The Influence of Friends and Experts on Privacy Decision Making in IoT Scenarios

PARDIS EMAMI-NAEINI, Carnegie Mellon University, USA
MARTIN DEGELING, Ruhr-University Bochum, Germany
LUJO BAUER, Carnegie Mellon University, USA
RICHARD CHOW, Intel Corporation, USA
LORRIE FAITH CRANOR, Carnegie Mellon University, USA
MOHAMMAD REZA HAGHIGHAT, Intel Corporation, USA
HEATHER PATTERSON, Intel Corporation, USA

As increasingly many Internet-of-Things (IoT) devices collect personal data, users face more privacy decisions. Personal privacy assistants can provide social cues and help users make informed decisions by presenting information about how others have decided in similar cases. To better understand which social cues are relevant and whose recommendations users are more likely to follow, we presented 1000 online participants with nine IoT data-collection scenarios. Some participants were told the percentage of experts or friends who allowed data collection in each scenario, while other participants were provided no social cue. At the conclusion of each scenario, participants were asked whether they would allow the described data collection. Our results help explain under what circumstances users are more or less likely to be swayed by the reported behavior of others in similar scenarios. For example, our results indicate that when friends denied data collection, our participants were more influenced than when friends allowed data collection. On the other hand, participants were more influenced by experts when they allowed data collection. We also observed that influence could get stronger or wear off when participants were exposed to a sequence of scenarios. For example, when experts and friends repeatedly allowed data collection in scenarios with clear risk or denied it in scenarios with clear benefits, participants were less likely to be influenced by them in subsequent scenarios.

CCS Concepts: • **Security and privacy → Social aspects of security and privacy**; • **Human-centered computing → User models**;

Keywords: Social Influence; Decision Making; Privacy; Internet of Things (IoT)

ACM Reference Format:
Pardis Emami-Naeini, Martin Degeling, Lujo Bauer, Richard Chow, Lorrie Faith Cranor, Mohammad Reza Haghighat, and Heather Patterson. 2018. The Influence of Friends and Experts on Privacy Decision Making in IoT Scenarios. In *Proceedings of the ACM on Human-Computer Interaction*, Vol. 2, CSCW, Article 48 (November 2018). ACM, New York, NY. 26 pages. https://doi.org/10.1145/3274317

Authors' addresses: Pardis Emami-Naeini, Carnegie Mellon University, Pittsburgh, PA, USA, pardis@cmu.edu; Martin Degeling, Ruhr-University Bochum, Bochum, Germany, martin.degeling@rub.de; Lujo Bauer, Carnegie Mellon University, Pittsburgh, PA, USA, lbauer@cmu.edu; Richard Chow, Intel Corporation, Santa Clara, CA, USA, richard.chow@intel.com; Lorrie Faith Cranor, Carnegie Mellon University, Pittsburgh, PA, USA, lorrie@cmu.edu; Mohammad Reza Haghighat, Intel Corporation, Santa Clara, CA, USA, mohammad.r.haghighat@intel.com; Heather Patterson, Intel Corporation, Santa Clara, CA, USA, heather.m.patterson@intel.com.

2573-0142/2018/11-ART48
https://doi.org/10.1145/3274317

1 INTRODUCTION

The interconnection of a myriad of smart devices and their pervasive ability to collect, store, and transfer information about people's private lives give rise to significant privacy concerns [90]. Because of the scope of data collection and its potential consequences, people find specifying their privacy preferences and making privacy decisions regarding IoT devices to be overwhelming [86]. People are also hindered in making informed privacy-related decisions due to the difficulty of obtaining information about the data collection [70]. These challenges become more burdensome as the number of IoT-related privacy decisions increases [22]. In fact, every smart light bulb, thermostat, camera, or any other sensor that people encounter may require a decision about whether to allow collection of user data. Although IoT data collection can pose a privacy risk, it may also provide users with benefits. However, the best decision is not always evident in the midst of these trade-offs. For instance, people may not be comfortable with law enforcement using video from security cameras to learn their whereabouts in a building, yet the perceived benefit of maintaining building security can outweigh their privacy concerns. There are numerous examples of this type that shed light on why privacy decisions in IoT settings can be very challenging, and distinct from those in other domains.

From birth on, people learn that others' opinions and judgments are frequently a reliable source of evidence about reality [25]. In many social and biological systems, individuals rely on other members' perceptions and observations to make decisions or adapt their behaviors accordingly [12, 15, 20, 74]. As defined by Latané, such influence is the result of "the real, implied, or imagined presence or actions of other individuals" [53]. Social influence has been demonstrated to have a strong impact on people's decision making in many domains [17] as people look at others' behaviors and opinions to inform and improve their own judgments [3, 30].

Research demonstrates that social cues also have an effect on users' information-sharing behaviors on social-networking sites (SNSs). Spotwood and Hancock found that SNS users' privacy-related behaviors and decisions are influenced by explicit social cues. For example, when users are made aware that most users select a privacy-protective setting, they are more likely to choose a more private setting themselves [83].

To mitigate the challenge of privacy-related decision making in IoT settings, we sought to understand the manifestation of social influence in IoT scenarios. Researchers have shown that reliance on social influence increases as the uncertainty in individuals' judgments and decisions rises [85]. Such uncertainty is especially prevalent in privacy decision making in the IoT world, where data collection introduces inevitable trade-offs between risks and benefits. We envision that personal privacy assistants built into IoT devices, such as smartphones and smartwatches, will help manage these decisions [16, 22]. Social cues provided by these assistants may help users make faster, more informed decisions by presenting information about how others have decided in similar cases.

Social influence can be categorized as either *direct* or *indirect*. *Direct* social influence is based on persuasion, whereas *indirect* social influence is a subtler psychological process, which occurs as a result of knowing aggregate information about others' actions [69].

In this paper, we investigate the effects of indirect social influence on user decisions about whether to allow data collection by IoT devices. In our large-scale online study, we exposed Mechanical Turk (MTurk) participants to various hypothetical data-collection scenarios and asked whether they would allow or disallow data collection in each scenario. Participants were randomly assigned to one of five conditions, which varied in the source and type of social influence. In four of these conditions, participants were told what percentage of their *friends* or of *privacy experts* had assented to the data collection. For each source of influence (friends and privacy experts), we further varied

the type of the influence to be *consistent* or *inconsistent* with the decisions made in that scenario by unprimed participants in a 500-person pre-study. For example, an *inconsistent experts* scenario would include a statement that 85% of privacy experts allowed data collection, whereas the majority of our pre-study participants who were exposed to the scenario without social influence had chosen to deny data collection. In the control condition, participants were not exposed to any social influence.

Our study design enabled us to understand the impact of indirect social influence on privacy decision making. We found that in general, displaying aggregate information about the behaviors of friends and privacy experts' sways participants' privacy-related decisions. Moreover, we found that social cues help participants make decisions significantly faster.

To better understand the variables that predict people's response to social cues when making privacy-related decisions, we studied factors emerging from the literature such as expertise [36, 42, 72], level of consensus [67], opinion difference [72], and task difficulty [7, 19, 25].

We found that participants are influenced by both privacy experts and their friends, but in different ways. When friends denied data collection, our participants were more influenced than when friends allowed data collection. On the other hand, and perhaps surprisingly, participants were influenced by privacy experts more when they allowed data collection.

We also observed that the influence of social cues could wear off or get stronger, depending on whether the cues were consistent or inconsistent with pre-study participants' opinions for several scenarios in a row. For example, 40% of participants who were shown a single inconsistent social cue would follow that cue; but only 32% of participants who had previously been exposed to one inconsistent social cue, and 29% of those who had seen two inconsistent social cues, would follow the subsequent (again inconsistent) cue. On the other hand, if the social influence is consistent over several scenarios, then participants are more likely to be affected by it in future scenarios.

In addition, the majority of our participants specified that technology expertise was the quality that would influence them the most when making privacy-related decisions (77% of participants), whereas the least selected influential quality was having a friendship history (19% of participants). This suggests that participants' decisions would be affected significantly more by advice from individuals who are known to have more expertise than by naive cues from people with whom they have a friendship history.

The rest of this paper is organized as follows. In Section 2, we discuss related work on social influence and on privacy preferences and decision making. In Section 3, we describe our methodology. In Section 4, we present the outcomes of our data analysis and the resulting model. In Section 5, we interpret our results. In Section 6, we discuss the limitations of our study methodology and some potential approaches to mitigate these limitations. Finally, we conclude the paper in Section 7.

2 RELATED WORK

Researchers have studied the power of social influence on individual opinions since the early 20th century. Jenness first studied conformity in 1932 and ran an experiment to understand how human estimation changes based on the influence of the majority; he observed that almost all the participants changed their opinions to be close to the group estimate [41]. In Sherif's well-known Autokinetic Effect experiment, when participants were unsure about the light's movement, they relied on information from others to form their own opinion [81]. In another classic psychology experiment, Asch studied the extent to which majority opinion could affect individual decisions and judgments. In his famous conformity experiment, he asked participants in a group setting to perform a judgment task, in which they had to guess the closest line to the target line. He found that an individual would conform to the majority's opinion even when the correct answer was obvious.

He showed that social influence can make people question their decision when it is different from the majority or simply exhibit public conformity to avoid contradicting group norms [4].

Kelman identified three different types of conformity: *compliance*, *internalization*, and *identification* [45]. *Compliance*, is conformance to meet a specific requirement or to avoid a specific punishment. In contrast, *internalization* occurs when individuals conform to something they believe in and consider a useful solution to their problem. Participants in Sherif's conformity experiment mostly fell into this category. *Identification*, applies to individuals who conform to fulfill their desire for a relationship with another person or group.

Deutsch and Gerard distinguish *normative social influence* and *informational social influence* [25]. *Normative social influence* occurs mostly when individuals want to fit in with a group, a famous example of which is Asch's aforementioned line experiment. *Informational social influence*, occurs when people seek information and guidance. For example, in Sherif's Autokinetic experiment, when participants were unsure about the correct answer, they observed other members of the group to inform their own decisions. Informational influence serves as a "cognitive repair" that lessens the harm of depending too much on one's own judgments [37].

We studied the impact of *indirect informational social influence* in the context of privacy-related decision making and identified factors that can predict individual responses to social influence. We presented our participants with vignettes that described IoT data collection scenarios. The vignette methodology, which is a technique to elicit opinions and attitudes of individuals by analyzing their responses to different scenarios [8], is particularly useful for determining significant factors and explaining the extent of their impact [1].

