Data Privacy and Pluralistic Ignorance

Emilee Rader
Michigan State University
emilee@msu.edu

Abstract

This paper presents the results of an online survey experiment with 746 participants that investigated whether social norms influence people's choices about using technologies that can infer information they might not want to disclose. The results show both correlational and causal evidence that empirical expectations (beliefs about what others do) and normative expectations (beliefs about what others believe) influence choices to use mobile devices in ways that generate data that could be used to make sensitive inferences. However, participants also reported concern about data privacy, and lower behavioral intentions for vignettes involving more invasive inferences. Pluralistic ignorance is a phenomenon where individuals behave in ways they privately disagree with, because they see others around them behaving the same way and assume this is evidence most people approve of the behavior. These results are consistent with the existence of pluralistic ignorance related to data privacy, and suggest that interventions focused on transparency about data practices are not enough to encourage people to make different privacy choices.

1 Introduction

Every time someone decides to use their mobile device to set an alarm, send a text message, look up directions, or any other action, they are making a choice that has privacy implications, even if they are not explicitly aware of it. Using many technologies for normal, everyday purposes generates data that supports making inferences about a user's body, activities and personal characteristics that are difficult to anticipate and can be surprising, unsettling or harmful when used for unexpected purposes [10, 19, 21].

Many conceptualizations of privacy treat it as an individual right, which means that individuals are responsible for controlling their own information according to their concerns and preferences [24]. With respect to data privacy, this perspective is codified in the logic of "privacy self-management", or notice and choice [32], where privacy is a transaction between the technology user and the platform, system, or organization that is on the receiving end of the data. This individualistic framework limits the potential avenues for influencing people's data privacy decisions to individual-level approaches: increasing knowledge about the implications of making different choices, providing more fine-grained controls and widgets for expressing privacy preferences, and disclosing information about the collecting party's data practices.

But, privacy is inherently social [23]. It is well established that social norms play an important role in interpersonal disclosure decisions [17]. For example, in interpersonal privacy, disclosure rules and boundaries are often norm-based. People learn about appropriate and inappropriate disclosure behavior from others in their family or organizations they belong to, and subsequently form beliefs about what private information looks like and how it should be managed [18].

There is also evidence that behaviors that promote data privacy are subject to social judgments and influence. For example, previous research has found that using encryption to protect one's communications is perceived to be a behavior that makes one seem paranoid [8,33], a stigmatized delusional state characterized by extreme suspicion [22]. And, Solove wrote that the "nothing to hide" argument is an extremely common response to finding out about unwanted data collection (e.g., "I'm not doing anything wrong, so I have nothing to hide"). This argument implies an assumption that only bad actors have reasons to want to protect their data privacy [31], so wanting privacy means one must be a bad actor. These examples illustrate that wanting data privacy is something that could cause a person to be judged negatively by others. However, people still say they value data privacy [20].

USENIX Symposium on Usable Privacy and Security (SOUPS) 2023.
August 6–8, 2023, Anaheim, CA, USA

A norm-based phenomenon called pluralistic ignorance occurs when people engage in a behavior they privately do not believe in or approve of, but they do it anyway because they believe that everyone else approves of it [14]. With respect to data privacy, this could look like privately being concerned about protecting data about oneself, but choosing to use technologies that can generate invasive inferences due to social expectations. If data privacy decisions are subject to this type of social influence, it could mean that interventions that are intended to help people manage their data privacy but are based on individualistic assumptions would fail for reasons that would be hard to identify at the individual level.

This paper presents an experiment investigating whether social norms that conflict with personal privacy preferences influence technology use decisions that have data privacy implications. The results show that social expectations do influence choices to use potentially invasive technologies, despite participants' private concern about data privacy. These results support an interpretation that pluralistic ignorance exists related to data privacy, and suggest that awareness interventions intended to change people's behavior by increasing their knowledge about threats to privacy may be ineffective. This paper contributes novel results to the research literature about social influences on data privacy by showing through a controlled experiment that people may use technologies they feel privacy concern about because of their beliefs about others' approval or disapproval of those technologies.

2 Related Work

2.1 Social Influences on Privacy Choices

Recent research has focused on the idea that information about the behavior and choices of others may be helpful for people faced with making privacy and security decisions. For example, through a participatory design study Chouhan et al. [5] evaluated a mechanism to help people seek security and privacy advice from their community, and found that participants felt like the necessary expertise to help them make decisions did not exist in their circle of close family and friends. Nissen et al. [16] explored participants' reactions to the idea of delegating consent decisions to third parties, and found that trust in the expertise of the third party was an important factor in whether they would delegate or not. Naeini et al. [15] investigated the influence of information about others' privacy choices on participants' choices, and found that information from friends and experts had different effects—the most influential social cues occurred in scenarios where friends denied data collection, and experts allowed it. And, Krsek et al. [13] found that being shown suggestions for security and privacy settings from unknown experts and members of the public influenced participants to self-report that they would choose settings resembling what had been suggested to them. These papers share a common focus on providing social input to specific security and privacy decisions, through providing information about the experiences and behavior of others.

In contrast, the focus of this paper is on the potential that social norms might implicitly influence participants' willingness to use technologies that collect data about them and are capable of making invasive inferences. This paper explores, in a broader sense, whether the influence of social norms may help explain why people continue to use technologies that are bad for privacy, even while they say privacy is important to them.

2.2 Theoretical Background

Research on norms in social psychology focuses on what are referred to as descriptive and injunctive norms. Descriptive norms are based on observing the behavior of others and using that as an example of what one should do [11]. Injunctive norms refer to behaviors that are either reinforced or discouraged through feedback from other people regarding their approval or disapproval of the behavior [6].

It can be difficult to tell whether a collective behavior—one that is observed among many members of a group or community—is caused by one's beliefs about the behavior or beliefs of others. For example, if everyone outside is using an umbrella it may simply be because it is raining and they don't want to get wet. But, there may be a social norm that umbrellas are more acceptable than raincoats in that situation, and it is impossible to know whether the choice to use an umbrella is a result of an individual's personal preference or due to their beliefs about what others would think about their choice of rain gear.

Bicchieri [1] argues that collective behaviors can be independent or interdependent. Independent but similar behaviors arise due to situational factors, whereas interdependent behaviors arise due to social influences. Those social influences can be empirical (e.g., using an umbrella and not a raincoat because that's what one sees others doing) or normative (e.g., using an umbrella because one believes people would think someone who uses a raincoat instead of an umbrella is weird).

A key concept in assigning causation for a collective practice is the type of beliefs that guide behavior. If a group of people behaves in the same way coincidentally, that behavior is not influenced by a social norm. Therefore, we can identify that a collective behavior results from a social norm and is not just coincidental if individuals engage in the behavior because they believe it is commonly done by others, or if they do it because they believe that others approve of the behavior. Bicchieri [1] describes these two types of social influence this way:

- *Empirical expectations* are beliefs about how most others will behave in similar situations, and depend on observing others' behavior (similar to descriptive norms)

- *Normative expectations* are beliefs about what others approve/disapprove of in similar situations (similar injunctive norms)

Pluralistic ignorance is a situation in which people's beliefs about what others approve/disapprove of are incorrect. It occurs where there is a common behavior that people engage in because they see others doing it and believe that this is evidence that they all approve of it. But privately, in fact, most people dislike or disagree with the behavior. Pluralistic ignorance has been implicated in social phenomena as varied as the bystander effect [26], campus alcohol abuse [7], and climate change inaction [12]. Pluralistic ignorance is a visibility problem, where the information people have access to about others' behaviors leads to incorrect assumptions about their beliefs [28]. In the context of climate change, this would look like seeing most people around you driving gasoline engine pickup trucks and assuming that you're the only one who cares about greenhouse gas emissions [12].

