# Towards Effective Differential Privacy Communication for Users' Data Sharing Decision and Comprehension

Aiping Xiong
Penn State University

Tianhao Wang
Purdue University

Ninghui Li
Purdue University

Somesh Jha
University of Wisconsin-Madison

***Abstract*—Differential privacy protects an individual's privacy by perturbing data on an aggregated level (DP) or individual level (LDP). We report four online human-subject experiments investigating the effects of using different approaches to communicate differential privacy techniques to laypersons in a health app data collection setting. Experiments 1 and 2 investigated participants' data disclosure decisions for low-sensitive and high-sensitive personal information when given different DP or LDP descriptions. Experiments 3 and 4 uncovered reasons behind participants' data sharing decisions, and examined participants' subjective and objective comprehensions of these DP or LDP descriptions. When shown descriptions that explain the implications instead of the definition/processes of DP or LDP technique, participants demonstrated better comprehension and showed more willingness to share information with LDP than with DP, indicating their understanding of LDP's stronger privacy guarantee compared with DP.**

## I. Introduction

The proliferation and ubiquitousness of pervasive computing has brought an unprecedented amount of collection and analysis of personal information. While such data can be used for personal and societal benefit, improving sustainability, public health, etc., they can also be used in undesired and unexpected ways. These usages can cause adverse consequences for data participants' reputation, insurability, etc., leading to hosts of privacy concerns. People distrust current tools [33], and utilize a variety of measures to protect privacy [26], [38], such as withholding personal information or deliberately providing false personal information, which is detrimental to the utility of the collected data.

To protect data privacy and ensure utility in the context of data publishing, the concept of differential privacy (DP) has been proposed [14], which adds noise to the aggregated result such that the amount of revealed information for any individual is bounded. DP techniques have been deployed by government agencies such as the US Census Bureau for the 2020 census [1]. In recent years, local differential privacy (LDP) has been proposed. LDP differs from DP in that random noise is added at an individual user level before sending the data to the server. Thus, under LDP users do not need to rely on the trustworthiness of the company or the server. LDP has been deployed by companies such as Google [17], Apple [4], and Microsoft [13]. With the increasing deployment of DP and LDP techniques, an interesting and important open question is whether users understand these techniques, trust them, and consequently, increase their data disclosure when these techniques are deployed.

Our work takes a step towards understanding how to *effectively* communicate DP and LDP techniques in order to facilitate users' data disclosure decisions. Centering on textual descriptions of differential privacy techniques, we set out to answer the following five research questions (RQs):

- **RQ 1:** Will participants increase their data disclosure, especially for high-sensitive information, when informed that DP or LDP techniques have been deployed?
- **RQ 2:** To what extent will participants' data disclosure decisions depend on how the privacy techniques are communicated, e.g., descriptions focus on definition or implication?
- **RQ 3:** What factors caused participants to decide whether or not to share their personal information when given description(s) of differential privacy?
- **RQ 4:** Do participants feel that they understand the description(s)? Moreover, which part(s) of the description(s) is difficult for them to understand?
- **RQ 5:** In which ways are participants' objective comprehension of DP and LDP affected by how the techniques are described?

**RQ1** and **RQ2** are about users' *data sharing decisions* when informed that DP and LDP techniques have been deployed. To address them, we conducted Experiments 1 and 2, in which participants made hypothetical data disclosure decisions in a health app survey setting (see Fig. 1). Participants were asked to imagine that they just installed the health app, which needs to collect personal information from them. They were then shown 14 questions asking for personal information, among which half are considered high-sensitive, and the other half are low-sensitive. We varied the presence and absence, as well as the ways of describing privacy techniques.

