobility system: The case of self-driving taxi," in *Information and communication technologies in tourism 2017*. Springer, 2017, pp. 755–766.
[79] R. Uittenbogaard, C. Sebastian, J. Vijverberg, B. Boom, D. M. Gavrila *et al.*, "Privacy protection in street-view panoramas using depth and multi-view imagery," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 10 581–10 590.
[80] A. C. P. Uy, R. A. Bedruz, A. R. Quiros, A. Bandala, and E. P. Dadios, "Machine vision for traffic violation detection system through genetic algorithm," in *2015 International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*. IEEE, 2015, pp. 1–7.
[81] M. Volkamer and K. Renaud, "Mental models–general introduction and review of their application to human-centred security," in *Number Theory and Cryptography*. Springer, 2013, pp. 255–280.
[82] V. Wan and W. M. Campbell, "Support vector machines for speaker verification and identification," in *Neural Networks for Signal Processing X. Proceedings of the 2000 IEEE Signal Processing Society Workshop (Cat. No. 00TH8501)*, vol. 2. IEEE, 2000, pp. 775–784.
[83] D. Wang, M. Pei, and L. Zhu, "Detecting driver use of mobile phone based on in-car camera," in *2014 Tenth International Conference on Computational Intelligence and Security*. IEEE, 2014, pp. 148–151.
[84] Z.-Q. Wang and I. Tashev, "Learning utterance-level representations for speech emotion and age/gender recognition using deep neural networks," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, pp. 5150–5154.
[85] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A discriminative feature learning approach for deep face recognition," in *European Conference on Computer Vision*. Springer, 2016, pp. 499–515.
[86] G. Wusk and H. Gabler, "Non-invasive detection of respiration and heart rate with a vehicle seat sensor," *Sensors*, vol. 18, no. 5, p. 1463, 2018.
[87] A. Xiong, R. W. Proctor, W. Yang, and N. Li, "Embedding training within warnings improves skills of identifying phishing webpages," *Human Factors*, vol. 61, no. 4, pp. 577–595, 2019.
[88] A. Xiong, T. Wang, N. Li, and S. Jha, "Towards effective differential privacy communication for users' data sharing decision and comprehension," *IEEE Symposium on Security and Privacy (S&P)*, 2020.
[89] W. Yang, A. Xiong, J. Chen, R. W. Proctor, and N. Li, "Use of phishing training to improve security warning compliance: evidence from a field experiment," in *Proceedings of the Hot Topics in Science of Security: Symposium and Bootcamp*, 2017, pp. 52–61.

[90] Y.-H. Yang, Y.-C. Lin, Y.-F. Su, and H. H. Chen, "A regression approach to music emotion recognition," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, no. 2, pp. 448–457, 2008.
[91] S. Zheng, N. Apthorpe, M. Chetty, and N. Feamster, "User perceptions of smart home iot privacy," *Proceedings of the ACM on Human-Computer Interaction*, vol. 2, no. CSCW, pp. 1–20, 2018.
[92] X. Zhou, S. Demetriou, D. He, M. Naveed, X. Pan, X. Wang, C. A. Gunter, and K. Nahrstedt, "Identity, location, disease and more: Inferring your secrets from android public resources," in *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security*, 2013, pp. 1017–1028.
[93] L. Zhu, F. R. Yu, Y. Wang, B. Ning, and T. Tang, "Big data analytics in intelligent transportation systems: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 1, pp. 383–398, 2018.

APPENDIX

A. Scenarios for Privacy Decision-Making

In this appendix, we list the textual descriptions and factor values of all scenarios (S). Possible and impossible inferences are used as options in the inference-selection question of each scenario. S01 to S04 are convenience-related, and S05 to S08 are safety- or security-related. S09 is for the attention check and S10 is health-related.

S01. When you are driving the connected autonomous vehicle (CAV), the camera inside the CAV collects **your photos** to infer your **mood**, thereby allowing improvement of your driving experience along the journey, such as adjusting the ambient light or inside temperature, for your **convenience**.