A characteristic of pluralistic ignorance is that people believe that they know others' private opinions, but are actually incorrect about what those opinions are [7]. This is recognizable in the discourse about data privacy as the belief that nobody cares about privacy anymore, or that privacy is dead [25], when in fact people do care about privacy [30]. In the data privacy context, this could look like seeing others around you using always-on voice assistants and assuming they must not care about privacy, because if they did they wouldn't use them. It is important to discover whether people's data privacy decisions are affected by pluralistic ignorance, because this would help researchers and practitioners understand what kind of informational interventions would be likely to make a difference. For example, individualistic interventions focused on knowledge about data practices and privacy harms would be less effective in a situation where people's data privacy choices depend on normative assumptions about the behavior and beliefs of others.

2.3 Research Questions

For pluralistic ignorance to exist related to a behavior that an individual engages in, three things must be true. First, individuals have to perceive that the behavior is common among other people. In other words, there has to be an empirical expectation—a belief about what others do—that supports the behavior. Second, individuals have to believe that the behavior is something others approve of. There has to be a normative expectation—a belief about what others believe—that also supports the behavior. And third, individuals have to have a personal expectation—that is, their own, private belief—disliking the behavior, even if they do it (or are likely to do it) anyway themselves. Investigating these three conditions that amount to pluralistic ignorance is the purpose of this study, and each one has an associated research question.

First, the study investigates the relationship between participants' existing empirical and normative expectations and their use of their mobile device in a context that could produce unwanted inferences. The first research question asks whether a person's beliefs about what others do (empirical expectations) and beliefs about what others believe (normative expectations) influence an actual privacy-related behavior.

> RQ1: Is there a relationship between self-reported empirical and/or normative expectations and using a technology that has privacy implications?

Next, the study uses a vignette about a hypothetical mobile device user, similar to the participant, to investigate the influence of empirical and normative expectations that are experimentally manipulated via the vignette on compliance with the use of a mobile device that can make potentially invasive inferences. In other words, the experiment investigates whether norms have a causal influence on the likelihood of complying with the behavior described in the vignette. Using a vignette reduces the impact of social desirability bias, and makes it possible to ask about situations that may contradict the situation in the participant's real life.

> RQ2: Do empirical and normative expectations affect likelihood of compliance with a behavior that produces potentially invasive, unwanted inferences?

Because the vignette includes information about a possible inference that can result from using the mobile device as described, it acts as a kind of awareness intervention explicitly informing participants about this possibility. The third research question asks whether there is a relationship between this intervention and participants' behavioral intentions after learning about the inference.

> RQ3: Is there a relationship between the technology use context, including possible inferences, and behavioral intentions?

3 Method

A survey-based experiment was conducted with 746 participants, hosted on the Qualtrics platform. Data collection took place online during September 20-30, 2021. The Institutional Review Board which oversaw this experiment determined the research to be exempt.

3.1 Vignettes

The experiment involved presenting a very short text vignette (M = 62 words) to participants which described a hypothetical situation. The vignettes each began with the statement,

"Please imagine the following situation and answer the question that follows: Somebody like you lives in a very similar area of the country." The vignettes described the use of a mobile device that collects some type of data and makes inferences about the main character of the vignette, described as someone similar to the participant. The vignettes manipulated the context of the situation, information about the extent to which others use their mobile devices as described (the empirical expectation, a description of what others do), and whether others believe is is OK for people to use their mobile devices in that way (the normative expectation, a description of what others approve of). Details about characteristics of the main character of the vignette were deliberately left vague, to allow the participant to imagine someone similar to them in the ways that were most important or relevant to each individual participant.

The vignettes and experiment design were based on a study by Bicchieri et al. about normative influences on masking and social distancing during the early days of the COVID-19 pandemic [2]. They wrote that it can be difficult to measure the influence of norms on behavior via a survey asking participants about their own behavior, because self-report responses can be affected by social desirability bias. This means that participants' answers may reflect normative beliefs about how one should behave, rather than how the participant thinks they would behave in the situation. There is some evidence in prior work that privacy-preserving behaviors are labeled by others as paranoid or crazy [8, 20, 33], so social desirability bias could be a real problem in this research. By making the vignette about someone else, participants' responses are about others' behavior, not their own. Therefore, they may be willing to answer in a less biased way.

In addition, the vignettes are deliberately simple. The goal of this study was to investigate whether the behavior in the vignette, as imagined by the participant, is subject to social influence. For this study, it does not matter if each participant understands the technology or the vignette behavior slightly differently. The focus of the study is whether there is a social component (empirical or normative expectation) that influences expected compliance with the behavior.

3.2 Experiment Conditions

The experiment had three categorical independent variables: context (3 levels) x empirical expectations (2 levels) x normative expectations (2 levels). It used a full factorial design for a total of 12 between-subjects conditions. Participants were each assigned at random to one of the twelve conditions.

The context dimension refers to the description in the vignette of how the mobile device would be used, and inferences that would be possible due to this use. Three different contexts were used, because previous research has established that privacy is contextual, and privacy-related choices and behaviors depend on context [17]. The contexts in this experiment were based on scenarios from Rader [20], an interview study that investigated participants' reactions to hypothetical scenarios involving unexpected inferences made from sensor-based technologies. The contexts used in this experiment are as follows:

- The *alarm* context focused on using a mobile device as a wake-up alarm. The inference presented in the vignette was that the system could detect how often the user snoozes or sleeps through the alarm. This context was selected because it is a common use case for mobile devices, and an inference that participants in Rader's interview study [20] viewed as directly related to the purpose as a wake-up alarm. They also felt it could be seen as helpful information for changing one's sleep routine or habits.

- The *cookbook* context involved using one's mobile device as a digital cookbook. The inference presented in the vignette was that the system could analyze the foods the user likes to eat and determine how healthy the user is. This context was chosen because keeping recipes on a mobile device is something that is currently possible, but the inference about the user's health is not directly related to the purpose as a cookbook. Participants in Rader's study [20] appreciated the idea of being able to hands-free cooking or possible suggested ingredient substitutions or recipe recommendations to encourage healthier eating habits, but were concerned about unwanted inferences affecting their health insurance or otherwise indicating that they were being judged or evaluated as unhealthy because of the foods they eat.

- The *location* context involved allowing one's mobile device to collect location data that could be used to infer how often the user visits the restroom. This context was chosen because most mobile users allow location data to be collected by apps or their mobile operating system. However, the inference is not tied to a specific purpose for using the mobile device, and is something people would may be uncomfortable with because it violates a taboo about sharing information about one's bathroom behavior. In Rader's study, while some participants imagined that this inference could be useful for one's doctor if the goal was to collect data on a medical condition, nearly all participants had very strong, negative reactions to the idea of a mobile device making inferences about their bathroom habits.