We ask **RQ3**, **RQ4**, and **RQ5** to understand the *reasons behind users' data sharing decisions*. To address **RQ3** and **RQ4**, we conducted an open-question survey in Experiment 3. The procedure was similar to prior experiments; however, participants made only one high-sensitive data disclosure decision. Following the data disclosure decision, we asked each participant to explain *why* they decided to share or not share their personal information. Participants also rated whether the given differential privacy description was easy to comprehend. For participants who indicated the description was not easy to understand, we asked them to highlight the part or parts that were difficult to comprehend. We conducted Experiment 4 to assess participants' *objective* comprehension of DP and LDP

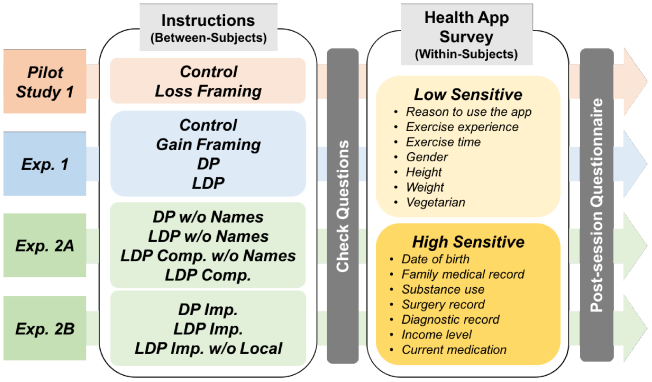


Fig. 1: Flow chart shows the experimental design for experiments of data sharing decisions (Group 1). Pilot Study 1 validated the health app data collection setting. Experiments 1, 2A, and 2B addressed RQ1 and RQ2. "Instructions" box presents the conditions in all those experiments. "Health App Survey" box lists the seven low-sensitive and the seven high-sensitive questions.

(**RQ5**). Based on lessons learned from previous experiments, we also added new descriptions for each technique that explain data flow and implication inferences. Participants were shown descriptions of DP or LDP, then answered five questions about the privacy and utility consequences. We compared the correct answer rates of those questions between the two new descriptions and descriptions from prior experiments.

DP and LDP are typically used in different settings. Each participant in our study was exposed to only one of them. In terms of privacy, LDP provides stronger protection than DP does, because LDP does not need to trust the server. Our experimental setting (the health app data collection testbed) also focused on the privacy protection provided by the techniques to minimize other confound factors in the experiment design. In this case, we expect a higher data disclosure rate under LDP. When that did not happen, it could be an indicator that many participants do not really understand the nature of protection from the descriptions. We obtained answers for each question as follows:

**RQ1:** *Data Sharing under Differential Privacy.* Participants increased data sharing for the high-sensitive questions when they were informed of protection from differential privacy, indicating a positive effect of communicating privacy techniques.

**RQ2:** *Data Sharing with Different Descriptions.* When descriptions focused on definition and/or data perturbation processes, participates' data sharing rates were not larger for LDP than for DP. Nevertheless, higher data disclosure rates were obtained for the LDP conditions than for the DP condition when the implications were communicated, i.e., whether the privacy protection relies on the trustworthiness of the company or the server.

**RQ3:** *Reasons behind Sharing Decisions.* About half of the participants chose to share their personal information. Most of them explained that they made the decision because of the described privacy protection, suggesting that trust in privacy protection techniques led to the decisions to share. Participants who decided not to share cited various concerns, top three of which are 1) the requested information are too sensitive to share, 2) distrust of the described differential privacy techniques, and 3) risks of data breach in the future.

**RQ4:** *Subjective Measure of Comprehension.* Only 13% of the participants indicated that they had difficulty in understanding the described techniques, and the difficult parts mentioned the most were about the data perturbation processes.

**RQ5:** *Objective Measures of Comprehension.* Better comprehension results were obtained for descriptions that provide implication inferences than those which do not.

Finally, we discuss how the obtained results inform our understanding of effective differential privacy communication and highlight implications of our findings in Section VIII.

To summarize, our work makes the following contributions:

- We provide *quantitative* and *qualitative* evidence showing benefits (increasing data sharing) of communicating differential privacy to users.
- We identify the data perturbation processes as the most difficult parts for laypeople to understand and provide evidence showing implication descriptions as one effective way (i.e., larger data disclosure rates and better objective comprehension results) for DP and LDP communication.
- We further uncover the effect of implication descriptions on comprehending differential privacy with data flow descriptions which afford privacy and utility implication inferences.
- We reveal a robust effect of information sensitivity in participants' data disclosure decisions even with privacy-enhancing techniques, suggesting biased responses within non-mandatory data collecting.