- Purpose: Convenience
- Core Inference:
  - Photos ⇒ Your mood
- Possible Inference(s):
  - Photos ⇒ Your demographics, such as age, gender, race
  - Photos ⇒ Your non-driving activity, such as sleeping, reading
  - Photos ⇒ Your personality types, such as openness, agreeableness
  - Photos ⇒ Your identity
- Impossible Inference(s):
  - Photos ⇒ Your whereabouts

S02. When you are driving the connected autonomous vehicle (CAV), the camera outside the CAV collects **photos of road scene** to improve **scene understanding**, thereby allowing advanced analysis and decision making of the CAV, for your **convenience**.

- Purpose: Convenience
- Core Inference:
  - Photos ⇒ Road scene
- Possible Inference(s):
  - Photos ⇒ Your frequency of visits on some location, such as grocery store, shopping mall
  - Photos ⇒ Your whereabouts
  - Photos ⇒ Your income level
  - Photos ⇒ Your choice preference, such as restaurant, grocery store
- Impossible Inference(s):
  - Photos ⇒ Your mood

S03. When you are driving the connected autonomous vehicle (CAV), the infotainment system inside the CAV collects **your playlists** to infer your **music preference**, thereby allowing music recommendation, for your **convenience**.

- Purpose: Convenience
- Core Inference:
  - Playlists ⇒ Your music preference
- Possible Inference(s):
  - Playlists ⇒ Your demographics, such as age, gender, race
  - Playlists ⇒ Your mood, such as happy, sad
  - Playlists ⇒ Your personality types, such as openness, agreeableness
- Impossible Inference(s):
  - Playlists ⇒ Your income level
  - Playlists ⇒ Your whereabouts

S04. When you enter the connected autonomous vehicle (CAV), the camera inside the CAV collects **your photos** to infer your **identity**, thereby loading your personal setting of the vehicle, for your **convenience**.

- Purpose: Convenience
- Core Inference:
  - Photos ⇒ Your identity
- Possible Inference(s):
  - Photos ⇒ Your demographics, such as age, gender, race
  - Photos ⇒ Your mood, such as happy, sad
  - Photos ⇒ Your personality types, such as openness, agreeableness
- Impossible Inference(s):
  - Photos ⇒ Your whereabouts
  - Photos ⇒ Your accident history

S05. When you enter the connected autonomous vehicle (CAV), the voice control inside the CAV collects **your voice** to infer your **identity**, thereby authorizing you to control the vehicle with voice commands, for your **security**.

- Purpose: Security
- Core Inference:
  - Voice ⇒ Your identity
- Possible Inference(s):
  - Voice ⇒ Your demographics such as age, gender, race
  - Voice ⇒ Your mood, such as happy, sad
  - Voice ⇒ Your personality types, such as openness, agreeableness
  - Voice ⇒ Your physical status, such as healthy, ill health
- Impossible Inference(s):
  - Voice ⇒ Your income level

S06. When you are driving the connected autonomous vehicle (CAV), the camera and LiDAR outside the CAV collects **2D and 3D photos** of other vehicles on the road, thereby allowing prediction of the other vehicles' trajectory and plan the CAV's next action, for your **safety**.

- Purpose: Safety
- Core Inference:
  - Photos ⇒ Other vehicles' trajectories

- Possible Inference(s):
  - Photos ⇒ Other vehicles' presence on some location
  - Photos ⇒ Other vehicles' whereabouts
  - Photos ⇒ Traffic rule violation of other vehicles, such as red light violation
  - Photos ⇒ Driving Style of other vehicles such as aggressive driving
- Impossible Inference(s):
  - Photos ⇒ **Mood of other vehicles' drivers**

S07. When you are driving the connected autonomous vehicle (CAV), the camera and LiDAR outside the CAV collects **2D and 3D photos of the environment** to enable object detection, such as **bicyclists or pedestrians**, thereby allowing prediction of **object's trajectory** and plan the vehicle's next action, for your **safety**.

- Purpose: Safety
- Core Inference:
  - Photos ⇒ Bicyclists or pedestrians' trajectories
- Possible Inference(s):
  - Photos ⇒ Demographics of bicyclists or pedestrians, such as age, gender, race
  - Photos ⇒ bicyclists or pedestrians' presence on some location
  - Photos ⇒ Social relationship of bicyclists or pedestrians
  - Photos ⇒ Identity of bicyclists or pedestrians
- Impossible Inference(s):
  - Photo ⇒ Political view of bicyclists or pedestrians

S08. When you are driving the connected autonomous vehicle (CAV), the CAV collects the **on-board diagnostic data** to let you track, monitor or share data of **your vehicle condition** with vehicle maintenance, for your **safety**.