The empirical expectation dimension refers to information in the vignette about how common it is that other residents in the hypothetical community use the mobile device for the purpose described in the vignette. The normative expectation dimension refers to information about how common it is that others approve of using the mobile device for that purpose. In

other words, the vignettes provided information about what other people do in the hypothetical situation (empirical expectation), and also about what other people believe should be done in that situation (normative expectation). Empirical and normative expectations each had two levels, "most" versus "few" other people. An example vignette is shown below, for the alarm (context) x high (empirical expectations) x high (normative expectations) condition. The text of the vignette closely follows the scenarios used in Rader's study [20]. See the replication materials, available online, for the full text of the 12 vignettes used in the experiment.

> Somebody like you lives in a very similar area of the country. **Most / Few** [empirical expectation][1] residents are using their mobile device as their wake-up alarm, which means it is possible for the system to detect how often they snooze or sleep through the alarm. **Most / Few** [normative expectation] residents also believe that it is OK for people to use their mobile device as their wake-up alarm.

3.3 Participants

Participants were recruited using the Qualtrics panel service. Eligible participants were mobile device users 18 years old or older who lived in the United States, and who had not had formal training or worked in a high-tech related field or discipline. The experiment used quotas for age (4.7% 18-20 years old, 41.3% 21-44 years old, 32.9% 45-64 years old, 21.1% 65+ years old) and gender (51% women) based on the 2019 U.S. Census Bureau Current Population Survey, Annual Social and Economic Supplement[2].

2539 participants started the survey by viewing the consent form. 194 declined consent, and 1523 were determined to be ineligible based on their answers to the screening questions. Eight additional participants were excluded when they did not agree to a quality commitment question. Finally, 64 responses were excluded before finishing the experiment where participants reported having "Good" or "Full" familiarity with a made-up word, and 4 more were excluded for answering all of the questions in less than 2 minutes. The final dataset for analysis includes 746 participants who completed the experiment. Participants ranged in age from 18 to 93 (M = 48, SD = 18). 50.7% were women, and 80% reported "White" as one of the ethnicity categories that described them. See the table in the Appendix for additional demographic details about the participants.

On average, it took participants 8 minutes to complete the experiment (SD = 6 min). They were allowed up to 24 hours to finish from the time they started reading the consent form. They received an incentive for completing the experiment in the form of gift cards or in-app credits equivalent to about $2 USD. This amount was determined by representatives of the Qualtrics panel service.

3.4 Procedure

Potential participants received a study invitation via an email message, and clicked on a link that directed them to the online survey. They first viewed the consent form, and after consenting were directed to a series of screening questions to determine their eligibility to participate. Eligible participants were assigned to one of 12 experiment conditions using a random number generator built in to the Qualtrics platform, and subsequently saw and answered questions about only one vignette.

The target size for each condition was 60 participants. The actual number of participants in each condition ranged from 54 to 72 (M = 62, SD = 5.8). The unequal n across conditions resulted from the method of random assignment, and a small number of participants that were excluded after assignment for data quality reasons (i.e., attention check, speeding through the survey). There was no correlation between the number of participants per condition and the number excluded in each condition. See Table 1 for how many participants were assigned to each condition, and the number per condition that were excluded after assignment.

Next, participants were asked a set of 7 questions that varied based on the experiment condition the participant was assigned to. These survey questions were designed based on the Bicchieri et al. study [2], as were the vignettes. Two questions first asked about participants' own past behavior and their beliefs about others' behavior, related to the context:

- *personal behavior*: "Do you [use your mobile device as your wake-up alarm | use an app on your mobile device as a digital cookbook | allow your mobile device to track your location]?" (Yes, No)

- *personal beliefs*: "Do you believe that it is OK for people to [use their mobile device as their wake-up alarm | use an app on their mobile device as a digital cookbook | allow their mobile device to track their location]?" (Yes, No)

An additional two questions asked about their perception of norms related to the context, in the form of empirical expectations about the behavior of others and the prevalence of believing it is OK to behave that way:

- *perceived empirical expectations*: "Please estimate the percentage of fellow residents in your area who [use their mobile device as their wake-up alarm | use an app on their mobile device as a digital cookbook | allow their mobile device to track their location]." (0-100 in increments of 10)

[1] The text in brackets was not presented to participants.

[2] See https://www2.census.gov/programs-surveys/demo/tables/age-and-sex/2019/age-sex-composition/2019gender_table1.xlsx

Empirical Expectations	Context	Normative Expectations: *Low* N	*Low* Excl.	*High* N	*High* Excl.
Low	Alarm	72	2	62	2
	Cookbook	63	7	72	11
	Location	67	2	57	1
High	Alarm	54	7	62	10
	Cookbook	61	10	60	9
	Location	55	0	61	7

Table 1: Number of participants in each condition. *N* denotes the number of participants in each condition, and *Excl.* indicates ineligible participants excluded from each condition. Each participant assigned at random to one of twelve conditions.

- *perceived normative expectations*: "Please estimate the percentage of fellow residents in your area who believe it is OK for people to [use their mobile device as their wake-up alarm | use an app on their mobile device as a digital cookbook | allow their mobile device to track their location]." (0-100 in increments of 10)

Normative expectations are evaluative, and are a person's beliefs about what others believe about what behaviors are acceptable or unacceptable. At face value, it may seem strange to think that normative expectations may exist for different uses of one's mobile device, and even stranger to measure this by asking about the prevalence of people who "believe it is OK" to use their mobile device in a particular way. However, people often comply with norms without being consciously aware they are doing so [1]. And, if the research were about a behavior for which it may be more intuitively obvious that social expectations are important, measuring them in this way might seem more straightforward; e.g., "Please estimate the percentage of fellow residents in your area who believe it is OK for people to smoke cigarettes indoors in public places." This study uses similar phrasing and sentence structure to find out whether norms are at work regarding uses of mobile devices that can generate data with privacy implications.

Also, note that the above questions about participants' personal behavior and beliefs related to the technology use context were asked before the vignette was presented. Asking these questions before presenting the vignette ensures that the responses to questions about participants' current beliefs and behaviors were not affected by the experiment manipulation.

The vignette was presented next, followed by a question about the behavior of the main character in the vignette given the situation described:

- *compliance likelihood*: "How likely is this person to [use their mobile device as their wake-up alarm | use an app on their mobile device as a digital cookbook | allow their mobile device to track their location] in this situation?" (Extremely Unlikely (0) - Extremely Likely (10) in increments of 1)

Two additional questions were asked as follow-ups to the vignette, about the believability of the inference presented in the vignette, and the participant's assessment of whether they would use their mobile device as described in the vignette assuming the inference were possible:

- *believability*: "Please indicate your level of agreement with the statement below. If a person [uses their mobile device as their wake-up alarm | uses an app on their mobile device as a digital cookbook | allows their mobile device to track their location], I believe it is possible for the system to [detect how often they snooze or sleep through the alarm | analyze the foods they like to eat and determine how healthy they are | detect how often they visit the restroom]." (Strongly disagree (1) - Strongly agree (5) in increments of 1)

- *personal use likelihood*: "How likely would you be to [use your mobile device as your wake-up alarm | use an app on your mobile device as a digital cookbook | allow your mobile device to track your location], assuming it is possible for the system to [detect how often you snooze or sleep through your alarm | analyze the foods you like to eat and determine how healthy you are | detect how often you visit the restroom]?" (Extremely Unlikely (0) - Extremely Likely (10) in increments of 1)

The final section of the survey asked questions about participants' concern about data privacy using questions from the collection and use subscales of the Concern for Information Privacy Scale (CFIP) [29], past negative privacy/security related experiences, and internet literacy using items based on a measure by Hargittai [9], as well as other questions not used in this paper. The full survey instrument is available online, with the replication materials for the experiment.