The factors that we studied were identified in the psychology literature as relevant for understanding the impact of social influence on decision making, but had not previously been examined in the context of IoT privacy. We tested the significance of three factors by varying them in the vignettes: expertise [36, 42], consensus level [67], and opinion difference [72].

The literature shows that people heed advice from experts more than advice from less informed individuals. In our study, we tested the impact of expertise by presenting participants with social cues from either privacy experts or friends. In addition, research shows that the level of consensus influences an individual's decision making. For instance, Martin et al. found that people were more influenced when presented with a stronger consensus than when presented with a weaker consensus [67]. In another experiment, Mackie tested the impact of 64% versus 82% of consensus level on decision making [64]. Inspired by these experiments, we tested the impact of two levels of consensus (85% vs. 65%) on participants making privacy-related decisions regarding IoT data collections.

We studied how close the participants' initial opinions were to influencers' opinions when making privacy-related decisions. Meshi et al. used the term *opinion difference* to describe this factor. With their fMRI experiment, they found that advice utilization will increase as the opinion difference becomes smaller [72].

Past research focused on segmenting users based on their privacy preferences has led to heavily-cited differences between privacy fundamentalists or highly concerned, pragmatists or moderates, and marginally concerned or minimalists [49]. However, the narrative that consumers make rational decisions on privacy matters has been challenged [27]. Current approaches to classify users based on their privacy concerns have, therefore, concentrated on finding other indicators for privacy behavior such as knowledge or motivation [28], or simply on using previous choices to generate recommendations for future decisions [47, 60].

Where the appropriateness of information sharing is related to factors of its social and technological context, other researcher, who are inspired by the concept of privacy as contextual integrity [76], have studied the factors that influence privacy decisions specifically for IoT. For instance, Lee and

Kobsa [56] and Emami-Naeini et al. [29] examined the influence on privacy decisions of factors such as *who* is collecting the data, *where* the data is collected, *what* kind of data is being collected, the *reason* for collection, and the *persistence* of data. In addition, Lee et al. [58] investigated the *what* and *who* factors for wearable devices. More generally, Chow [16] describes how these factors might be communicated and understood with a conceptual "privacy stack."

Morton and Sasse [73] classified the information-seeking behavior of participants who try to decide whether or not they want to use a service. They found that only 15% of their 58 participants are "crowd followers" who are heavily influenced by "environmental cues" such as the advice of others or media reports when making privacy-relevant decisions. They identified four other groups — information controllers, security concerned, benefit seekers, and those looking for organizational assurance — that all have different demands and interests when making privacy decisions.

Mendel and Toch [71] studied the phenomenon of social influence on privacy behavior in online social networks. We confirm some of their results in the IoT privacy setting. For instance, user's susceptibility to privacy influence depends on his or her privacy knowledge and self-efficacy.

Several other projects have investigated the impact of social influence in privacy or security settings. Patil et al. [78] studied how the preferences of an individual's social circle affects the privacy settings that are selected when using an instant messaging application. They found that this additional information provided useful guidance, but this influence was secondary to the privacy-sensitivity of the setting itself. Besmer et al. [11] studied access control settings for third-party social network applications, measuring the influence of information about the percentage of other users who shared information with such applications. They found that this information could impact user decisions, but only if the cues were sufficiently strong. Balestra et al. [5] studied the impact of exposure to social annotations on privacy consent for a genomics application. These annotations had the general effect of making users feel more informed, but also less confident in their understanding of the application and less trusting in the institution soliciting the consent. Das et al. [23] studied how social influence affects Facebook security settings. They found that many friends adopting a particular feature would influence users towards adoption. Conversely, few friends adopting a feature may bias users away from adoption, which is viewed as not ideal for security features where the goal is to encourage adoption as much as possible.

Social influence has also been examined outside of privacy and security contexts. For example, collaborative filtering systems for product recommendations [48] and reputation management systems, such as the ones used by eBay [79], are examples of social influence in decision making.

Finally, our work has been inspired by systems such as the Privacy Assistant [22] and pawS [51]. In these systems, each user has a software agent that helps them understand and interact with their environment for privacy decisions. Building such systems requires a detailed understanding of the factors that influence users, consciously or unconsciously, when making privacy decisions, so that the software agent can highlight these factors for the user or automatically make decisions on the user's behalf.

3 METHODOLOGY

We conducted a mixed-design online study with 1000 Mechanical Turk (MTurk) participants from the United States in order to understand the impact of social influence on privacy-related decision making regarding IoT devices and identify the factors affecting this influence. In this section, we first discuss the design of the study and then describe the approaches we used to analyze our data.

3.1 Study Design

We conducted our study on Mechanical Turk so that we could recruit a large number of diverse participants quickly and economically. By recruiting 1000 participants we had adequate statistical

power to conduct the desired tests on the data and make comparisons across 5 treatment groups. MTurk has been used frequently for experimental studies on related topics such as understanding people's privacy concerns and biases in decision making [29, 35, 77]. Nonetheless, MTurk does introduce some biases, which we will discuss in Section 6. To improve the reliability of our results, we required MTurkers' Human Intelligence Task (HIT) rate to be above 90%.

Before launching the main study, we ran a pre-survey with 500 MTurk participants. In that survey, we presented participants with 28 hypothetical IoT data collection scenarios, each describing a location, a data collection device, the type of data being collected, how data will be used and shared, and how long data will be retained. After each scenario, we asked participants if they would allow or deny that data collection. The factors and their interactions that we selected to test in these scenarios have been shown by researchers to be among the factors that influence privacy concerns [9, 10, 46, 54, 55, 57, 59].

From the 28 pre-survey scenarios We selected three groups of scenarios for our main study, representing a range of privacy concern levels: three *allow* scenarios, where more than 80% of our pre-study participants agreed to allow data collection (without being swayed by social influence, which was not present in the pre-survey); three *deny* scenarios where fewer than 20% allowed data collection; and three *balanced* scenarios where 45% to 55% allowed data collection. In the main study, participants were exposed to these nine scenarios, which are included in Appendix A, in random order, with a series of questions after each scenario.

In our main study, we randomly assigned 1000 participants to one of four experimental conditions or the control condition, for a total of 200 participants per condition. Participants were not presented with any social cues in the control condition, whereas in the experimental conditions, we appended information about the percentage of *influencers* who allowed the data collection — described either as *friends who use this app* or *privacy experts*. For each decision in our study, we used the average opinion of the pre-study participants as a proxy for an initial opinion on that decision. To understand how opinion difference impacts privacy-related decision making regarding IoT devices, we tested the *consistency* of the social cue. A consistent social cue reflects a small opinion difference, whereas an inconsistent social cue exhibits a large opinion difference. In the two consistent conditions, participants were told that most influencers allowed data collection for *allow* scenarios and denied data collection for *deny* scenarios. Conversely, in the two inconsistent conditions, participants were told that most influencers allowed data collection for *deny* (resp., *allow*) scenarios. For two of the three *balanced* scenarios, participants in the consistent conditions were told that most influencers allowed data collection; for one of the balanced scenarios they were told that most influencers denied data collection. In the inconsistent conditions, the influencers' decisions were reversed. The description of each scenario presented to the participants in the control condition was identical to what was shown in the experimental conditions, except that the sentence indicating how friends or experts had behaved was not present.

We used two levels of consensus when describing the percentage of influencers who allowed data collection: *weak* and *strong*. The weak consensus was described as either "more than 65%" or "fewer than 35%" and the strong consensus was described as either "more than 85%" or "fewer than 15%." Participants in all four experimental conditions were exposed to the following combinations of scenarios with strong and weak levels of consensus. In the three *allow* scenarios, two scenarios had strong consensus and one had weak consensus. In the three *deny* scenarios, two scenarios had weak consensus and one had strong consensus. Finally, in the three *balanced* scenarios, two scenarios had strong consensus and one had weak consensus.

The following is scenario D1, with weak consensus, as shown to participants in the consistent friends condition:

"You are at a department store. This message is displayed on your smartphone: This store has a facial recognition system that takes pictures of customers' faces automatically as they enter the store in order to identify returning customers. This method is used to keep track of your orders and make suggestions based on your purchasing habits. Your picture will never be deleted. Fewer than 35% of your friends who use this app allowed this data collection."

After the participants read each scenario, we asked them to move to the next page of the survey, where we asked them six questions, as shown in Appendix B.1. They could return to the scenario by clicking on the back button. The first question was an attention check question designed to check whether participants understood basic information about the scenario they just read. For each participant, over the course of the nine scenarios, we asked three attention check questions about the type of data, three about retention time, and three about the location of the data collection. Following the attention check question, we asked participants whether they would allow or deny the data collection described in the scenario (the possible answer choices were allow, probably allow, deny, and probably deny) and the reasons behind their decision in a multiple choice question with 15 answer choices. In addition, we asked them on a five-point Likert scale to what extent they agree that the described data collection is beneficial to them and to what extent they agree it is beneficial to the society. Finally, we asked participants how confident they were about their decision to allow or deny the data collection.

After participants had seen all nine scenarios and answered the questions about each, we asked participants to self-report how much they were influenced by the decisions that the influencers made in the scenarios and also asked about the reasons they were or were not influenced.

For the last (ninth) scenario, we also asked participants how their decisions might change if they were shown the scenario with the same consensus level but with a different influencer (i.e., if they had privacy experts as their influencers throughout the survey, we asked them about their friends as the influencers and vice versa). We then asked them how their decisions might change if they were shown the scenario with the original influencer, but with the opposite majority decision (from more than 85% to fewer than 15% and vice versa or from more than 65% to fewer than 35% and vice versa).

We expected there could be more groups or individuals besides those we asked about in the survey questions. Hence, as an open-ended question, we asked the participants to name other potential influencers when making a privacy-related decision.

As trust is known to play a role in response to social influence [39], we asked participants to specify their level of trust in a number of potential influencers, such as privacy experts, their family, or their colleagues. The list of potential influencers was derived from a pre-survey question in which participants were asked in an open-ended question to describe people or organizations they would be interested in consulting to help them make a similar decision.

Next, we asked participants "What qualities would make you likely to be influenced by a specific group of people when you need to make decisions like the ones in our scenarios?" We provided a list of nine qualities and invited participants to specify their own.

At the end of the survey, we asked participants some general demographic questions about their age, gender, education level, and income range.