- Purpose: Safety
- Core Inference:
  - OBD Data ⇒ Your vehicle condition
- Possible Inference(s):
  - OBD Data ⇒ Your CAV use pattern, such as the number of occupants
  - OBD Data ⇒ Your frequency of driving
  - OBD Data ⇒ Your driving style, such as aggressive driving
  - OBD Data ⇒ Your personal settings and preference, such as inside temperature, driving speed
- Impossible Inference(s):
  - OBD Data ⇒ Your mood, such as happy, sad

[ATTENTION CHECK]

S09. When you are driving the connected autonomous vehicle (CAV), the camera outside the CAV collects **the photos of other vehicles' license plate**to infer their **identities**, thereby allowing to identify aggressive vehicles (or drivers), for your **safety**. Please **ignore the questions** and **only select the 2nd option** as your answers for all the following questions in this scenario. With the help of your responses to this scenario, you show us that you have read the scenario descriptions.

- Core Inference:
  - Photos ⇒ Other vehicles' identities
- Possible Inference(s):
  - Photos ⇒ Other vehicles' accidental history
  - Photos ⇒ Other vehicles' presence
  - Photos ⇒ Other vehicles' toll violation history
- Impossible Inference(s):
  - Photos ⇒ Other drivers' personality types
  - Photos ⇒ Other drivers' social relationship

[HEALTH SCENARIO]

S10. When you are driving the connected autonomous vehicle (CAV), the sensors inside the CAV collect **biometric data**, such as your body temperature, heart rate, to infer your **physical status**, thereby suggesting adjustments of the vehicle speed, inside temperature and others, for your **health**.

- Purpose: health
- Core Inference:
  - Biometrics ⇒ Your physical status

B. Survey Protocol

Instructions are bold.

1) Phase 1: **[QUALIFICATION] Please read the following description of CAVs closely, and answer three questions.**

CAVs are connected vehicles that have self-driving capabilities. The CAVs use reliable low latency wireless network, such as 5G, and a wide range of sensors, such as internal and external cameras, Lidar, Radar, Ultrasound sensors, GPS, to obtain relevant traffic and other information inside and outside the vehicle. At the same time, the CAVs' driving control occurs without direct input from the driver. For example, when a CAV breaks suddenly, it can transmit a notice to vehicles behind that enables those vehicles to warn their drivers to stop, or automatically apply brakes if a crash is imminent.

Q1. CAVs use wireless communication to share information about safety, the infrastructure, and other road users, such as pedestrians and bicyclists.

- ○ False
- ○ True
- ○ Prefer not to answer

Q2. Drivers still need to monitor the CAVs all the times in case there is a fatal error during operation.

- ○ False
- ○ True
- ○ Prefer not to answer

Q3. CAV is a term used to describe vehicles that are both connected and automated.

- ○ False
- ○ True
- ○ Prefer not to answer

2) Phase 2: **Suppose that you are driving one of the CAVs. Next, we will present ten different service scenarios of the CAVs. For each scenario, please 1) read the description carefully (we set a minimum viewing time for the scenario description); 2) then answer a few questions about it.**

[SCENARIOS 1-10] When you are driving the connected autonomous vehicle (CAV), the [device or sensor of] the CAV collects [data] to infer [core inference], thereby [purpose explanation], for your [purpose].

Q4. Besides the core inference (Photos ⇒ Your mood), please choose other possible inferences that you believe to be true based on the collected data (check all that apply):

- ☐ Photos ⇒ Your demographics, such as age, gender, race
- ☐ Photos ⇒ Your non-driving activities, such as sleeping, reading
- ☐ Photos ⇒ Your personality types, such as openness, agreeableness
- ☐ Photos ⇒ Your identity
- ☐ Photos ⇒ Your whereabouts
- ☐ Prefer not to answer

[Q4 QUESTION AND OPTIONS ARE BASED ON S01.]

Q5. How do you rate the balance between the utility benefits and privacy risks in the scenario?

- ○ Benefits are much less than risks
- ○ Benefits are less than risks
- ○ Benefits are almost equal to risks
- ○ Benefits are greater than risks
- ○ Benefits are much greater than risks
- ○ Prefer not to answer

Q6. Would you share the data to use the service described in the scenario?