4 Results

4.1 Self-Reported Behavior and Beliefs in Each Context

A majority of participants in both the alarm (177 of 250) and location (148 of 240) conditions answered Yes to the personal behavior question, indicating that they either use their mobile device as a wake-up alarm or allow it to track their location. Most participants across all three contexts also reported that they believe it is OK for people to use their mobile devices for these things. This pattern was most pronounced for the alarm conditions, where 71% of participants said they used their mobile device for this, and 98% said it is OK for others to do so. In contrast, 96% of participants in the cookbook conditions said it is OK to use their mobile device as a digital cookbook,

	Personal Behavior		Personal Beliefs	
	No	*Yes*	*No*	*Yes*
Alarm	29%	71%	2%	98%
Cookbook	75%	25%	4%	96%
Location	38%	62%	28%	72%

Table 2: Percent of participants in each context condition who reported doing the behavior (Personal Behavior) and believing it is OK for others to do the behavior (Personal Beliefs).

	Empirical Expect.			Normative Expect.		
Context	*Mean*	*Med.*	*SD*	*Mean*	*Med.*	*SD*
Alarm	56.9	60	27.3	68.6	80	26.9
Cookbook	32.7	30	23.4	61.5	60	30.9
Location	52	50	22.1	50	50	23.2

Table 3: Descriptive statistics for participants' empirical expectations (beliefs about what others do) and normative expectations (beliefs about what others believe) in each context condition, on a scale from 0 to 100 in increments of 10.

but only 25% of participants said they actually did. And while 72% of participants in the location conditions said it is OK to allow one's mobile device to track one's location, ten percent fewer (62%) said they personally allowed this. These results show that overall, most participants were comfortable with these contexts for using mobile devices, and saw nothing wrong with others using their mobile devices in these ways. Table 2 presents the percent of participants who answered Yes versus No to the questions about personal behavior and beliefs. Replication materials, including the data and code to reproduce the analyses, are available online.

Participants were also asked to estimate the how common the behavior is among other people in their area (perceived empirical expectation), and to what extent people believe that it is OK to do the behavior (perceived normative expectation). For example, participants who saw vignettes about the location context were asked to estimate the percentage of others in their area who allow their mobile device to track their location, and who believe that it is OK to allow this. In the alarm and cookbook contexts, the mean estimated percentage was higher for perceived normative expectations (alarm: 68.6, cookbook: 61.5) than for perceived empirical expectations (alarm: 56.9, cookbook: 32.7). In other words, participants believed that it is more common for people to approve of these uses of mobile devices than to actually engage in using them in these ways. But in the location context, the mean empirical and normative expectations were about the same (empirical: 52, normative: 50). These results show that participants in each context believed there are some social expectations associated with the behaviors; but, they believed fewer people believe location tracking is acceptable than the alarm and cookbook contexts. Table 3 shows the average estimated percentages for participants' empirical and normative expectations in each context.

4.2 Privacy-related choices are correlated with beliefs about others' behavior (RQ1)

Participants' answers to the questions about their personal beliefs and their perceived empirical and normative expectations can be used to identify correlations between these factors and their self-reported behavior. This allows us to investigate whether social expectations are associated with participants' own choices to use their mobile devices as wake-up alarms, digital cookbooks, or to allow location tracking.

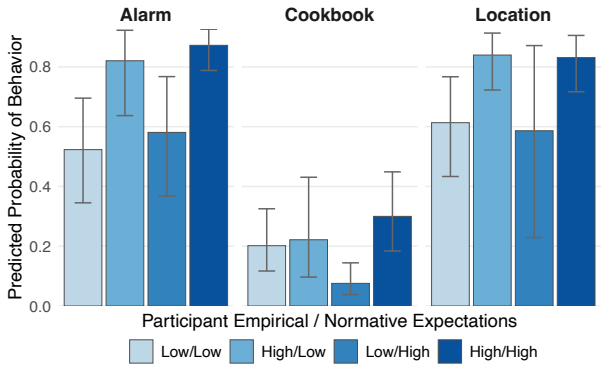

Figure 1: Predicted probabilities from the logistic regression model showing that perceived empirical and normative expectations are associated with a greater likelihood of complying with the behavior in the context condition participants were assigned to (RQ1). Error bars represent 95% confidence intervals.

The perceived empirical and normative expectations variables were each split into high vs. low categories at the median of each variable, so the results of this analysis would be more comparable with the results of the experimentally manipulated empirical and normative expectations analyses presented in the next section. The 2 x 2 high vs. low empirical and normative expectations variables were then recoded into a single categorical variable with four levels for use in the regression model (high empirical/high normative, high/low, low/high, low/low). Note that it is incorrect to assume that low perceived normative expectations (a low estimated percentage of others who believe the behavior is OK) means that a high proportion believe it is not OK. It could be that reporting a low percentage means that participants are not knowledgeable about others' beliefs, or that no norm exists.

A logistic regression model was used to identify factors associated with self-reported use, which was the dependent variable. The model has a categorical predictor for the context (alarm, cookbook or location) and a categorical predictor

for the combination of perceived empirical and normative expectations. It also includes the interaction between context and expectations, and controls for personal beliefs about whether the behavior is OK, demographic variables gender and age, the number of negative security/privacy experiences the participant reported, privacy concern, and internet literacy. If there is a relationship between perceived empirical and/or normative expectations and the dependent variable, then we can conclude that social expectations exist and may influence whether participants use their mobile devices as the study contexts described.

The intercept in the model represents the category combination of alarm (context), low/low (perceived empirical / normative expectations), no (personal belief), and man (gender). Positive coefficients in the model indicate greater odds that the participant would self-report doing the behavior, as compared to the intercept. The model results indicate that believing the behavior is OK has a strong, positive influence on the odds that the participant will report that they do the behavior. The odds that participants would report doing the behavior were 14.6 times higher (*coef* = 2.67) when they believe that it is OK to do the behavior.

However, even with that predictor in the model, a high perceived empirical expectation was associated with a greater odds of doing the behavior. The coefficient of 1.43 for high empirical/low normative expectations indicates that the odds are 4.2 times higher that participants report doing the behavior when they perceive that a high percentage of others do the behavior, even if the perceived normative expectations (belief that others approve of the behavior) are low. When participants perceive that both empirical and normative expectations are high, the odds of reporting the behavior are 6.2 times higher (*coef* = 1.82).