In the control condition, the questions we asked at the end of each scenario were identical to what we asked in the experimental conditions. Only in the ninth scenario did we ask participants to assume having privacy experts and their friends as influencers, and we again posed the same questions that we asked the participants in the experimental conditions.

To have a record of how much time participants spent throughout the study, we instrumented our surveys to collect the time by setting invisible timers before each question.

3.2 Data Analysis

One of our main goals in this study was to find out what factors explain whether or not participants follow the social cues they receive from privacy experts and their friends. The complete list of the factors that we analyzed in our study and their corresponding levels are described in Table 1.

Factor	Levels
influencer type	privacy experts, friends
how consistent the social cue is compared to the responses of pre-survey participants	consistent, inconsistent
strength of the social cue	strong (more than 85% or fewer than 15%), weak (more than 65% or fewer than 35%)
total number of prior scenarios	0 to 8
direction of the social cue	toward allow, toward deny
to what extent the data collection is beneficial to me	strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
to what extent the data collection is beneficial to society	strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
to what extent participants trust [specific groups] e.g., their friends, privacy experts, or colleagues	strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
to what extent participants agree that [privacy experts/friends] have more technical knowledge than they do	strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
to what extent participants agree that [privacy experts/friends] have more background information than they do	strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
to what extent participants agree that they generally make decisions on their own	strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
current scenario type	allow, deny, balanced
prior scenario type (note: this is determined by condition for allow and deny scenarios but will vary for balanced scenarios)	consistent, inconsistent
to what extent participants agree that they have sufficient knowledge about privacy	strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
to what extent participants agree that they have sufficient knowledge about technologies mentioned in the scenario	strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
general demographic information such as: age, gender, income range, and education level	the corresponding levels for each demographic factor are presented Appendix A

Table 1. Description of the data analysis factors.

We conducted a mixed between-subjects and within-subjects study with experimental factors between participants and repeated measure factors within participants. Thus, we applied a mixed-model logistic regression with both random intercept and random slope on a binary outcome to describe whether participants acted consistently with the influence (1) or not (0). To avoid over fitting, we performed an exploratory analysis on only the first 20% of the data [38]. We looked at the distribution of factors and the summary statistics to find trends in our collected data. We also applied the model selection on the first 20% of the data. In order to find the best model, we performed backward elimination and compared the models by their Bayesian information criterion (BIC), which is a general metric for the goodness of fit. We took the following steps to select the best model [43]:

(1) Start by building the model with all the factors and interaction terms.

(2) Remove the factor or the interaction of two factors which has the highest p-value. If the interaction of a factor that has the highest p-value is still in the model, it will not be removed until all its interactions are removed from the model.
(3) Repeat step (2) until the BIC does not decrease.

After finding the best model, we checked the performance of our model by the common approach of training the model on the first 80% and testing the model on the last 20% of our collected data [24].

3.3 Free Text Responses

As part of the survey, participants were asked to specify which other people or organizations would influence their decisions. The answers were collected as free text and later annotated by two researchers. Each annotator first independently developed a simple codebook [65] based on 250 answers, 50 from each of the five study conditions. The annotators then discussed, refined, and merged both codebooks into one, which took two iterations. The resulting codebook contained 23 codes and was used by both annotators to independently code all the responses. After finishing the coding, we decided to merge some codes (e.g., "spouses" and "parents" into the more general annotation "family"), as the number of their occurrences was very low and there was little conceptual difference between the codes. The resulting 14 codes are listed in Table 2. The annotator agreement as measured by Cohens Kappa was $\kappa = 0.81$, which is regarded as very good to excellent [50].

In addition to manual coding, we also examined the sentiment of the answers using an online service[1] that classifies the sentiment of a given text as *positive*, *negative*, or *neutral* to learn whether participants expressed any strong feelings toward the question.

4 RESULTS

In this section, we present our findings. We first report on the impact of social cues on the response time for decision making (Section 4.1). Next, we describe and evaluate our model that predicts whether participants will follow social cues (Section 4.2). We compare participants' self-reported perceptions of how they were influenced by social cues to their observed behavior and note interesting divergences (Section 4.3). We then elaborate on the extent to which participants report trusting different influencers and the characteristics of influencers that affect that trust (Section 4.4).

For our main study, we recruited 1000 Mechanical Turk (MTurk) participants from the United States. Participants took an average of 15 minutes to complete the survey. Participants' demographics are shown in Table 3.

Out of 1000 participants, only 14 participants made more than two mistakes on the nine attention-check questions. However, these participants' answers to other survey questions, and the amount of time they took to answer them, did not suggest inattention. Therefore, we did not exclude any of these participants from our analysis.

4.1 Faster Privacy-Related Decision Making

Research has shown that people are more likely to look for guidance and information from others when a task is perceived as difficult [7, 19, 25]. In our study, the task was the privacy-related decision to allow or deny the data collection in each scenario. We used mean response time (RT), the time it took participants to make the decision in each scenario, to measure the difficulty of each task. RT distributions are positively skewed, which contradicts the assumptions behind some common statistical tests. Hence, when using RT as a dependent variable in analyses, we apply a generalized linear mixed model (GLMM) on the raw timing data [61]. To model the influential factors that impact the RT of decision making, we applied GLMM with a random intercept for each

[1]We used text-processing.com that offers a sentiment classifier trained on tweets and movie reviews.

Code	Occurrences	Description
no one	265	Participants do not want any input
experts	217	Privacy experts were mentioned most often, but this group includes also other experts like safety or technology experts
family	178	Includes mentions of parents, spouses, siblings, or family in general
law enforcement	97	Mostly mentioned in a general way, but some participants referred to specific institutions like police, FBI, or NSA
media	52	Participants said they would be looking for news, some referring to specific platforms where they read online reviews
friends	47	Some participants tried to differentiate to emphasize that this group should consist of "close" or "trusted" friends
coworkers	39	Especially referring to workplace scenarios, participant would ask colleagues about their opinions
government	34	Those expected guidelines from government officials on what is appropriate
companies	23	Some wanted to know more about the reasons for a data collection, therefore referring to the companies asking for their data
non profits	22	Most often mentioned were EFF or ACLU
general public	19	Mentioned interest in what the "general public" or "society" would do in these scenarios
boss	17	Similar to "coworker," some participants would listen to what their managers or superiors at work would recommend
celebrity	15	Specific and unspecific mentions of celebrities. Most notably Edward Snowden (6 times)
don't know	12	Participants had no preference

Table 2. Descriptions of the codes used on free text answers. An occurrence is counted if both annotators used the same code.

Gender		Age		Education		Income	
Male	51.1%	Range	18-74	No high school degree	0.0%	< $25k	26.0%
Female	48.6%	Mean	35.1	High school degree	29.4%	$25k-$49k	34.6%
Other	0.0%	Std. Dev.	10.2	College degree	47.1%	$50k-$74k	24.7%
Prefer not to answer	0.3%			Professional degree (Masters/PhD/medical/law)	12.8%	$75k-$99k	10.4%
				Associates degree	9.5%	$100k-$124k	0.4%
				Prefer not to answer	0.0%	$125k-$149k	0.2%
						$150k-$174k	0.0%
						$175k-$199k	0.0%
						> $200k	0.0%
						Prefer not to answer	3.7%

Table 3. Participant demographics

user. Our dependent variable was the amount of time participants spent on making decisions to allow or deny the data collection and the independent variables were factors such as the study condition and scenario type. In our model, we used a Gamma distribution, as it is commonly used to statistically describe RT distributions [87]. Participants spent 3.83 seconds on average to make the decision to allow or deny each data collection. Our analysis showed that participants who were in the experimental conditions spent 3.78 seconds on average per decision and were significantly faster than the participants in the control condition (mean = 4.24s, std. dev. = 19.62s, coefficient $= -0.07$, p-value < 0.05).

Drilling down, we observed that participants on average made faster decisions in the *allow* scenarios (mean = 3.69s, std. dev. = 12.62s) than the *deny* scenarios (mean = 3.91s, std. dev. = 16.75s).

Compared to the *allow* scenarios, it took them significantly longer to make decisions in the *balanced* scenarios (mean = 4.02s, std. dev. = 18.17s, coefficient = 0.06, p-value < 0.05).

We observed that social cues resulted in faster decision making for all three types of scenarios. Our analysis specifically showed the significant impact of social cues on the *balanced* scenarios, which required more difficult decisions as they generally required participants to consider trade-offs between clear risks and clear benefits. Notably, participants made significantly faster decisions about *balanced* scenarios in the experimental conditions (mean = 3.89s, std. dev. = 10.03s) than the control condition (mean = 4.55s, std. dev. = 17.43s, coefficient = -0.09, p-value < 0.05), suggesting that social cues allowed participants to reach a decision more quickly. Summary statistics for the timing data are presented in Table 4.

Conditions	**Scenario type**					
	allow		*deny*		*balanced*	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
control	3.94	11.06	4.24	12.51	4.55	17.43
consistent experts	3.93	13.93	3.94	15.39	3.88	4.91
inconsistent experts	3.68	4.32	3.72	5.06	3.51	11.46
consistent friends	3.30	11.09	3.81	7.15	4.07	3.26
inconsistent friends	3.60	4.08	3.84	9.67	4.11	10.76
all conditions	3.69	12.62	3.91	16.75	4.02	18.17

Table 4. Summary statistics for response time (RT) of decision making

In general, our analysis demonstrates that providing participants with social cues, either from privacy experts or their friends, will help them make privacy-related decisions faster, especially in more complex scenarios, which exhibit inherent trade-offs between risks and benefits.

4.2 Inferred Influence

After presenting each of the scenarios, we asked participants whether they would allow the data collection. These responses allowed us to infer the amount of influence social cues had in each experimental condition. For more statistical power, we binned the answers as 0 (merging "probably deny" and "deny") or 1 (merging "probably allow" and "allow").