- ○ No
- ○ Yes
- ○ Prefer not to answer

Q7. How confident are you in the above data-sharing decision?

- ○ Very unconfident
- ○ Unconfident
- ○ Neutral
- ○ Confident
- ○ Very confident
- ○ Prefer not to answer

3) Phase 3: **This is the post-session questionnaire. First of all, please answer 11 questions about your own opinions on privacy issues in general.**

[IUIPC]

Q8. Companies seeking information should disclose the way the data are collected, processed, and used.

- ○ Strongly disagree
- ○ Disagree
- ○ Neither agree nor disagree
- ○ Agree
- ○ Strongly agree
- ○ Prefer not to answer

Q9. A good privacy policy should have a clear and conspicuous disclosure.

- ○ Strongly disagree
- ○ Disagree
- ○ Neither agree nor disagree
- ○ Agree
- ○ Strongly agree
- ○ Prefer not to answer

Q10. It is very important to me that I am aware and knowledgeable about how my personal information will be used.

- ○ Strongly disagree
- ○ Disagree
- ○ Neither agree nor disagree
- ○ Agree
- ○ Strongly agree
- ○ Prefer not to answer

Q11. It usually bothers me when companies ask me for personal information.

- ○ Strongly disagree
- ○ Disagree
- ○ Neither agree nor disagree
- ○ Agree
- ○ Strongly agree
- ○ Prefer not to answer

Q12. When companies ask me for personal information, I sometimes think twice before providing it.

- ○ Strongly disagree
- ○ Disagree
- ○ Neither agree nor disagree
- ○ Agree
- ○ Strongly agree
- ○ Prefer not to answer

Q13. It bothers me to give so many personal information to so many companies.

- ○ Strongly disagree
- ○ Disagree
- ○ Neither agree nor disagree
- ○ Agree
- ○ Strongly agree
- ○ Prefer not to answer

Q14. I'm concerned that companies are collecting too much personal information about me.

- ○ Strongly disagree
- ○ Disagree
- ○ Neither agree nor disagree
- ○ Agree
- ○ Strongly agree
- ○ Prefer not to answer

Q15. Online companies should not use personal information for any purpose unless it has been authorized by the individuals who provided information.

- ○ Strongly disagree
- ○ Disagree
- ○ Neither agree nor disagree
- ○ Agree
- ○ Strongly agree
- ○ Prefer not to answer

Q16. When people give personal information to an online company for some reason, the online company should never use the information for any other reason.
- ○ Strongly disagree
- ○ Disagree
- ○ Neither agree nor disagree
- ○ Agree
- ○ Strongly agree
- ○ Prefer not to answer

Q17. Online companies should never sell the personal information in their computer databases to other companies.
- ○ Strongly disagree
- ○ Disagree
- ○ Neither agree nor disagree
- ○ Agree
- ○ Strongly agree
- ○ Prefer not to answer

Q18. Online companies should never share personal information with other companies unless it has been authorized by the individuals who provided the information.
- ○ Strongly disagree
- ○ Disagree
- ○ Neither agree nor disagree
- ○ Agree
- ○ Strongly agree
- ○ Prefer not to answer

In the end, please answer questions about your demographic information and driving related experience.

[DEMOGRAPHICS]

Q19. What's your gender?
- ○ Male
- ○ Female
- ○ Other
- ○ Prefer not to answer

Q20. What's your age?
- ○ 18 - 24
- ○ 25 - 34
- ○ 35 - 44
- ○ 45 - 54
- ○ 55 or older
- ○ Prefer not to answer

Q21. What's your ethnicity?
- ○ American Indian / Alaska Native
- ○ African / African American
- ○ Native Hawaii / Pacific Islander
- ○ Hispanic / Latino
- ○ Caucasian
- ○ Asian
- ○ More than one race
- ○ Other / Unknown
- ○ Prefer not to answer

Q22. What is your highest degree you have earned?
- ○ No high school degree
- ○ High school degree
- ○ College degree
- ○ Associate degree
- ○ Bachelors
- ○ Professional degree (masters / Ph.D.)
- ○ Medical degree
- ○ Prefer not to answer

Q23. Are you majoring in or have a degree or job in computer science, computer engineering, information technology, or a related field?
- ○ No
- ○ Yes
- ○ Prefer not to answer

[DRIVING EXPERIENCE]

Q24. Do you have a valid driver's license?
- ○ No
- ○ Yes
- ○ Prefer not to answer