Also of note are the negative coefficients for the cookbook context, and the interaction between the cookbook context and the perceived empirical/normative expectations. While most of the interaction coefficients are not statistically significant, these coefficients do illustrate that participants were in general less likely to report using their mobile device as a cookbook and also perceiving that others do so. This is reflected in the model as lower odds of doing the behavior in the cookbook context even when the empirical expectation is high, in contrast to the the other two contexts. Table 4 presents the full regression results for this model, in the leftmost column.

These results show that participants' empirical expectations—their beliefs about how others use their mobile devices—are related to whether they use their mobile devices in similar ways. However, participants' own normative expectations—beliefs about whether others believe it is OK to do the behavior—did not have any effect beyond the effect of empirical expectations. This can clearly be seen in Figure 1, which presents the predicted probabilities from the model for men with personal beliefs that it is OK to do the behavior. Where empirical expectations

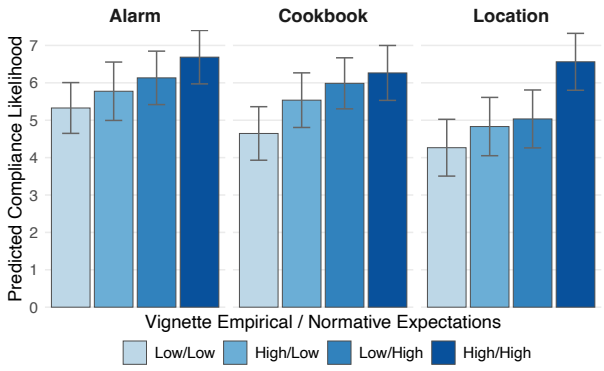

Figure 2: Predicted values from the OLS model showing that experimentally manipulated social expectations increase compliance likelihood (RQ2). Error bars represent 95% confidence intervals.

are low, the likelihood of doing the behavior is lower than where empirical expectations are high. Normative expectations are not associated with greater likelihood beyond empirical expectations. This means that seeing others around them doing the behavior—using their mobile devices as alarms, cookbooks, or allowing location tracking—is strongly associated with the participants doing the behavior themselves.

4.3 Norms support conforming with others' privacy-related choices (RQ2)

The previous model showed that social expectations are probably a factor in participants' choices to use their mobile phones for purposes that may allow sensitive data to be collected about them. However, participants were not asked about their awareness of the possibility of such inferences, and the previous model can only identify correlations between the predictors and the dependent variable. In order to determine whether social expectations are "causally relevant" [2] for participants' behavior, empirical and normative expectations were experimentally manipulated in the vignettes. Each vignette also presented a possible inference that could be made as a result of using one's mobile device as the vignette described. If participants reported a higher likelihood that the main character in the vignette would do the behavior when empirical and/or normative expectations in the vignette are high than when they are low, then we can conclude that social expectations affect the likelihood of compliance with the behavior.

An OLS regression model was used to find out if a causal relationship exists between the experimentally manipulated empirical and normative expectations presented in the vignette and the likelihood that the main character in the vignette would do the behavior. The dependent variable, compliance likelihood, was on an 11 point scale from Extremely Unlikely

	RQ1: Personal behavior (*logistic*)	RQ2: Compliance likelihood (*OLS*)	RQ3: Personal use likelihood (*OLS*)
Context: cookbook	-1.468** (0.479)	-0.681 (0.470)	-1.921*** (0.494)
Context: location	0.369 (0.495)	-1.062* (0.483)	-2.670*** (0.507)
Expectations: high empirical/low normative	1.428* (0.584)	0.448 (0.494)	-1.066* (0.519)
Expectations: low empirical/high normative	0.234 (0.548)	0.805• (0.471)	-0.907• (0.495)
Expectations: high empirical/high normative	1.828*** (0.457)	1.359** (0.474)	-0.606 (0.498)
Personal beliefs: Yes	2.679*** (0.402)	1.298*** (0.354)	2.313*** (0.372)
Gender: woman	0.212 (0.232)	0.241 (0.236)	0.735** (0.248)
Age	-0.043*** (0.008)	-0.024*** (0.007)	-0.031*** (0.008)
Num negative security/privacy experiences	0.216** (0.074)	-0.093 (0.075)	0.130• (0.079)
Privacy concern	-0.269 (0.169)	0.051 (0.166)	-0.362* (0.175)
Internet literacy	0.412** (0.137)	0.232• (0.133)	0.646*** (0.140)
Believability		0.492*** (0.093)	0.683*** (0.098)
cookbook * high empirical/low normative	-1.308 (0.809)	0.443 (0.695)	1.120 (0.730)
location * high empirical/low normative	-0.233 (0.740)	0.118 (0.706)	1.451• (0.742)
cookbook * low empirical/high normative	-1.361• (0.708)	0.535 (0.665)	0.959 (0.698)
location * low empirical/high normative	-0.346 (1.014)	-0.036 (0.684)	0.977 (0.719)
cookbook * high empirical/high normative	-1.301* (0.607)	0.260 (0.681)	1.128 (0.715)
location * high empirical/high normative	-0.696 (0.644)	0.938 (0.676)	0.716 (0.710)
Intercept	-0.731 (0.939)	2.866** (0.984)	3.195** (1.033)
Observations	739	739	739
R^2	0.38 (McFadden's)	0.194	0.333

• p<0.1; * p<0.1; ** p<0.05; *** p<0.01

Table 4: Regression coefficients (and standard errors) for the three regression models. RQ1 focuses on participants' current behavior, RQ2 on their estimate of compliance in the vignette situation, and RQ3 on their behavioral intentions given the information about the inference in the vignette. For RQ1, the empirical and normative expectations are the participant's self-report; for RQ2 and RQ3 they are experimentally manipulated via the vignette. Seven observations with gender category "Not Reported" were excluded from all regressions; these observations were evenly spread across the experiment conditions due to random assignment.

(0) to Extremely Likely (10) in increments of 1 (M = 5.6, *Median* = 6, SD = 3). This model has the same predictors as the previous model, except the empirical and normative expectations experimentally manipulated in the vignette are used instead of the self-reported perceived empirical and normative expectations. This model also has an additional predictor: believability, which represents the participant's evaluation of whether they believe that the inference in the vignette is possible (M = 3.4, *Median* = 4, SD = 1.2). Believability was lower for the location condition (M = 2.90) than the alarm (M = 3.63) or cookbook (M = 3.64) conditions. Table 4 presents the results of this model in the middle column.

High empirical and normative expectations presented in the vignette both caused an increase in compliance likelihood in the experiment. All social expectations categories (high/low, low/high, and high/high) had positive coefficients, indicating an increase in compliance likelihood when compared with the reference category of low/low. The coefficient was smallest and not statistically significant where normative expectations were low (*coef* = 0.44). However, in the low/high category, compliance likelihood was 0.80 points higher than in the low/low condition, and 1.35 points higher when both experimentally manipulated empirical and normative expectations were high (high/high). This indicates that a causal relationship exists between both types of social expectations and compliance likelihood in the experiment.