For each scenario, we compared participants' preferences to allow or deny the data collection in the experimental conditions with the preferences of participants in the control condition. We created a factor called *follow* that indicates whether participants decisions followed the presented social cue. This binary factor was either 0 (not follow) or 1 (follow). We observed that 63% of participants followed the social cues they received in the experimental conditions. Table 5 shows the differences between the fraction of participants who allowed data collections in each experimental condition and the fraction of those who allowed the same data collections in the control condition. To statistically analyze the extent of the influence in different conditions in our repeated measures study, we applied mixed-effects logistic regression with random intercept for each user. The dependent variable in our analysis was binary, indicating allow or deny, and the independent variables were the scenario type, study condition, and strength of the social cue. The goal of the regression was to determine which experimental conditions would significantly increase the likelihood of allowing or denying the data collection compared to the control condition. Using random intercept in these analyses enhances the credibility of the results as the method considers the correlation between multiple data points within each user. Our results showed that compared to the control condition, participants were most influenced in the *balanced* scenarios. Perhaps surprisingly, we found that strong inconsistent social cues from friends have the most influence on participants making more complex decisions.

Influencers	Consistency	Level of consensus	Scenario type			
			allow	*deny*	*balanced – allow*	*balanced – deny*
experts	consistent	strong	0.11	*	0.18	−0.08
		weak	*	*	*	—
	inconsistent	strong	*	*	−0.17	0.1
		weak	*	0.15	−0.13	—
friends	consistent	strong	0.05	−0.07	0.09	*
		weak	*	−0.06	*	—
	inconsistent	strong	*	*	−0.23	*
		weak	*	0.08	−0.09	—

Table 5. Differences between the fraction of participants who allowed data collections in the experimental conditions and the control condition. Positive numbers indicate more participants allowed data collection, whereas negative numbers indicate more participants denied data collection. We applied generalized linear mixed model (GLMM) regression with random intercept for each participant on users' preferences to allow or deny the data collection in order to find out whether different factors increase or decrease the likelihood of allowing or denying the data collection. The * signs inside the table indicate that the difference is not statistically significant. The — signs indicate scenarios not tested. All numeric values shown indicate statistically significant difference.

In order to understand the factors that contribute to following social cues, we ran a regression analysis to build a model that describes the participants' behavior. In this model, the dependent variable was *follow*. Besides the factors described in Table 1, we also included the participants' demographic information in the model. Based on the results of our model selection process, we identified the factors that predict whether participants follow social cues. Our model showed that participants follow inputs from their friends and privacy experts differently based on whether the influence is in the direction of allowing or denying data collection. If the presented social cue favors allowing the data collection, participants will follow privacy experts more than their friends. On the other hand, when the direction of the social cue is toward denying the data collection, they will be more influenced by their friends. This difference between following experts and friends is also reflected in Table 5. Another significant factor in our model was the strength of the social cue. As expected, we found that a strong cue influences participants more than a weak cure.

The consistency of the cue was another statistically significant factor that contributed to our model. The regression results indicated that participants will follow consistent social cues significantly more compared to inconsistent social cues. In addition, we found that participants will become more influenced after experiencing a repeated sequence of cues that are consistent with pre-study participants' decisions. On the other hand, participants will become less influenced after experiencing a sequence of social cues that are inconsistent with pre-study decisions, especially when the cues come from experts and suggest less privacy-protective decisions.

Among the factors that we tested during model selection, many turned out not to be statistically significant and some were removed during model selection. For example, none of the demographic factors were statistically significant. The detailed results of the logistic regression for our best model are presented in Table 6, along with the complete list of factors that we removed based on their contribution to the model.

To evaluate its performance, we trained our model on the first 80% of our dataset and tested on the last 20%. The model achieved a test AUC of 0.81, which is considered excellent [40].

4.3 Reported Influence

After participants had been exposed to all nine scenarios and answered the questions associated with each, we asked them to report on a five-point Likert scale how much their decision making

Factor	Estimate	Std Err	Z-value	p-value
(Intercept)	1.50	0.16	9.36	0.00^{***}
strong cue	0.09	0.11	0.85	0.39
social cues from friends	−0.04	0.09	−0.43	0.66
total number of prior scenarios	0.04	0.01	2.96	0.00^{**}
direction of the cue: toward deny	0.27	0.09	3.07	0.00^{**}
inconsistent prior scenarios	−0.82	0.11	−6.98	0.00^{***}
current scenario type: *deny*	0.00	0.12	0.00	0.99
current scenario type: *balanced*	−0.11	0.13	−0.81	0.41
making decisions on your own	−0.29	0.10	−2.84	0.00^{**}
inconsistent social cue	−1.39	0.07	−19.66	0.00^{***}
strong social cue in the *deny* scenarios	−0.20	0.17	−1.21	0.22
strong social cue in the *balanced* scenarios	0.15	0.16	0.97	0.02^{*}
friends' behavior toward denying the scenario	0.25	0.12	1.99	0.04^{*}
increase in the number of inconsistent prior scenarios	−0.13	0.02	−5.86	0.00^{***}
Observations	6400			
Log-Likelihood	−3046.915			
Akaike Inf. Crit.	6119.831			
Bayesian Inf. Crit.	6206.292			
Note:	$^{*}p < 0.05$ $^{**}p < 0.01$ $^{***}p < 0.001$			
Insignificant factors from Table 1 in model selection:	(6, 7, 8, 9, 10, 14, 15, 16)			

Table 6. Regression results for the model to explain *follow* (i.e., whether participants follow advice). These results are reported on the last 80% of the dataset (the first 20% having been used for model selection).

was (or would have been, in the control condition) influenced by knowing privacy experts' or their friends' decisions. We found that the percentage of participants who reported that they would be influenced by privacy experts was similar in the control condition (56%) and the condition in which participants received consistent cues from privacy experts (52%). On the other hand, the percentage of participants who reported being influenced by consistent cues from privacy experts (54%) is significantly larger than the percentage of participants who reported being influenced in the other three experimental conditions (consistent (21%) and inconsistent cues from friends (5.5%) and inconsistent cues from privacy experts (24%)) (p-value < 0.05). Many more participants reported being influenced by consistent social cues from either friends or experts (119 participants) than inconsistent ones (49 participants) (p-value < 0.05 for differences between consistent and inconsistent cues from friends; and from experts). The extent of participants' reported influence in each study condition is presented in Figure 1.

After asking participants to report how much they were influenced, we asked them to provide us with their reasons. Two top reasons for participants who reported being influenced were "I generally like to find out what other people have done when making a decision" (69%) and "I think my friends/privacy experts have more technical knowledge than me" (54%). Among the participants who reported not being influenced by the social cues, the most common reasons were "I generally make decisions on my own" (81%) and "I generally make these kinds of decisions on my own" (76%).

4.4 Willingness to Trust Influencers

After exposing the participants to all nine scenarios, we asked them whom they would trust to give them good advice when making privacy-related decisions regarding data collection. Participants were instructed to choose an answer on a five-point Likert scale from "strongly agree" to "strongly disagree" for each of the following groups: privacy experts, family, real-life friends, people working in technical fields, colleagues, social-network friends, and no one except themselves.

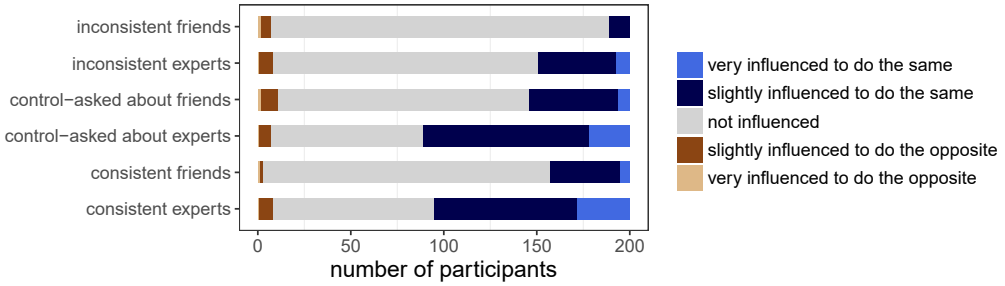

Fig. 1. Participants' reported influence in our five study conditions from "very influenced to do the same" to "very influenced to do the opposite."

Among our five conditions, we did not observe any statistically significant difference between trusting advice from family, people working in technical fields, colleagues, and social-network friends. However, the differences between the conditions were statistically significant in trusting privacy experts, friends, and no one but myself, as discussed next.

Our results showed that most participants trusted privacy experts to give them privacy-related advice, except when presented with inconsistent social cues from privacy experts. Participants significantly lost their trust in privacy experts when their behaviors were not consistent with most participants, as determined by our pre-survey results (66% of participants in the control condition trusted privacy experts, compared to 45% of the participants who received inconsistent cues from privacy experts).

Looking at the number of prior scenarios each participant saw, we found that participants were increasingly less affected by social cues as they saw more inconsistent behaviors from the influencers, especially from privacy experts. The percentage of participants who followed inconsistent social cues from privacy experts decreased by 10% (from 41% to 31%) after seeing only one inconsistent cue. For participants who received inconsistent cues from their friends, we observed a decrease of 2% (from 39% to 37%). We observed that significantly fewer participants in the consistent-experts condition specified they trust no one except themselves in making privacy-related decisions (44%) compared to participants who received inconsistent cues from privacy experts (57%).

To better understand which qualities increased the likelihood that participants were influenced, we asked individuals to select the features they seek in their influencer. The subset of qualities that covers 98% of the responses is: some background in technology, no ulterior motive, reliability, and honesty.

As we mentioned in Section 3.3, we asked participants to specify other groups that would influence their privacy decisions and coded the responses; the breakdown is shown in Table 2. The majority of answers were only coded with one code. Although we specifically asked participants about groups other than friends and experts, most participants nevertheless mentioned that they would (or would not) be influenced by friends or experts. Other frequently mentioned groups or organizations that were not listed in previous questions were law enforcement agencies like the police or the FBI, government officials, media reporting, and non-profit organizations.

The participants' responses also highlighted that it is important to think about what constitutes a friend or an expert. One participant wrote "I'd have to know who the 'privacy experts' and specific 'friends' were" (FI104). Another person stated "I would need more information on what defines this term. What factors are they evaluating to make this decision? How much of an 'expert' are they?" (FC50). Some participants mentioned they would not be influenced by general experts but would be

influenced by domain-specific experts. Others further specified that they would be influenced only by "trusted" friends, or that it would matter whether those friends have experience with a specific technology. As one participant put it, "Friends would be the most influential because I interact with them the most and can hear directly how it affects their lives" (FI60).

A surprisingly large number of participants (178) mentioned family members as important influencers. Their responses mentioned that this includes "someone from the field," a sibling who works in the technology sector, and children who are more knowledgeable. Family members are also most trusted since, as one participant wrote, "they look out for my best interest and can be fully trusted" (EC145).