## II. Background and Related Work

### A. Background on Differential Privacy

Differential Privacy [15] applies in the setting where there is a trusted data curator, who gathers data from individual users, processes the data in a way that satisfies DP, and then publishes the results.

**Definition 1** (Differential Privacy)**.** *An algorithm* $\mathbf{A}$ *satisfies* $\epsilon$*-DP, where* $\epsilon \geq 0$*, if and only if for any two datasets* $D$ *and* $D'$ *that differ in at most one record, and any set* $\mathbf{R}$ *of possible outputs of* $\mathbf{A}$*, we have*

$$\Pr[\mathbf{A}(D) \in \mathbf{R}] \leq e^{\epsilon} \Pr[\mathbf{A}(D') \in \mathbf{R}]$$

The definition prevents a strong adversary who knows all but one record in the database from inferring the last one after seeing the output. To ensure that, $\mathbf{A}$ first obtains the true result from $D$, and then adds noise to the result.

In the local setting, each user perturbs the input value $v$ using an algorithm $\mathbf{A}$ and reports $\mathbf{A}(v)$ to the aggregator.

**Definition 2** (Local Differential Privacy)**.** *An algorithm* $\mathbf{A}(\cdot)$ *satisfies* $\epsilon$*-local differential privacy (*$\epsilon$*-LDP), where* $\epsilon \geq 0$*, if and only if for any input* $v, v'$*, and any set* $\mathbf{R}$ *of possible outputs of* $\mathbf{A}$*, we have*

$$\Pr[\mathbf{A}(v) \in \mathbf{R}] \leq e^{\epsilon} \Pr[\mathbf{A}(v') \in \mathbf{R}]$$

In both DP and LDP, $\epsilon$ plays an important role as it measures the randomness of the process. A large $\epsilon$ leads to insufficient noise, which does not provide much privacy protection.

**Difference between DP and LDP.** In DP, the server has access to the true sensitive values of the users, while in LDP, the aggregator does not see the actual private data of each individual. Instead, users send randomized information to the aggregator, who infers the data distribution based on that. However, the better trust model also comes at the cost of utility: with the same privacy guarantee, measured by the parameter $\epsilon$, the utility of LDP is worse than DP by a factor of $\Theta(\sqrt{n})$ [10], where $n$ is the number of users.

**Deployment of DP and LDP.** Although DP was proposed more than a decade ago, the first public deployments of this concept are related to LDP, e.g., companies like Apple [4], Google [17], and Microsoft [13]. Exemplary use cases include collecting users' default browser homepage and search engine to understand the unwanted or malicious hijacking of user settings; or gathering frequently typed emojis and words to help predict keyboard typing.

More recently, DP is also deployed in industry, government, and academic. In particular, Uber released an open-source project for SQL query with differential privacy [24]; LinkedIn proposed a system to analyze user information with DP [29]; the US Census Bureau has deployed DP technologies for the 2020 census [1]; and Harvard built a system prototype for researchers to share data using DP [19].

With the deployments of DP and LDP, we started seeing companies and organizations begin to communicate DP and LDP techniques to the public, including Apple, Google, Microsoft, Uber, and US Census Bureau. We took descriptions from the companies and organizations mentioned above, made minor modifications to fit our context, and used them in our experiments.

### *B. Related Work*

**Usability of DP and LDP.** One primary goal of addressing the RQs is to achieve "usable differential privacy". The most closely related prior work is Bullek et al. [8], which studied people's comprehension of the random response method [45] for LDP and preference of the privacy parameter. In that study, each participant perturbed answers for sensitive questions with three probabilities, corresponding to three $\epsilon$ values. For a final high-sensitive question, participants were asked to first choose the perturbation probability and then answer. 75% of the participants chose the largest perturbation (which obscured their true answers the most), 5% chose the intermediate one, and 20% chose the least perturbation. Subjective reasons provided for selecting the least protection focused on a desire to respond truthfully. One interpretation of these results is that most people want strong privacy protection, but a cognitive bias to equate such data perturbation with "lying" (data-obfuscation) can sway privacy-related decisions.