Like the previous model, the influence of the participant's personal belief that it is OK to use one's mobile device as described in the vignette had a strong, positive influence on compliance likelihood, which was 1.29 points higher when personal belief was Yes than when it was No. Believability was also important: compliance likelihood was 0.49 points higher for each 1-point increase in believability.

Because context was randomly assigned to participant, we can also draw causal conclusions about the impact of the use context on compliance likelihood. The coefficients for both the cookbook context and the location context are negative, indicating that compliance likelihood was lower than in the baseline alarm context. For the location context, the coefficient was large and statistically significant, indicating that compliance likelihood is 1.06 points lower for location vignettes than alarm vignettes. None of the coefficients for the interaction between context and empirical/normative expectations were statistically significant.

Figure 2 shows the pattern of these results in the form of predicted values calculated from the model. In all three context conditions, compliance likelihood is highest where both empirical and normative expectations are high, and lowest

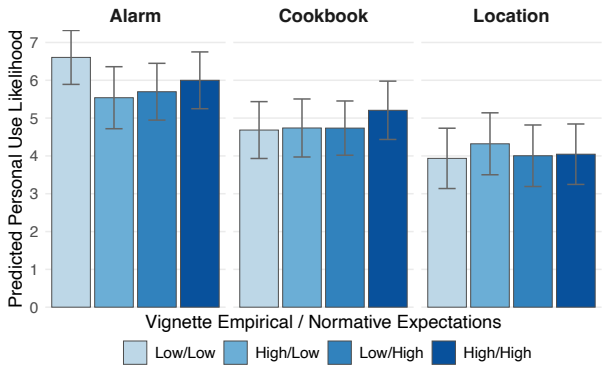

Figure 3: Predicted values from the OLS model showing that participants' estimate of how likely they would be to use their mobile device as described in the vignette decreases as the inferences in the vignette become more invasive and potentially harmful (RQ3). Error bars represent 95% confidence intervals.

where both are low, after controlling for the other variables in the model, including the participant's personal beliefs. The vignettes asked participants to imagine what someone like them would do, assuming the situation in the vignette were true. Compliance likelihood can therefore be interpreted as a measure of what the participants themselves would do [1]. This means that this model provides evidence that not only are social expectations related to participants' behavior, they are causally related. In other words, believing that others use technologies in ways that allow invasive data collection and also approve of doing so increases the likelihood that an individual will also allow this themselves. This provides further evidence that norms exist supporting the use of apps on mobile devices that make potentially unwanted, invasive inferences.

4.4 More invasive inferences are associated with lower behavioral intentions (RQ3)

A final question remains about whether the inferences in the vignettes are really unwanted. Participants in the study showed overall concern regarding data privacy, as measured by items from the collection and use subscales of the Concern for Information Privacy survey instrument (CFIP). The overall mean across all of the questions asked in this experiment was 4.3 out of 5 (higher means more concerned), and there were no differences across the context conditions. But, this does not necessarily mean an objection to the specific inferences mentioned in this study. See the replication materials, available online, for descriptive statistics about participants' responses to the CFIP questions.

In addition to measuring the compliance likelihood of the main character in the vignette, the survey also asked participants to estimate how likely they themselves would be to do the behavior described in the vignette, assuming the inferences were actually possible. This question was essentially about participants' behavioral intentions, measured after an awareness intervention (the vignette) informing them about possible inferences. If participants were opposed to the inferences in the vignette they read, this would be reflected in their measured behavioral intentions.

Context	M	Median	SD
alarm	6.4	7	3.3
cookbook	5.4	5.5	3.1
location	3.5	3	3.2

Table 5: Personal use likelihood descriptive statistics.

An OLS regression model was used to identify whether a relationship exists between the experimentally manipulated empirical and normative expectations and participants' self-report of how likely they would be to use their mobile devices in the way describes in the vignette, (personal use likelihood) if the inferences presented in the vignettes were possible. The model has personal use likelihood as the dependent variable, which used the same response category structure as the compliance likelihood measure in the previous section. The predictors in the model are also identical to the compliance likelihood OLS model. Table 4 presents the results of this model in the rightmost column.

The inference in the alarm context (the baseline context condition) focused on tracking oversleeping, which was expected to be the least concerning inference to participants based on Rader's interview study [20]. In the cookbook context, the inference was about how healthy the participant is based on the foods in the recipes they cook, which could be more concerning but also potentially helpful to someone who wants to adopt healthier eating habits. The inference in the location condition about detecting bathroom behavior was expected to be fairly unacceptable to participants, because it was unacceptable to most of Rader's interview participants. The coefficients in the model for the cookbook context (*coef* = -1.92) and location context (*coef* = -2.67) are both negative, large, and statistically significant, showing the expected pattern.

The means for the personal use likelihood variable, presented in Table 5, clearly illustrate this relationship. On average, personal use likelihood was highest for the alarm context (M = 6.4), lower for the cookbook context (M = 5.4), and lowest in the location context (M = 3.5). These results are an indication that where the inferences are more invasive, participants reported that they would be less likely to behave in ways that would enable the inferences to be made.

In contrast to the other two models, the coefficients for the experimentally manipulated empirical and normative expectations conditions are negative. Only the coefficient for high

empirical / low normative expectations is statistically significant, but it is fairly large (*coef* = -1.06). These results present an inconsistent pattern: if social expectations (empirical or normative) were consistently influential, then this would be apparent in the coefficient for high empirical / high normative as well. In addition, none of the coefficients for the interaction between context and social expectations are statistically significant. The only conclusion that can be drawn from this model is that the social expectations presented in the vignettes do not have a clear relationship with participants' behavioral intentions related to technologies that generate potentially invasive inferences.

As in the previous OLS model for RQ2, believability of the inference had a positive, statistically significant effect on personal use likelihood (*coef* = 0.68). Rader [20] wrote that when inferences were thought to be useful, for example for helping users to correct bad habits or improve their health, the reaction to the inferences was more positive. This could be an indication that participants who believed the inferences were possible may have been more enthusiastic about potential benefits from the inferences. Finally, in this model, like the other two, personal expectation (whether the participant believes it is OK to use their mobile device as an alarm, a cookbook, or to track their location) had a strong, positive effect.

Figure 3 clearly shows that the highest predicted use likelihood is for the alarm conditions, followed by the cookbook conditions and then the location conditions. This is different from the results of the RQ1 logistic model (see Figure 1), where the predicted probability of using one's mobile device as a cookbook was much lower than the other two context conditions. Bicchieri et. al [2] argue that using hypothetical scenarios about a protagonist that is not the participant but is similar to them frees participants from considering the details of their own lives and situations when considering how the person in the vignette would react given the described situation. It could be that participants considered other contextual factors beyond the experimentally manipulated social expectations in the vignette when answering about their own behavioral intentions. These may reflect a positive reaction to the idea of a digital cookbook, but participants may lack actual opportunities to use their devices in this way. It is not surprising that participants in the location condition would be the least comfortable with the associated inference, which was about detecting bathroom visits. Overall, this model provides evidence that the inference awareness intervention presented in the vignette was associated with lower participant willingness to do the behavior where the inferences were more invasive.

4.5 Limitations

This research has several limitations. First, because this is a survey-based experiment, participants' answers to the questions are self-report and reflect their beliefs and their perceptions of their own behavior, but should not be interpreted as direct evidence of their actual or future behaviors.