The fact that coworkers and bosses were mentioned in a number of responses is likely related to the fact that participants were presented with some scenarios that take place in a work environment. Here, for some participants, what "influencing" means seems to shift. One participant said, for example, that management would influence her decisions "because they can force you to do it as part of the job requirement" (EC7).

While the majority of participants replied in a neutral manner, many of those who said they make decisions without any influence (annotated with code "no one") used more negative wording to emphasize their autonomy with respect to questions of privacy. While there was no significant difference in sentiment between the conditions in general, the data shows differences in sentiment when comparing the responses on the annotation level across conditions. On average, all responses except those annotated with "no one" were classified as 16% negative, while the answers with the code "no one" were classified as having a negative sentiment in 62% of cases. Often, the participants with negative sentiments not only rejected the idea that they would be influenced by others in making privacy decisions, but also stressed that they would not allow any of the data collections because of what they seem to represent on a societal level. As one participant put it, "What influences me in these types of situations are the authors of books such as '1984' and 'Brave New World.' Nothing good comes of keeping a population under so much close surveillance" (FI57). Others emphasized their autonomy regarding questions of privacy with statements like "I believe in making decisions all on my own, not based on what others would do" (FC48).

5 DISCUSSION

The literature shows that social cues serve as an effective approach to help people make informed privacy and security decisions [26, 33, 34]. Inspired by past work, we conducted a mixed-design study to test if and to what extent people are influenced by knowing the decisions of their friends or privacy experts in different scenarios. Our results indicated that social cues from both privacy experts and friends influence privacy-related decision making about IoT data collections. A number of factors impact the extent of this influence. For example, we found that a stronger social cue has more influence, especially in balanced scenarios that expose participants to a trade-off between risks and benefits.

5.1 Privacy Experts or Friends?

This study focused on the impact of influence from friends and privacy experts. The wording that we used throughout our scenarios for friends was "friends who use this app," without specifying whether these are friends in real life or friends on social media. Future studies are needed to investigate which groups of friends are more influential than others.

In the current study, we asked participants what qualities would make them most likely to be influenced by a specific group of people. The most frequently mentioned qualities were having a background in technology and not having an ulterior motive. As supported both by prior research and our results, people are, in general, influenced by both privacy experts and their friends, but

differently by each group. We hypothesize that people believe that experts have the knowledge needed to make good privacy decisions related to IoT, but they believe their friends are less likely to have an ulterior motive. The trust people have in experts can be destroyed quickly: Our study participants lost trust in inconsistent social cues from privacy experts significantly faster than they lost trust in inconsistent cues from friends.

Although our analysis did not reveal a significant difference between the extent of the *inferred* influence of privacy experts and friends, our participants *reported* being influenced by privacy experts significantly more than by their friends. Also, participants noted that they are most influenced by the quality of having a technology background, which is usually more prevalent among privacy experts than among friends. A possible explanation for the lack of significant difference in inferred influence is a phenomenon called *expert effect*. Researchers have shown that people's confidence in their own opinions and decisions gradually increase as they are shown social cues from a group of experts [74]. Hence, when people share similar opinions with the experts, they may become less influenced over time by their social cues.

By examining the decision response time, we observed that receiving social cues from privacy experts and friends helped participants make decisions faster. While faster decision time might improve the perceived usability of privacy-choice interfaces and privacy-assistant tools, we must pay attention to the quality and credibility of the social cues. If a decision is made more easily because a user trusted the influencer, individuals may feel betrayed if the recommendation turns out to be against the user's best interest. As the number of daily privacy decisions increases, users may rely more on cues that can speed their decision making. Unlike review-based platforms such as Amazon, where people read other users' comments on a product before making purchasing decisions, people are unlikely to spend much time on each decision when they need to make a large number of real-time privacy-related decisions about pervasive IoT data collection. However, we could imagine a privacy assistant that included social cues as well as links to more information about risks and benefits. Such information might be particularly useful the first time someone encounters a new type of device that is collecting their data, or when they are surprised by the recommendation of the influencers.

5.2 Wisdom of Crowds

"Wisdom of the crowd" refers to the phenomenon that a group of individuals is in aggregate more and better informed than most individuals [68]. Researchers have shown that when uncertain, people look to other people's opinions for information to form their own [25, 30]. As Allen observed, individuals may go along with decisions and beliefs that are expressed by the majority because they think that a crowd's opinion is more likely to be correct than theirs. In some situations, conformity is constructive and appropriate, while in other situations it is not [2] and can even be detrimental [62]. In our study, we found that people are indeed influenced by social cues from the crowd (i.e., friends and privacy experts). However, crowds are not always wiser than individuals. Thus, we need to be careful about the crowds whose opinions we are collecting in order to present to people not to mislead users with incorrect or incomplete information. There are different factors that could make a crowd wiser and more accurate than individuals, including expertise and diversity [84].

A better-informed crowd is likely to provide more useful information. Therefore, it is imperative for crowds to have some expertise and background in privacy and technology, which are also the qualities desired most by participants. Moreover, participants specifically mentioned that they want to know who the experts are and what their level of expertise is. Privacy assistants may leverage the expertise of the crowd in a variety of ways. One approach might be for developers to provide an option to users that allows them to choose the types of expertise and the crowds

from which they want to receive social cues. For instance, as repeatedly mentioned in the open-ended responses, some users trust social cues from non-profit organizations such as EFF or ACLU. There may be other users who want to receive cues from government officials. Another approach would be to incorporate a reputation system similar to Amazon's rating system. In addition, to participate as an expert, participants might be required to have some privacy and technology-related certifications, such as the Certified Information Privacy Technologist certification from the International Association of Privacy Professionals. Qualified experts could then be rated by other system users.

Another important factor to make a crowd wiser is diversity of opinions. Researchers from different fields have found benefit in having different viewpoints within a group [18, 21] and reducing the redundancy of perspectives [14]. Presenting participants with diverse perspectives about a particular data collection and its risks and benefits gives them a broader understanding of the situation and helps them make more informed decisions. In our study, the consistency of (friends' or experts') opinions was a between-subjects factor, while a more diverse platform could include both consistent and inconsistent information.

5.3 Social Influence in Action

Currently, there is no deployed privacy assistant platform that can give its users the ability to opt-out from various IoT data collection across many devices. Our findings provide guidance for the design of effective privacy assistants that can help users make more informed privacy decisions quickly.

In our study, social cues from privacy experts as well as friends influenced people in their decision making regarding IoT data-collection scenarios. However, we also found that this impact is dependent on factors such as the behavior of the influencer, task difficulty, consistency of the social cue, strength of the cue, and self-efficacy. We found that for the scenarios in which the benefits are generally seen to outweigh the risks, people are more influenced by cues from privacy experts, whereas in the scenarios in which the risk to the privacy is dominant, providing people with their friends' cues will have more impact. We also found that people are significantly more likely to be influenced when making decisions about balanced scenarios, which present clear trade-offs between benefits and risks. Other research about social influence has had similar findings [32].

In addition, people will follow social cues significantly more when the influencer is acting consistently with the average response (87% will follow) than when they act inconsistently with the average (21% will follow). This observation is consistent with the term *confirmation bias*. In psychology, this phenomenon is defined as the tendency that people pay more attention to information confirming their own beliefs than information they disagree with [6, 75]. As shown in the literature, people cherry pick the advice from the crowd by focusing on the opinions that are consistent with their own [88]. From the wisdom-of-the-crowd point of view, this approach can be harmful as it will block people from incorporating the majority's perspective. Several explanations have been advanced for why people may resist being swayed by outside influence. For example, it has been shown that people can incorrectly believe that the average judgment in a crowd is no more informed than the average individual's. Holding this belief is significantly related to ignoring the opinions of other people [52]. Another reason why people give more weight to their initial decision is that they know the reasons for their own judgments, but not those behind judgments of the majority. Letting go of one's own judgments and changing one's opinions has been shown to be painful and cause regrets [82]. One way to alleviate this issue is to provide more detailed, yet not overwhelming, information about the decisions.

Another important factor that affects the weight people place on other's opinions is the size of the crowd [66]. In other words, people are more influenced by larger crowds. This is something that should be taken into account by the designers of privacy assistants that provide social cues.

Self-efficacy has been shown to be a reason that people ignore influence from others. Self-efficacy makes people feel that it is unnecessary to yield to others' decisions [63]. We studied this factor and found that participants who expressed a desire to make decisions on their own were significantly less influenced by experts and friends.

The impact of a social cue is context dependent. There were several participants who said they were influenced by their colleagues in the work-related scenarios or influenced by the fire department in the scenarios in which we specifically mentioned fire-hazard prevention as the benefit.

Prior studies have shown that decisions about adopting new technologies are related to trust [31], especially when facing uncertainty or risk [89]. Our results demonstrated that after being exposed to a sequence of cues which are inconsistent with the average behavior, participants lose trust in privacy experts faster than they lose trust in their friends. Many participants also expressed that, while they would listen to arguments, they want to be independent in their decision making—sometimes with strong statements against a perceived bias in cues or towards anyone who would try to make the decision *for* them. This could be partly related to the current news cycle that is dominated by headlines about micro-targeting and companies trying to use personal data to influence web users. However, it also highlights the importance of developing trust in the influencers and the systems presenting the cues to counter potential negative perceptions.

6 LIMITATIONS

We conducted our study using the Mechanical Turk platform. Although the demographic information of our sample of MTurkers was close to the average U.S. population, our sample was not representative of the U.S. population. For instance, MTurkers are both younger, more educated, and more privacy-sensitive than the overall U.S. population [44, 80].

Researchers also worry that MTurkers do not devote full attention to the questions they are asked [35]. To mitigate this issue, we instrumented our surveys with attention check questions for each scenario. Upon examining participants' responses to the attention check questions and their response times, we confirmed the success of our approach. Despite all the limitations of the MTurk population, prior work has confirmed the reliability of the responses [13]. In addition, research has shown that the MTurk population exhibits the same decision making biases as the general population [35].

Another limitation of this study was that we asked participants to imagine themselves in nine hypothetical data collection scenarios followed by social cues. The main reason that we applied the vignette-based methodology was to control the factors that we were interested in studying. We acknowledge that the context of the vignettes was not as detailed or realistic as real-life scenarios. However, we wanted to conduct a carefully controlled study and examine specific relevant factors in simplified data collection scenarios, whereas adding more context to the scenarios would have introduced some confounding factors that we could not control in our statistical analysis. Our study provided statistical evidence that social influence indeed plays an important role in privacy-related decision making. Now that we have demonstrated which effects exist in these scenarios, future work should explore richer and more realistic contexts.