Q25. What is your average mileages per year?
- ○ $<$ 2,000 miles
- ○ 2,000 - 5,000 miles
- ○ 5,000 - 10,000 miles
- ○ 10,000 - 20,000 miles
- ○ $>$ 20,000 miles
- ○ Prefer not to answer

[EXPERIENCE WITH DRIVING ASSISTANCE AND CONNECTIVITY FUNCTIONS, ACCEPTANCE OF CAV]

Q26. Have you ever used connectivity functions inside the vehicles, such as Google Android Auto, Apple CarPlay, GM OnStar, or Ford SYNC?
- ○ No, not at all
- ○ No, rarely
- ○ Yes, sometimes
- ○ Yes, quite often
- ○ Prefer not to answer

Q27. Have you ever used driving assistance functions, such as automatic parking, cruise control or adaptive cruise control (ACC)?
- ○ No, not at all
- ○ No, rarely
- ○ Yes, sometimes
- ○ Yes, quite often
- ○ Prefer not to answer

Q28. If CAVs are available in the near future, please indicate your willingness to use CAVs:
- ○ No, never
- ○ No, rarely
- ○ Yes, for some cases
- ○ Yes, for sure
- ○ Prefer not to answer

C. Tables

In this appendix, we list the descriptive and inferential statistics of dependent measures.

TABLE IV: ANOVA Results for Three Privacy-Decision Measures (Utility-Privacy Tradeoff, Data-Sharing Decision, and Confidence Rating) at Phase 2 as a Function of Scenario (Safety/Security, Convenience) as a Within-Subject Factor and Condition (Control, Priming, Feedback) as a Between-Subject Factor

Scenario	Effect	Utility-Privacy Tradeoff			Data-Sharing Decision			Confidence Rating		
		F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
Safety/Security & Convenience	Scenario (1, 378)	**264.49**	**<.001**	**.412**	**283.99**	**<.001**	**.429**	1.15	.284	.003
	Condition (2, 378)	**4.93**	**.008**	**.025**	2.46	.086	.013	<1.0		
	Scenario × Condition (2, 378)	**5.05**	**.007**	**.026**	**3.62**	**.028**	**.019**	2.89	.057	.015
		F	p	η_p^2	χ^2	p		F	p	η_p^2
Health	Condition (2, 378)	2.49	.084	.013	**8.83**	**.012**		<1.0		

TABLE V: ANOVA Results of Privacy-Decision Measures (Utility-Privacy Tradeoff, Data-Sharing Decision, and Confidence Rating) at Phase 2 with Experience as An Extra Factor

Scenarios	Effect	Utility-Privacy Tradeoff			Data-Sharing Decision			Confidence Rating		
		F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
Safety/Security & Convenience	Scenario(1, 372)	**276.62**	**<.001**	**.426**	**292.65**	**<.001**	**.44**	<1.0		
	Condition(2, 372)	**5.28**	**.005**	**.028**	2.71	.068	.014	1.094	.336	.006
	Experience(2, 372)	1.66	.191	.009	**8.44**	**<.001**	**.043**	<1.0		
	Scenario × Condition(2, 372)	**4.43**	**.013**	**.023**	**3.36**	**.036**	**.018**	2.06	.129	.011
	Experience × Scenario(2, 372)	**7.49**	**<.001**	**.039**	**3.94**	**.02**	**.021**	<1.0		
	Experience × Condition(4, 372)	<1.0			<1.0			<1.0		
	Experience × Scenario × Condition(4, 372)	<1.0			2.33	.055	.024	1.54	.19	.016
		F	p	η_p^2	χ^2	p		F	p	η_p^2
Health	Condition(2, 372)	1.9	.151	.010	**8.83**	**.012**		<1.0		
	Experience(2, 372)	**4.7**	**.01**	**.025**	**29.95**	**<.001**		2.74	.066	.014
	Condition × Experience(4, 372)	1.1	.366	.011	<1.0			<1.0		

TABLE VI: Analyses of Variances Summary of General Privacy Attitude and Willingness to Use CAVs at Phase 3

Effect	General Privacy Attitude			Willingness to Use CAVs		
	F	p	η_p^2	F	p	η_p^2
Condition(2, 372)	**3.10**	**.046**	**.016**	<1.0		
Experience(2, 372)	**3.25**	**.040**	**.017**	**21.72**	**<.001**	**.105**
Condition × Experience(4, 372)	<1.0			<1.0		