Second, sampling choices limit the generalizability of the results, in a couple of ways. Participation was limited to people who reported that they did not have high-tech related expertise, because normative influences may be more important factors for non-experts in their choices to use certain technologies or allow data collection. People with training or work experience in a high-tech related field may react differently to the vignettes due to knowing more about how inferences are generated. Also, the experiment used two recruiting quotas, age and gender. This means that the sample is not statistically representative of the internet-using population of the United States for other demographic characteristics like ethnicity, income, education, etc. While this sample has more external validity than a sample where participants were selected entirely based on convenience, this study did not use probability sampling to select participants. Therefore, results should not be generalized to experts, or used to make claims about the broader U.S. population.

In addition, while the selection of the three vignette contexts was based on prior research [20], the specifics of the vignettes participants were asked to react to in the experiment undoubtedly had an influence on the results. It is possible that social influence on data privacy decisions varies by context, and if different contexts had been selected for the experiment the results might have been different.

Finally, the three vignette contexts (alarm, cookbook, location) are different from each other in a number of ways. For example, the alarm and cookbook contexts include a purpose for the behavior (wake-up alarm and digital cookbook), while the location context focuses on a type of information / data (location) without specifying a purpose. The three inferences across the vignettes also vary, in that information about oversleeping may be seen as less of a privacy violation than information about bathroom behavior. This means that while the experiment design supports an interpretation that contextual factors influence privacy-related choices, it is not possible to determine from this experiment which specific contextual factors are responsible for the effect.

5 Discussion

Pluralistic ignorance occurs where individuals engage in a behavior they dislike because of social expectations that they do so. This study investigated whether pluralistic ignorance may be occurring with respect to data privacy and the use of mobile devices. Under conditions of pluralistic ignorance, incorrect beliefs that others approve of the use of potentially invasive technologies would conflict with individuals' privacy preferences and concern. Complying with the social pressure and adhering to the norm would mean behaving contrary to one's own beliefs about privacy.

The results of this study show that a positive correlation exists between using one's mobile device in ways that could compromise one's privacy, and a belief that others use their mobile devices in similar ways. This is one aspect of pluralistic ignorance: believing that a behavior one engages in is commonly done by others too. In addition, the study found evidence of a causal relationship between both empirical and normative expectations and the likelihood of complying with the norm and engaging in the behavior described in the hypothetical vignette. Normative expectations are beliefs about what others believe should be done. The fact that these beliefs were causally related with compliance likelihood in the experiment means that the social influence is more than just imitation—compliance is also affected by the approval/disapproval of others. This is another aspect of pluralistic ignorance: believing that others approve of the behavior.

The third aspect of pluralistic ignorance is a private dislike of the behavior. Participants in the experiment expressed general privacy concern, and also said they would be less likely to use their mobile device as the vignette described where the inferences in the vignette were more invasive and potentially harmful. Taken together, these results suggest that pluralistic ignorance is taking place with respect to data privacy.

A common approach to intervening to help end users make privacy choices that are more consistent with their preferences is providing information to increase awareness of data collection and inferences. This approach is based on the idea that data privacy is an individual decision about control, and education or literacy campaigns can change the inputs to those decisions as a way to change the outcome of the decisions. Many informational campaigns seek to change attitudes towards a behavior by providing previously unknown information about the dangers or harms of that behavior (e.g., alcohol abuse campaigns that focus on persuasion by conveying information about the dangers of alcohol [7,27]). However, the results of this study show that while data privacy results from individual choices, there are social influences on those choices. This means that data privacy is also social, and interventions are needed that take this into account.

Recent research has suggested that peer information sharing might help people make better privacy decisions [5,13,15,16]. Unfortunately, in a pluralistic ignorance situation, sharing information about others' behavior is likely to backfire, because pluralistic ignorance is inherently self-reinforcing. If everyone uses invasive technologies despite privately disliking doing so, sharing information about others' behavior would reinforce the very norm which influences people to behave contrary to their true preferences. In other words, more information about what others do would actually perpetuate the problem by strengthening the norm. Under conditions of pluralistic ignorance, more transparency about others' true beliefs, not their behavior, is needed. However, many approaches to social cybersecurity and privacy do indeed focus on helping individuals via more transparency about others' behavior.

Changing norms under conditions of pluralistic ignorance is hard, because nobody wants to be the first one to express their true preferences and behave contrary to the norm [1]. This prevents people from realizing they are not alone in their dislike of the norm; and it also prevents collective action towards a solution. People cannot coordinate if they all believe they are the only one who thinks they way they do [12]. To encourage people to make different choices, it is necessary to counteract each person's incorrect assumption that everyone approves of the invasive data collection but them.

The most common approach to changing the norm under these conditions is to expose the pluralistic ignorance: people need to know they are not alone, and that others also disapprove and want to protect their privacy and their data. Previous research has found some success in informational campaigns focused not on reasons to adopt the behavior change, but on the true beliefs of others [27]. The goal of this type of intervention is to correct the misperception that each individual is the only one who dislikes the behavior. In addition to informational campaigns, "trendsetters" can also be successful, but only if the conditions are right. A successful trendsetter for counteracting pluralistic ignorance must be an independent thinker, not sensitive to being judged by others, and believe that going against the norm will do some good. They also need to be well positioned in their social network to reach enough people such that when they go against the norm, it makes deviant behavior seem less risky for enough people that the norm falls apart [3].

And here is the final challenge: protecting one's privacy is by nature an action that is not very visible. As a way of combatting pluralistic ignorance, mechanisms must be developed to make protecting privacy more visible without compromising it. For example, the Facebook "I Voted" button was reported to have significantly increased voter turnout in the U.S. in 2010 even in light of voter apathy [4]. Voting is ostensibly a behavior that is private and not observed, potentially making this instance an analogy to interventions focused on expressing one's true beliefs about data privacy. Providing a mechanism that would allow visibility into people's desire to protect their information would also require efforts to put a positive spin choices to protect one's information, to avoid assumptions based on the 'nothing to hide' myth [31] that only bad people have reasons to want to protect their privacy.

Clearly, there are other barriers to improving data privacy than pluralistic ignorance. As much literature has discussed, people do not have a lot of options for truly protecting their data, especially in workplace and education settings. And, notice and choice ensures that people are asked for consent before they have a good idea of what the privacy implications of their consent might be. But, solutions to these issues will arguably be less successful and may even fail if they ignore the role that social expectations play in data privacy.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. CNS-1524296.

References

[1] Cristina Bicchieri. *Norms in the Wild*. Oxford University Press, 2016.

[2] Cristina Bicchieri, Enrique Fatas, Abraham Aldama, Andrés Casas, Ishwari Deshpande, Mariagiulia Lauro, Cristina Parilli, Max Spohn, Paula Pereira, and Ruiling Wen. In science we (should) trust: Expectations and compliance across nine countries during the COVID-19 pandemic. *PLOS ONE*, 16(6):e0252892, 2021.

[3] Cristina Bicchieri and Alexander Funcke. Norm Change: Trendsetters and Social Structure. *Social Research: An International Quarterly*, 85(1):1 – 22, 09 2018.