The focus of our study was to understand the impact of social cues from friends and privacy experts. However, there are other interesting groups or individuals mentioned by our participants that are worth investigating in future studies, such as family members or colleagues.

Finally, the described data collections in our study were hypothetical, and hence did not impose any actual risk to the privacy of participants. Therefore, real-world concerns and decisions about IoT data collections may be different from reported behaviors based on perceived risks and benefits. Despite these limitations, we believe our results provide useful insights that can inform privacy assistant design.

7 CONCLUDING REMARKS

As IoT devices become more widespread and people confront choices about personal data collection, the number of decisions to be made may be overwhelming. In this paper, we explored the impact of social cues in helping people make faster, more informed decisions regarding their privacy. To understand how people will be influenced by social cues from privacy experts and friends, we conducted an online user study with 1000 Mechanical Turk participants, randomly assigned to five conditions. We presented each participant with nine hypothetical data collection scenarios. In four conditions, we showed participants what percentage of experts or friends allowed the data collection. In the fifth condition, we described the data collections without any additional information. Our statistical results confirmed the impact of social cues on people making privacy decisions. We also found that the extent of this influence is dependent on various factors such as the level of privacy protectiveness of influencers' decisions and the strength of the social cues. We believe our findings can inform the design of more effective privacy assistants.

ACKNOWLEDGMENTS

This research has been supported in part by DARPA and the Air Force Research Laboratory under agreement number FA8750-15-2-0277. The US Government is authorized to reproduce and distribute reprints for Governmental purposes not withstanding any copyright notation thereon. Additional support has also been provided by Google. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of DARPA, the Air Force Research Laboratory, Google, or the US Government. In addition, the authors would like to thank Professor Alex Davis and Professor Howard Seltman for providing insightful suggestions to improve the statistical methods used to analyze the data.

REFERENCES

[1] Herman Aguinis and Kyle J Bradley. 2014. Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods* 17, 4 (2014), 351–371.

[2] Vernon L Allen. 1965. Situational factors In conformity1. In *Advances in experimental social psychology*. Vol. 2. Elsevier, 133–175.

[3] Vernon L Allen and John M Levine. 1971. Social support and conformity: The role of independent assessment of reality. *Journal of Experimental Social Psychology* 7, 1 (1971), 48–58.

[4] Solomon E Asch. 1956. Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological monographs: General and applied* 70, 9 (1956), 1.

[5] Martina Balestra, Orit Shaer, Johanna Okerlund, Madeleine Ball, and Oded Nov. 2016. The Effect of Exposure to Social Annotation on Online Informed Consent Beliefs and Behavior. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, New York, NY, USA, 900–912. https://doi.org/10.1145/2818048.2820012

[6] Jonathan Baron. 2000. *Thinking and deciding*. Cambridge University Press.

[7] Robert S Baron, Joseph A Vandello, and Bethany Brunsman. 1996. The forgotten variable in conformity research: Impact of task importance on social influence. *Journal of personality and social psychology* 71, 5 (1996), 915.

[8] Christine Barter and Emma Renold. 1999. The use of vignettes in qualitative research. *Social research update* 25, 9 (1999), 1–6.

[9] Debjanee Barua, Judy Kay, and Cécile Paris. 2013. Viewing and controlling personal sensor data: what do users want?. In *International Conference on Persuasive Technology*. Springer, 15–26.

[10] Michael Benisch, Patrick Gage Kelley, Norman Sadeh, and Lorrie Faith Cranor. 2011. Capturing location-privacy preferences: quantifying accuracy and user-burden tradeoffs. *Personal and Ubiquitous Computing* 15, 7 (2011), 679–694.
[11] Andrew Besmer, Jason Watson, and Heather Richter Lipford. 2010. The Impact of Social Navigation on Privacy Policy Configuration. In *Proceedings of the Sixth Symposium on Usable Privacy and Security (SOUPS '10)*. ACM, New York, NY, USA, Article 7, 10 pages. https://doi.org/10.1145/1837110.1837120
[12] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy* 100, 5 (1992), 992–1026.
[13] Michael Buhrmester, Tracy Kwang, and Samuel D Gosling. 2011. Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on psychological science* 6, 1 (2011), 3–5.
[14] Ronald S Burt. 2017. Structural holes versus network closure as social capital. In *Social capital*. Routledge, 31–56.
[15] Scott Camazine. 2003. *Self-organization in biological systems*. Princeton University Press.
[16] Richard Chow. 2017. The Last Mile for IoT Privacy. *IEEE Security & Privacy* 15, 6 (2017), 73–76. https://doi.org/10.1109/MSP.2017.4251118
[17] Robert B Cialdini and Noah J Goldstein. 2004. Social influence: Compliance and conformity. *Annu. Rev. Psychol.* 55 (2004), 591–621.
[18] Robert T Clemen and Robert L Winkler. 1986. Combining economic forecasts. *Journal of Business & Economic Statistics* 4, 1 (1986), 39–46.
[19] Janet Fagan Coleman, Robert R Blake, and Jane Srygley Mouton. 1958. Task difficulty and conformity pressures. *The Journal of Abnormal and Social Psychology* 57, 1 (1958), 120.
[20] Iain D Couzin, Christos C Ioannou, Güven Demirel, Thilo Gross, Colin J Torney, Andrew Hartnett, Larissa Conradt, Simon A Levin, and Naomi E Leonard. 2011. Uninformed individuals promote democratic consensus in animal groups. *science* 334, 6062 (2011), 1578–1580.
[21] Matthew A Cronin and Laurie R Weingart. 2007. Representational gaps, information processing, and conflict in functionally diverse teams. *Academy of Management Review* 32, 3 (2007), 761–773.
[22] Anupam Das, Martin Degeling, Daniel Smullen, and Norman Sadeh. 2018. Personalized Privacy Assistants for the Internet of Things. (2018).
[23] Sauvik Das, Adam D.I. Kramer, Laura A. Dabbish, and Jason I. Hong. 2015. The Role of Social Influence in Security Feature Adoption. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. ACM, New York, NY, USA, 1416–1426. https://doi.org/10.1145/2675133.2675225
[24] Thomas H Davenport. 2013. *Analytics in Healthcare and the Life Sciences: Strategies, Implementation Methods, and Best Practices*. Pearson Education.
[25] Morton Deutsch and Harold B Gerard. 1955. A study of normative and informational social influences upon individual judgment. *The journal of abnormal and social psychology* 51, 3 (1955), 629.
[26] Paul DiGioia and Paul Dourish. 2005. Social navigation as a model for usable security. In *Proceedings of the 2005 symposium on Usable privacy and security*. ACM, 101–108.
[27] Nora A. Draper. 2017. From Privacy Pragmatist to Privacy Resigned: Challenging Narratives of Rational Choice in Digital Privacy Debates: Challenging Rational Choice in Digital Privacy Debates. *Policy & Internet* 9, 2 (June 2017), 232–251. https://doi.org/10.1002/poi3.142
[28] Janna Lynn Dupree, Richard Devries, Daniel M. Berry, and Edward Lank. 2016. Privacy Personas: Clustering Users via Attitudes and Behaviors Toward Security Practices. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 5228–5239. https://doi.org/10.1145/2858036.2858214
[29] Pardis Emami-Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Cranor, and Norman Sadeh. 2017. Privacy expectations and preferences in an IoT world. In *SOUPS '17: Proceedings of the 13th Symposium on Usable Privacy and Security*.
[30] Leon Festinger. 1954. A theory of social comparison processes. *Human relations* 7, 2 (1954), 117–140.
[31] David Gefen, Izak Benbasat, and Paula Pavlou. 2008. A research agenda for trust in online environments. *Journal of Management Information Systems* 24, 4 (2008), 275–286.
[32] Francesca Gino and Don A Moore. 2007. Effects of task difficulty on use of advice. *Journal of Behavioral Decision Making* 20, 1 (2007), 21–35.
[33] Jeremy Goecks, W Keith Edwards, and Elizabeth D Mynatt. 2009. Challenges in supporting end-user privacy and security management with social navigation. In *Proceedings of the 5th Symposium on Usable Privacy and Security*. ACM, 5.
[34] Jeremy Goecks and Elizabeth D Mynatt. 2005. Supporting privacy management via community experience and expertise. In *Communities and Technologies 2005*. Springer, 397–417.
[35] Joseph K Goodman, Cynthia E Cryder, and Amar Cheema. 2013. Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making* 26, 3 (2013), 213–224.