Our work is orthogonal to [8]. Instead of focusing on the parameter value, we strive to convey the qualitative nature of differential privacy. We study people's willingness to share personal information when given different descriptions of DP or LDP, and the reasons behind those decisions. While the privacy parameter $\epsilon$ critically affects the level of privacy protection, it seems unlikely that end users can choose $\epsilon$ based on understanding how the mechanism works and assess the impacts of different $\epsilon$ values. It is more likely that they will rely on expert assessment of the appropriateness of deployed $\epsilon$ values, and accessible explanations of the consequences. This is similar to how privacy policies are used. In practice, people are not reading privacy policies because they are long and full of legalese [32]. At the same time, when a company's privacy policies and practices are inadequate, this will often be discovered by experts, and lawsuits may ensure [9], [34].

**Usability of Privacy Notices.** A large body of work has been conducted to inform users about privacy techniques [27], [28] and to facilitate their privacy decisions [20], [42]. For example, when privacy policies of online shopping sites were made more prominent and accessible within a shopping search engine interface, participants increased their purchase intention with the sites offering better privacy protection [42].

To improve the usability of privacy from the user's perspective, "privacy by control" through notice and choice have become essential for privacy protection [39]. Notice and consent as a principle is widely recognized by law and society. For example, companies such as Google and Apple have implemented permission dialogs in Android and iOS to request access to hardware and data from users. Felt and colleagues [18] investigated the effectiveness of the Android permission system in warning users of app installation risks. Their results showed that most participants did not pay attention to permission warnings or did not understand what the permissions mean. Lin et al. [31] examined permission warnings in helping participants manage the privacy of installed apps. Their results showed that designs that highlight privacy implications, e.g., unexpected data collection practices, were effective in helping participants avoid intrusive apps.

The communication of DP or LDP also deals with the usability of privacy notices. However, it has a unique challenge because of its mathematical complexity. Thus, we start from definitions and explore how to remove the technical details while preserving the fidelity of the communication.

**Decision Making of Online Data Sharing.** Our study centers on people's data sharing decisions, which are affected by various factors. People's decision making in risk contexts are influenced by how a problem is framed [43], [44]. Specifically, if the outcomes are described in terms of potential loss (negative framing), people are risk-seeking. However, people are risk-averse when the outcomes are presented in terms of potential gains (positive framing).

A different way of establishing a frame of reference involves emphasis framing, which accentuates a subset of potentially relevant considerations [16]. For example, the consequence of online data sharing can be framed positively in terms of free product and service, or negatively in terms of loss

due to privacy concerns. While some prior studies provided evidence that the emphasis framing influenced people's privacy decisions [3], some did not [21], requiring further research. A scrutiny of the prior studies revealed differences on information sensitivity level. While highly intrusive information, such as drug use, was asked in [3], most information examined in [21], such as height and time of exercise, were less sensitive.

Privacy issues arise in the specific contexts [11], [35], [36]. Prior studies revealed that many health apps had privacy risks to users [12], [23], and caused low engagement of users due to privacy concern [30], [41]. So we chose a health app survey setting as the testbed to evaluate participants' data sharing decisions in the current study. While DP provides better utility, we note that this is mainly beneficial to the server, rather than the users. LDP provides better privacy promise than DP does in the health app data sharing context, which would be preferred by users [8]. We also varied the sensitivity level of the health information across survey questions and evaluated the effect of a negative framing in terms of privacy risk or a positive framing in terms of benefit on participants' data sharing decisions.