[4] Robert M Bond, Christopher J Fariss, Jason J Jones, Adam D I Kramer, Cameron Marlow, Jaime E Settle, and James H Fowler. A 61-million-person experiment in social influence and political mobilization. *Nature*, 489(7415):295 – 298, 2012.

[5] Chhaya Chouhan, Christy M LaPerriere, Zaina Aljallad, Jess Kropczynski, Heather Lipford, and Pamela J Wisniewski. Co-designing for Community Oversight: Helping People Make Privacy and Security Decisions Together. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–31, 2019.

[6] Robert B Cialdini, Linda J Demaine, Brad J Sagarin, Daniel W Barrett, Kelton Rhoads, and Patricia L Winter. Managing social norms for persuasive impact. *Social Influence*, 1(1):3–15, March 2006.

[7] Dale T. Miller Deborah A. Prentice. Pluralistic ignorance and the perpetuation of social norms by unwitting actors. *Advances in Experimental Social Psychology*, 28:161–209, 1996.

[8] Shirley Gaw, Edward W. Felten, and Patricia Fernandez-Kelly. Secrecy, flagging, and paranoia: Adoption criteria in encrypted email. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 591–600, 2006.

[9] E Hargittai. An Update on Survey Measures of Web-Oriented Digital Literacy. *Social Science Computer Review*, 27(1):130 – 137, 2008.

[10] Samantha Hautea, Anjali Munasinghe, and Emilee Rader. That's not me: Surprising algorithmic inferences. In *Poster presented at the 2020 Symposium on Usable Privacy and Security*, 2020.

[11] Kyle Irwin and Brent Simpson. Do Descriptive Norms Solve Social Dilemmas? Conformity and Contributions in Collective Action Groups. *Social Forces*, 91(3):1057–1084, February 2013.

[12] Esther Michelsen Kjeldahl and Vincent F Hendricks. The sense of social influence: pluralistic ignorance in climate change. *EMBO Reports*, 19:e47185, 2018.

[13] Isadora Krsek, Kimi Wenzel, Sauvik Das, Jason I Hong, and Laura Dabbish. To Self-Persuade or be Persuaded: Examining Interventions for Users' Privacy Setting Selection. *CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2022.

[14] Dale T. Miller and Cathy McFarland. Pluralistic Ignorance: When Similarity is Interpreted as Dissimilarity. *Journal of Personality and Social Psychology*, 53(2):298–305, 1987.

[15] Pardis Emami Naeini, Martin Degeling, Lujo Bauer, Richard Chow, Lorrie Faith Cranor, Mohammad Reza Haghighat, and Heather Patterson. The Influence of Friends and Experts on Privacy Decision Making in IoT Scenarios. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):1 – 26, 11 2018.

[16] Bettina Nissen, Victoria Neumann, Mateusz Mikusz, Rory Gianni, Sarah Clinch, Chris Speed, and Nigel Davies. Should I Agree? Delegating Consent Decisions Beyond the Individual. Proceedings of the CHI Conference on Human Factors in Computing Systems, pages 1 – 13, 2019.

[17] Helen Nissenbaum. Privacy as Contextual Integrity. *Washington Law Review*, 79:119–158, 2004.

[18] Sandra Petronio. *Boundaries of Privacy: Dialectics of Disclosure*. State University of New York Press, Albany, NY, 2002.

[19] President's Council of Advisors on Science and Technology. Big data and privacy: a technological perspective. Technical report, May 2014.

[20] Emilee Rader. Normative and Non-Social beliefs about sensor data: Implications for collective privacy management. In *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*, Boston, MA, August 2022. USENIX Association.

[21] Emilee Rader, Samantha Hautea, and Anjali Munasinghe. I have a narrow thought process: Constraints on explanations connecting inferences and self-perceptions. In *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*. USENIX Association, 2020.

[22] Nichola J. Raihani and Vaughan Bell. An evolutionary perspective on paranoia. *Nature Human Behaviour*, 3(2):114–121, 2019.

[23] Priscilla M Regan. *Legislating Privacy: Technology, Social Values, and Public Policy*. The University of North Carolina Press, Chapel Hill, NC, 1995.

[24] Priscilla M Regan. Privacy as a Common Good in the Digital World. *Information, Communication & Society*, 5(3):382–405, 2002.

[25] Neil M. Richards. Four privacy myths. In Austin Sarat, editor, *A World Without Privacy: What Law Can and Should Do?*, pages 33–82. Cambridge University Press, 2015.

[26] Rikki H. Sargent and Leonard S. Newman. Pluralistic ignorance research in psychology: A scoping review of topic and method variation and directions for future research. *Review of General Psychology*, 25(2):163–184, 2021.

[27] Christine M. Schroeder and Deborah A. Prentice. Exposing pluralistic ignorance to reduce alcohol use among college students. *Journal of Applied Social Psychology*, 28(23):2150–2180, 1998.

[28] Jacob Shamir and Michal Shamir. Pluralistic Ignorance Across Issues and Over Time: Information Cues and Biases. *The Public Opinion Quarterly*, 61(2):227–260, 1997.

[29] H. Jeff Smith, Sandra J. Milberg, and Sandra J. Burke. Information privacy: Measuring individuals' concerns about organizational practices. *MIS Quarterly*, 20(2):167–196, 1996.

[30] Daniel Solove. The myth of the privacy paradox. *Geo. Wash. L. Rev.*, 89(1), 2021.

[31] Daniel J Solove. "i've got nothing to hide" and other misunderstandings of privacy. *San Diego L. Rev.*, 44:745, 2007.

[32] Daniel J Solove. Introduction: Privacy self-management and the consent dilemma. *126 Harvard Law Review*, pages 1880–1903, 2013.

[33] Justin Wu and Daniel Zappala. When is a Tree Really a Truck? Exploring Mental Models of Encryption. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*, pages 1–16, 2018.

Appendix

	N	%		*N*	%
Age			**Employment**		
18–20	36	5%	Employed full time	243	33%
21–44	305	41%	Employed part time	112	15%
45–64	232	31%	Unemployed or Disabled	182	24%
65+	173	23%	Unemployed or Disabled	182	24%
Gender			Retired	179	24%
Man	361	48%	Student	30	4%
Woman	378	51%	**Income (USD)**		
No Gender Reported	7	1%	Less than $25,000	167	22%
Education			$25,000 to $34,999	122	16%
Some High School	22	3%	$35,000 to $49,999	115	15%
High School Grad	500	67%	$50,000 to $74,999	155	21%
College Grad	152	20%	$75,000 to $99,999	89	12%
Postgraduate degree	72	10%	$100,000 to $149,999	75	10%
Ethnicity			$150,000 to $199,999	13	2%
White	575	77%	$200,000 or more	10	1%
Black or African American	75	10%	**Residential Area**		
Multiple Ethnicities	31	4%	Village or Countryside	113	15%
Hispanic, Latino or Spanish	29	4%	Small or Mid-Size Town	418	56%
Asian or Pacific Islander	26	4%	Large City	215	29%
Native American Alaskan	5	1%			
Other or Not Specified	5	1%			

Table 1: Participant demographics.