[36] Nigel Harvey and Ilan Fischer. 1997. Taking advice: Accepting help, improving judgment, and sharing responsibility. *Organizational Behavior and Human Decision Processes* 70, 2 (1997), 117–133.
[37] Chip Heath, Richard P Larrick, and Joshua Klayman. 1998. Cognitive repairs: How organizational practices can compensate for individual shortcomings. In *Review of Organizational Behavior*. Citeseer.
[38] Paul Hewson. 2016. Statistical Rethinking: a Bayesian Course with Examples in R and Stan R. McElreath, 2015 Boca Raton Chapman and Hall–CRC 470 pp.,£ 60.99 ISBN 978-1-482-25344-3. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 179, 4 (2016), 1131–1132.
[39] Bert H Hodges, Benjamin R Meagher, Daniel J Norton, Ryan McBain, and Ariane Sroubek. 2014. Speaking from ignorance: Not agreeing with others we believe are correct. *Journal of Personality and Social Psychology* 106, 2 (2014), 218.
[40] David W Hosmer Jr, Stanley Lemeshow, and Rodney X Sturdivant. 2013. *Applied logistic regression*. Vol. 398. John Wiley & Sons.
[41] Arthur Jenness. 1932. The role of discussion in changing opinion regarding a matter of fact. *The Journal of Abnormal and Social Psychology* 27, 3 (1932), 279.
[42] Helmut Jungermann and Katrin Fischer. 2005. Using expertise and experience for giving and taking advice. *The routines of decision making* (2005), 157–173.
[43] Joseph B Kadane and Nicole A Lazar. 2004. Methods and Criteria for Model Selection. *J. Amer. Statist. Assoc.* 99, 465 (2004), 279–290. https://doi.org/10.1198/016214504000000269 arXiv:https://doi.org/10.1198/016214504000000269
[44] Ruogu Kang, Stephanie Brown, Laura Dabbish, and Sara Kiesler. 2014. Privacy attitudes of mechanical turk workers and the us public. In *Symposium on Usable Privacy and Security (SOUPS)*, Vol. 4. 1.
[45] Herbert C Kelman. 1958. Compliance, identification, and internalization three processes of attitude change. *Journal of conflict resolution* 2, 1 (1958), 51–60.
[46] Predrag Klasnja, Sunny Consolvo, Tanzeem Choudhury, Richard Beckwith, and Jeffrey Hightower. 2009. Exploring privacy concerns about personal sensing. In *International Conference on Pervasive Computing*. Springer, 176–183.
[47] B. P. Knijnenburg. 2017. Privacy? I CanâĂŹt Even! Making a Case for User-Tailored Privacy. *IEEE Security & Privacy* 15, 4 (2017), 62–67. doi.ieeecomputersociety.org/10.1109/MSP.2017.3151331
[48] Joseph A. Konstan and John Riedl. 2003. Designing Information Spaces. Springer-Verlag, London, UK, UK, Chapter Collaborative Filtering: Supporting Social Navigation in Large, Crowded Infospaces, 43–82. http://dl.acm.org/citation.cfm?id=937956.937959
[49] Ponnurangam Kumaraguru and Lorrie Cranor. 2005. Privacy indexes : a survey of Westin's studies. *Institute for Software Research* (Jan. 2005). http://repository.cmu.edu/isr/856
[50] J. Richard Landis and Gary G. Koch. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33, 1 (March 1977), 159. https://doi.org/10.2307/2529310
[51] Marc Langheinrich. 2002. A Privacy Awareness System for Ubiquitous Computing Environments. In *Proceedings of the 4th International Conference on Ubiquitous Computing (UbiComp '02)*. Springer-Verlag, London, UK, UK, 237–245. http://dl.acm.org/citation.cfm?id=647988.741491
[52] Richard P Larrick and Jack B Soll. 2006. Intuitions about combining opinions: Misappreciation of the averaging principle. *Management science* 52, 1 (2006), 111–127.
[53] Bibb Latané. 1981. The psychology of social impact. *American psychologist* 36, 4 (1981), 343.
[54] Scott Lederer, Jennifer Mankoff, and Anind K Dey. 2003. Who wants to know what when? privacy preference determinants in ubiquitous computing. In *CHI'03 extended abstracts on Human factors in computing systems*. ACM, 724–725.
[55] Hosub Lee and Alfred Kobsa. 2016. Understanding user privacy in Internet of Things environments. In *Internet of Things (WF-IoT), 2016 IEEE 3rd World Forum on*. IEEE, 407–412.
[56] Hosub Lee and Alfred Kobsa. 2017. Privacy preference modeling and prediction in a simulated campuswide IoT environment. In *2017 IEEE International Conference on Pervasive Computing and Communications, PerCom 2017, Hawaii, USA, March 13-17, 2017*. 276–285.
[57] Hosub Lee and Alfred Kobsa. 2017. Privacy preference modeling and prediction in a simulated campuswide IoT environment. In *Pervasive Computing and Communications (PerCom), 2017 IEEE International Conference on*. IEEE, 276–285.
[58] Linda Lee, JoongHwa Lee, Serge Egelman, and David Wagner. [n. d.]. Information Disclosure Concerns in The Age of Wearable Computing. In *Proceedings of the NDSS Workshop on Usable Security, USEC 2016, Internet Society. February 10, 2016.*
[59] Jialiu Lin, Shahriyar Amini, Jason I Hong, Norman Sadeh, Janne Lindqvist, and Joy Zhang. 2012. Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, 501–510.

[60] Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhimedi, Shikun Zhang, Norman M. Sadeh, Yuvraj Agarwal, and Alessandro Acquisti. 2016. Follow My Recommendations: A Personalized Privacy Assistant for Mobile App Permissions. In *Twelfth Symposium on Usable Privacy and Security, SOUPS 2016, Denver, CO, USA, June 22-24, 2016.* 27–41. https://www.usenix.org/conference/soups2016/technical-sessions/presentation/liu

[61] Steson Lo and Sally Andrews. 2015. To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology* 6 (2015), 1171.

[62] Jan Lorenz, Heiko Rauhut, Frank Schweitzer, and Dirk Helbing. 2011. How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences* 108, 22 (2011), 9020–9025.

[63] Todd Lucas, Sheldon Alexander, Ira J Firestone, and Boris B Baltes. 2006. Self-efficacy and independence from social influence: Discovery of an efficacy–difficulty effect. *Social Influence* 1, 1 (2006), 58–80.

[64] Diane M Mackie. 1987. Systematic and nonsystematic processing of majority and minority persuasive communications. *Journal of Personality and Social Psychology* 53, 1 (1987), 41.

[65] Kathleen M. MacQueen, Eleanor McLellan, Kelly Kay, and Bobby Milstein. 1998. Codebook Development for Team-Based Qualitative Analysis. *CAM Journal* 10, 2 (May 1998), 31–36. https://doi.org/10.1177/1525822X980100020301

[66] Albert E Mannes. 2009. Are we wise about the wisdom of crowds? The use of group judgments in belief revision. *Management Science* 55, 8 (2009), 1267–1279.

[67] Robin Martin, Antonis Gardikiotis, and Miles Hewstone. 2002. Levels of consensus and majority and minority influence. *European Journal of Social Psychology* 32, 5 (2002), 645–665.

[68] Pavlin Mavrodiev, Claudio J Tessone, and Frank Schweitzer. 2012. Effects of social influence on the wisdom of crowds. *arXiv preprint arXiv:1204.3463* (2012).

[69] Pavlin Mavrodiev, Claudio J Tessone, and Frank Schweitzer. 2013. Quantifying the effects of social influence. *Scientific reports* 3 (2013), 1360.

[70] Alex Mayle, Neda Hajiakhoond Bidoki, Sina Masnadi, Ladislau Boeloeni, and Damla Turgut. 2017. Investigating the Value of Privacy within the Internet of Things. In *GLOBECOM 2017-2017 IEEE Global Communications Conference.* IEEE, 1–6.

[71] Tamir Mendel and Eran Toch. 2017. Susceptibility to Social Influence of Privacy Behaviors: Peer Versus Authoritative Sources. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17).* ACM, New York, NY, USA, 581–593. https://doi.org/10.1145/2998181.2998323

[72] Dar Meshi, Guido Biele, Christoph W Korn, and Hauke R Heekeren. 2012. How expert advice influences decision making. *PLoS One* 7, 11 (2012), e49748.

[73] A. Morton and M. A. Sasse. 2014. Desperately seeking assurances: Segmenting users by their information-seeking preferences. In *2014 Twelfth Annual International Conference on Privacy, Security and Trust.* 102–111. https://doi.org/10.1109/PST.2014.6890929

[74] Mehdi Moussaid, Simon Garnier, Guy Theraulaz, and Dirk Helbing. 2009. Collective information processing and pattern formation in swarms, flocks, and crowds. *Topics in Cognitive Science* 1, 3 (2009), 469–497.

[75] Raymond S Nickerson. 1998. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology* 2, 2 (1998), 175.

[76] Helen Nissenbaum. 2004. Privacy as Contextual Integrity. *Washington Law Review* 79 (2004), 119. http://nyu.edu/projects/nissenbaum/papers/washingtonlawreview.pdf

[77] Gabriele Paolacci, Jesse Chandler, and Panagiotis G Ipeirotis. 2010. Running experiments on amazon mechanical turk. (2010).

[78] Sameer Patil, Xinru Page, and Alfred Kobsa. 2011. With a little help from my friends: can social navigation inform interpersonal privacy preferences?. In *Proceedings of the 2011 ACM Conference on Computer Supported Cooperative Work, CSCW 2011, Hangzhou, China, March 19-23, 2011.* 391–394. https://doi.org/10.1145/1958824.1958885

[79] Paul Resnick, Ko Kuwabara, Richard Zeckhauser, and Eric Friedman. 2000. Reputation Systems. *Commun. ACM* 43, 12 (Dec. 2000), 45–48. https://doi.org/10.1145/355112.355122

[80] Joel Ross, Andrew Zaldivar, Lilly Irani, and Bill Tomlinson. 2009. Who are the Turkers? Worker Demographics in Amazon Mechanical Turk. (01 2009).

[81] Muzafer Sherif. 1935. A study of some social factors in perception. *Archives of Psychology (Columbia University)* (1935).

[82] Jack B Soll and Albert E Mannes. 2011. Judgmental aggregation strategies depend on whether the self is involved. *International Journal of Forecasting* 27, 1 (2011), 81–102.

[83] Erin L Spottswood and Jeffrey T Hancock. 2017. Should I share that? Prompting social norms that influence privacy behaviors on a social networking site. *Journal of Computer-Mediated Communication* 22, 2 (2017), 55–70.

[84] James Surowiecki. 2005. *The wisdom of crowds.* Anchor.

[85] Ulf Toelch, Marjolijn J van Delft, Matthew J Bruce, Rogier Donders, Marius TH Meeus, and Simon M Reader. 2009. Decreased environmental variability induces a bias for social information use in humans. *Evolution and Human Behavior* 30, 1 (2009), 32–40.

[86] Ilaria Torre, Odnan Ref Sanchez, Frosina Koceva, and Giovanni Adorni. 2018. Supporting users to take informed decisions on privacy settings of personal devices. *Personal and Ubiquitous Computing* 22, 2 (2018), 345–364.
[87] Trisha Van Zandt and Roger Ratcliff. 1995. Statistical mimicking of reaction time data: Single-process models, parameter variability, and mixtures. *Psychonomic Bulletin & Review* 2, 1 (1995), 20–54.
[88] Ilan Yaniv and Eli Kleinberger. 2000. Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational behavior and human decision processes* 83, 2 (2000), 260–281.
[89] Tao Zhou. 2011. An empirical examination of initial trust in mobile banking. *Internet Research* 21, 5 (2011), 527–540.
[90] Jan Henrik Ziegeldorf, Oscar Garcia Morchon, and Klaus Wehrle. 2014. Privacy in the Internet of Things: threats and challenges. *Security and Communication Networks* 7, 12 (2014), 2728–2742.