## III. Overview of Experiment Design

We ran a series of online experiments, which can be divided into two groups. Experiments in Group 1 (including Pilot Study 1 and Experiments 1, 2A, 2B) focus on *decision* measures of whether participants were willing to share their personal information under different conditions. Experiments in Group 2 (including Experiments 3, Pilot Study 2, and Experiment 4) focus on more in-depth understanding of the *reasons* behind participants' data sharing decisions and their *comprehension* of DP and LDP. We ran multiple studies in part because findings in earlier studies led to interesting questions that we sought to answer with additional studies.

In this section, we describe the method of participant recruitment, design of differential privacy descriptions, and the testbed of health app data collection. Experiments in Group 1 all used the same procedure, which is described in Section IV. Experiments in Group 2 used different task procedures, which are explained in Sections VI and VII, respectively.

### A. Participant Recruitment

All experiments were conducted on Amazon Mechanical Turk (MTurk), and the human intelligent task (HIT) was posted with restrictions to US workers with at least 95% approval rate and 100 or more approved HITs. We made these restrictions in the studies to accurately represent sample restrictions of most recent MTurk research [22]. All experiments complied with the American Psychological Association Code of Ethics and were approved by the Institutional Review Board at the authors' institutes. Informed consent was obtained for each participant. Data of the experiments were anonymized before analysis.

### B. Differential Privacy Communication Design

To come up with the descriptions of DP and LDP to be used in the study, we started from the descriptions published by the companies and organizations that deployed these techniques, and then conducted multiple rounds of internal discussion and review of the descriptions. In the discussions, we involved experts of differential privacy to ensure that our descriptions of DP and LDP are technically accurate, and laypeople to help ensure that they can be understood. As mathematical rigour is one key strength of DP and LDP, we decided not to shy away from using mathematical terms such as "probability" or "aggregated data" in some of the descriptions. The full descriptions of all conditions used in the experiments are given in Table XII from Appendix D.

To verify and enhance participants' understanding, we added one check question asking participants to recognize the presented technique immediately after each textual description (see Appendix A). For participants who did not answer the question correctly, we presented the corresponding description again. We asked the same check question in the post-session questionnaire evaluating the effect of the second presentation.

### C. Heath App Data Collection Setting

For each experiment, we presented the same health app data collection scenario in which each participant was instructed to play the role of an health app user in three steps (see Appendix A for the details). Within the instructions, we present examples of collected data (e.g., age and gender) at the local app and the app server to let participants better understand the health app data collection and then situate themselves in the hypothetical setting we created.

## IV. Experiment 1

Before Experiment 1, we conducted Pilot Study 1, which used the health app data collection setting described above. We had the following findings from Pilot Study 1. Participants showed less willingness to answer the high-sensitive questions than the low-sensitive ones. When the loss framing (explaining privacy threat of data sharing) was presented, participants' data disclosure was reduced regardless of question sensitivity, and the reduction rate was larger for the high-sensitive questions. Thus, we obtained the effect of question sensitivity and the framing effect [3], [5], [21], confirming the health app data collection setting and hypothetical willingness to disclose personal information as a testbed to evaluate privacy decisions. See Appendix B for additional details of Pilot Study 1.

Since the benefit of data disclosure in lieu of privacy threat is often emphasized in the wild, we chose to mention the benefit of data sharing (i.e., the gain framing) in Experiment 1. We evaluated the effect of communicating differential privacy techniques on participants' data sharing decisions (**RQ1**) with four between-subjects conditions: *Control*, *Gain Framing*, *DP*, and *LDP*. We predicted a main effect of question sensitivity as Pilot Study 1. With an emphasis on possible benefit, we expected that participants in the *Gain Framing* condition would become less concerned about their privacy, and thus would increase data sharing compared to the *Control* condition. With extra privacy protection in the *DP* and *LDP* conditions,

participants would increase their data disclosure further, more for the *LDP* condition with better privacy guarantee.

### A. Participants and Stimuli

We recruited 598 Amazon MTurk workers. Each participant was paid 1 US dollar for completing the study (median completion time: about 289 seconds) The payment rate was the same for all experiments except Experiment 3 (see details in Table XI). The descriptions of *DP* and *LDP* used in the study focused on definitions, and we listed the organizations which have implemented the techniques (see Table XII).