TABLE VII: Results of Three Privacy-Decision Measures (Utility-Privacy Tradeoff, Data-Sharing Decision, and Confidence Rating) in Phase 2 as a Function of Condition (Feedback, Priming, and Control), Scenario (Safety/Security, Convenience), and Experience (Little, Some, Much)

Experience	Condition	Scenarios	Utility-Privacy Tradeoff	Data-Sharing Decision	Confidence Rating
Little(106)	Feedback(40)	Convenience	2.76(0.15)	0.48(0.05)	4.28(0.10)
		Safety/Security	3.50(0.15)	0.72(0.04)	4.19(0.10)
	Priming(34)	Convenience	2.82(0.16)	0.40(0.06)	4.22(0.11)
		Safety/security	3.67(0.16)	0.72(0.05)	4.27(0.11)
	Control(32)	Convenience	3.07(0.16)	0.56(0.06)	4.12(0.12)
		Safety/security	3.95(0.17)	0.81(0.05)	4.09(0.11)
Some(142)	Feedback(38)	Convenience	2.89(0.15)	0.54(0.05)	4.01(0.11)
		Safety/Security	3.53(0.16)	0.74(0.04)	3.93(0.10)
	Priming(46)	Convenience	2.93(0.14)	0.53(0.05)	4.27(0.10)
		Safety/security	3.71(0.14)	0.83(0.04)	4.23(0.10)
	Control(58)	Convenience	3.05(0.12)	0.60(0.04)	4.03(0.09)
		Safety/security	3.90(0.13)	0.88(0.04)	4.25(0.08)
Much(133)	Feedback(47)	Convenience	3.21(0.14)	0.68(0.05)	4.13(0.10)
		Safety/Security	3.43(0.14)	0.80(0.04)	4.13(0.09)
	Priming(48)	Convenience	3.21(0.14)	0.71(0.05)	4.12(0.10)
		Safety/security	3.60(0.14)	0.84(0.04)	4.19(0.09)
	Control(38)	Convenience	3.34(0.15)	0.63(0.05)	4.13(0.11)
		Safety/security	4.11(0.16)	0.92(0.04)	4.17(0.10)

Note. The number in the parentheses of the first two columns indicates the number of participants for each condition. The number in the parentheses of the last three columns shows the standard errors of corresponding cell.

TABLE VIII: Results of Three Privacy-Decision Measures in Phase 2 as a Function of Condition (Feedback, Priming, and Control) and Scenario (Safety/Security, Convenience)

Condition	Scenario	Utility-Privacy Tradeoff	Data-Sharing Decision	Confidence Rating
Feedback(125)	Convenience	2.97 (0.08)	0.57 (0.03)	4.14 (0.06)
	Safety/Security	3.48 (0.09)	0.76 (0.02)	4.09 (0.06)
Priming(128)	Convenience	3.01 (0.08)	0.56 (0.03)	4.20 (0.06)
	Safety/Security	3.66 (0.08)	0.80 (0.02)	4.23 (0.06)
Control(128)	Convenience	3.14 (0.08)	0.60 (0.03)	4.08 (0.06)
	Safety/Security	3.97 (0.08)	0.88 (0.02)	4.19 (0.06)

Note. The number in the parentheses of the first column indicates the number of participants for each condition. The number in the parentheses of the last three columns shows the standard errors of corresponding cell.

TABLE IX: Result of Three Privacy-Decision Measures of Health Scenario

Condition	Utility-Privacy Tradeoff	Data-Sharing Decision	Confidence Rating
Feedback(125)	3.06 (0.11)	0.59 (0.04)	4.03 (0.09)
Priming(128)	2.98 (0.12)	0.52 (0.04)	4.15 (0.08)
Control(128)	2.71 (0.12)	0.41 (0.04)	4.06 (0.08)

Note. The number in the parentheses of the first column indicates the number of participants for each condition. The number in the parentheses of the last three columns shows the standard errors of corresponding cell.

TABLE X: Results of Generalized Linear Mixed-Effects Regression on Data-Sharing Decision.

Effect	*df*	χ^2	*p*
Scenario	1	5.58	.018
Condition	2	2.40	.302
Scenario $\times$ Condition	2	6.60	.037