Appendices

A SURVEY SCENARIOS

The following is the list of scenarios that were presented to the participants in the control condition. We had three *allow* scenarios (A1 – A3), three *deny* scenarios (D1 – D3), and three *balanced* scenarios (B1 – B3). The actual order of the scenarios were randomized for each participant.

- (A1) You are at a department store. This message is displayed on your smartphone: This store has temperature sensors that check for abnormal temperatures, which indicate potential hazards, e.g., fire. This data will be kept for one day.
- (A2) You are at work. This message is displayed on your smartphone: This building has temperature sensors that check for abnormal temperatures, which indicate potential hazards, e.g., fire. This data will be kept for one day.
- (A3) You are at a library. This message is displayed on your smartphone: This library has presence sensors in each room that are used to determine when to switch the lights on and off to reduce costs and save energy. This data will be kept until the room is no longer occupied.
- (D1) You are at a department store. This message is displayed on your smartphone: This store has a facial recognition system that takes pictures of customers' faces automatically as they enter the store in order to identify returning customers. This method is used to keep track of your orders and make suggestions based on your purchasing habits. Your picture will never be deleted.
- (D2) You are at a library. This message is displayed on your smartphone: This library has an iris scanner that scans customers' irises automatically as they enter the library in order to identify returning visitors. This is used to keep track of your visits and make suggestions based on your habits. Your iris scan will never be deleted.
- (D3) You are in a public restroom. This message is displayed on your smartphone: This restroom has cameras that are recording video of the entire room. The video is shared with law enforcement to improve public safety. This video will never be deleted.
- (B1) You are at the library. This message is displayed on your smartphone: Your smartwatch is keeping track of your specific position. Your position is used by the smartwatch to determine possible escape routes in the case of an emergency. This data will never be deleted.
- (B2) You are at work. This message is displayed on your smartphone: This building uses fingerprint scanners instead of keys to unlock office doors and the break room door. This data is also used to track where employees are in the building. Your fingerprint data will never be deleted.
- (B3) You are in a public restroom. This message is displayed on your smartphone: This restroom has presence sensors to detect whether someone is present. This data is shared with law enforcement to improve public safety and they will keep it for one year.

B SAMPLE SURVEY QUESTIONS

These are the questions that we asked the participants in the experimental condition, which included consistent social cues from privacy experts in the scenarios. Here is a sample scenario in this condition:

- (A1) You are at a department store. This message is displayed on your smartphone: This store has temperature sensors that check for abnormal temperatures, which indicate potential hazards, e.g., fire. This data will be kept for one day. More than 85% of privacy experts allowed this data collection.

B.1 Questions Posed at the End of Each Scenario

- Q1. What type of data is being collected in the scenario? (In three scenarios, we asked about the data type with the following choices: video, audio, specific position, presence, temperature, fingerprint, image of iris, image of face, and other (please specify). In three other scenarios, we asked about the location of data collection with the following choices: coffee shop, workplace, home, library, public restroom, school, department store, and other (please specify). In the remaining three scenarios, we asked about the retention time and the choices were: 1 day, 1 week, 6 months, 1 year, until the room is no longer occupied, until you leave, until the end of the shift, it will never be deleted, and other (please specify).)
- Q2. If you had the choice, would you allow or deny this data collection? (choices: allow, probably allow, probably deny, and deny)
- Q3. (If the answer to Q2 is allow or probably allow) Why would you allow this data collection? (check as many as apply) (choices: I am comfortable with the type of data being collected, I am comfortable with the purpose of data collection, I am comfortable with the length of time for which the data is being kept, I am comfortable with the location where the data collection is happening, I think the data collection is beneficial to me, I think the data collection is beneficial to society, I don't think the data collection will reveal my identity, I think my collected data will be kept securely, I don't see any risk in the data collection, I don't have enough information to make an informed decision, This is what most privacy experts would do, This is what most of my friends would do, The benefits to me outweigh the risks, I think the data collection is required in this situation, other (please specify))
- Q3. (If the answer to Q2 is deny or probably deny) Why would you deny this data collection? (check as many as apply) (choices: I am uncomfortable with the type of data being collected, I am uncomfortable with the purpose of data collection, I am uncomfortable with the length of time for which the data is being kept, I am uncomfortable with the location where the data collection is happening, I think the data collection is not beneficial to me, I think the data collection is not beneficial to society, I think the data collection will reveal my identity, I think my collected data will not be kept securely, I see potential risks in the data collection, I don't have enough information to make an informed decision, This is what most privacy experts would do, This is what most of my friends would do, The risks outweigh the benefits to me, I don't think the data collection is required in this situation, other (please specify))
- Q4. This use of my data would be beneficial to me. (choices: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree)
- Q5. This use of my data would be beneficial to society. (choices: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree)
- Q6. Regardless of whether you would allow or deny the data collection, how confident are you that this was the right decision for you? (answered on a five point scale from extremely confident to not confident at all)

B.2 Questions Posed at the End of Nine Scenarios

- Q7. When considering the 9 scenarios above, how much were you influenced by the decisions that privacy experts made in these scenarios? (we asked about friends in the conditions in which we showed the social cues from friends) (choices: very influenced to do what the experts did, Slightly influenced to do what the experts did, Not influenced at all, Slightly influenced to do opposite of what the experts did, Very influenced to do opposite of what the experts did)
- Q8. (If the answer to Q7 is very influenced to do what the experts did or slightly influenced to do what the experts did) What are the reason(s) you were influenced to do what the privacy experts did when deciding to allow or deny the data collection? (check as many as apply) (choices: I didn't have a strong opinion about allowing or denying the data collection, I generally trust privacy experts when making this kind of decision, I think privacy experts have more technical knowledge about the data collection, I think privacy experts have more background information about the data collection, I usually agreed with the actions that were taken by the privacy experts in this survey, I generally like to find out what other people have done when making a decision, It is easier to do what other people have done than to make the decision on my own, I am not sure why I was influenced, other (please specify))
- Q8. (If the answer to Q7 is very influenced to do opposite of what the experts did or slightly influenced to do opposite of what the experts did) What are the reason(s) you were influenced to do the opposite of what the privacy experts did when deciding to allow or deny the data collection? (check as many as apply) (choices: I didn't have a strong opinion about allowing or denying the data collection, I generally don't trust privacy experts when making this kind of decision, I think I have more technical knowledge about the data collection, I think I have more background information about the data collection, I usually disagreed with the actions that were taken by the privacy experts, I generally like to do the things that are different from what other people do, I am not sure why I was influenced, Other (please specify))
- Q8. (If the answer to Q7 is not influenced at all) What are the reason(s) you were not influenced by privacy experts' actions when deciding to allow or deny the data collection? (check as many as apply) (choices: I didn't have a strong opinion about allowing or denying the data collection, I generally don't trust privacy experts when making this kind of decision, I think I have more technical knowledge about the data collection, I think I have more background information about the data collection, I generally make decisions on my own, I make these kinds of decisions on my own, I usually disagreed with the actions that were taken by the privacy experts in this survey, I would want to know more about the people whose actions are being shown to me before I would trust them, I am not sure why I wasn't influenced, Other (please specify))

(Only for the last scenario: e.g., if the last scenarios was A1)

This is the last scenario you were shown: You are at a department store. This message is displayed on your smartphone: This store has temperature sensors that check for abnormal temperatures, which indicate potential hazards, e.g., fire. This data will be kept for one day.

- Q9. (We keep the level of consensus as before and change the influencers from privacy experts to friends or from friends to privacy experts) If you were told that more than 85% of your friends who use this app allowed the data collection in this scenario, would you allow or deny this data collection? (choices: allow, probably allow, probably deny, deny)
- Q10. (We keep the influencers the same and change the consensus level to the opposite majority decision from more than 85% to fewer than 15% or from more than 65% to fewer than 35%) If you were given the same scenario but told that fewer than 15% of privacy experts allowed the data collection, would you allow or deny this data collection? (choices: allow, probably allow, probably deny, deny)

- Q11. (open-ended question) We have previously shown you how privacy experts and your friends who use this app acted in similar situations. Who are the other people or organizations whose actions would influence yours in scenarios like these? Which would be most influential?
- Q12. For each type of person described below, please specify your level of agreement with the following statement: I would trust [blank] to give me good advice when I need to make a decision about allowing devices to collect and use my information. (choices for blank: privacy experts, my family, my real-life friends, people working in technical fields, my colleagues, my social network friends, no one except myself) (answer choices: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree, not applicable)
- Q13. Please specify your level of agreement with the following statements. (choices: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree)
  - I think privacy experts have more technical knowledge about the data collection than I do.
  - I think my friends have more technical knowledge about the data collection than I do.
  - I think privacy experts have more background information about the data collection than I do.
  - I think my friends have more background information about the data collection than I do.
  - I generally like to find out what other people have done when making a decision.
  - It is easier to do what other people have done than to make the decision on my own.
  - I generally make decisions on my own.
  - I would want to know more about the people whose actions are being shown to me before I would trust them.
- Q14. Please specify your level of agreement with the following statements. (choices: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree)
  - I have sufficient knowledge about privacy to make a decision about allowing my information to be collected and used.
  - I have sufficient knowledge about the technologies mentioned in the scenarios to make a decision about allowing my information to be collected and used.
  - The scenarios generally provided sufficient information about how data would be used to make a decision about allowing my information to be collected and used.
- Q15. What qualities would make you likely to be influenced by a specific group of people when you need to make decisions like the ones in our scenarios? (check as many as apply) (choices: Having some background in technology, Being related to them by blood, i.e. family members, Having some friendship history with them, Knowing them well, Being close friends or family, Being reliable, Being honest, Caring about me, Having no ulterior motive, Other (please specify))

(Demographic Questions)

- Q16. What is your age? (open-ended)
- Q17. What is your gender? (choice: male, female, other, prefer not to answer)
- Q18. What is the highest degree you have earned? (choices: No high school degree, High school degree, College degree, Professional degree (masters/PhD/medical/law), Associates degree, Prefer not to answer)
- Q19. What is your income range? (choices: Less than $25,000/year, $25,000/year - $49,999/year, $50,000/year - $74,999/year, $75,000/year - $99,999/year, $100,000/year - $124,999/year, $125,000/year - $149,999/year, $150,000/year - $174,999/year, $175,000/year - $199,999/year, $200,000/year and above, Prefer not to answer)

Received April 2018; revised July 2018; accepted September 2018