### B. Procedure

After accepting the HIT, all participants were directed from MTurk to a survey on Qualtrics, and were assigned to one condition randomly. At the beginning of all conditions except *Control*, we emphasized the benefit of sharing personal information. The study continued with a goal description. Following the three-step health-app data collection scenario, the corresponding differential privacy description was presented in the *DP* or *LDP* condition (see Appendix A for the detailed descriptions). Then, participants in all conditions answered their data sharing decisions for 14 questions. Consistent with [3], [21], we divide them into seven *low-sensitive* questions (i.e., the reason to use the health app, exercise experience, exercise time, gender, height, weight, vegetarian) and seven *high-sensitive* questions (i.e., date of birth, family medical record, substance use, surgery record, diagnostic record, income level, current medication).

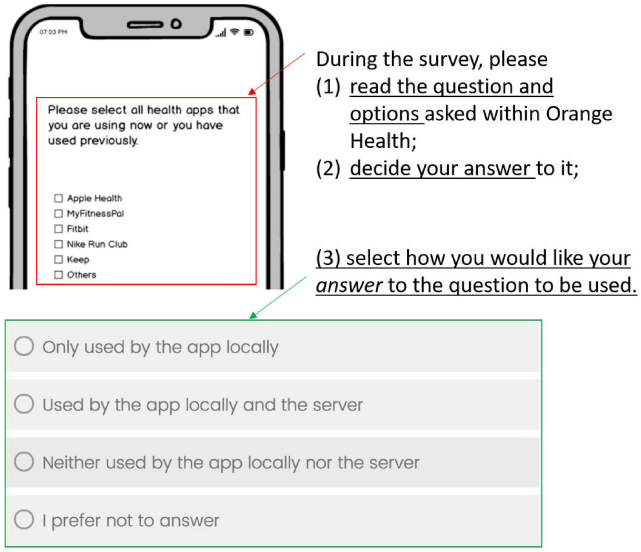


Fig. 2: Instructions of survey questions in Experiments 1, 2A and 2B. "with DP" was added to end of options 2 and 3 for all DP related conditions, and "with LDP" was added in the same way for all LDP related conditions.

Each question with its options was presented within a smartphone layout (see Fig. 2). Participants were instructed that their task was to read the survey question, decide their answer to it, and select how they would like their answer to be used. Note that we did not ask participants to actually provide the answers. We distinguished two types of usage for the collected data, local and at the app server. So for each question, participants were asked to decide whether they would like the data being used (a) locally only, (b) both locally and at the app server, (c) neither, or (d) prefer not to answer.

The 14 questions were presented randomly in each condition. After answering the 14 questions, participants completed a questionnaire that asked for demographic information (e.g., age, gender, education, and computer-science background). We also asked participants to indicate their agreement with statements on their trust for the app and the server on a 7-point Likert scale (1: strongly disagree, 7: strongly agree), respectively (see Appendix A).

For the *DP* and *LDP* conditions, a check question (see Appendix A) was presented after the differential privacy description and before data sharing decision-making. For participants who did not answer the question correctly, we presented the corresponding description again. Participants answered the check question again at the end of the post-session questionnaire. They also indicated their trust level for DP or LDP on the 7-point Likert scale. We did not obtain any significant difference of the trust evaluations except that there was a main effect of condition in the current experiment, $\chi^2_{(3)} = 11.44, p = .009$. Participants in the *DP* condition (65.3%) showed more trust than participants in the *Gain Framing* condition did (50.5%), $p_{adj} = .007$. Thus, we omit the results of trust evaluation but discuss them in the General Discussion (Section VIII).

### C. Results

Participants were excluded from data analysis using two criteria: duplicate IP address and overall completion time less than 120 seconds. Due to the main interest in the effect of differential privacy communication, we also excluded participants who did not answer the second check question correctly. The number of participants excluded from data analysis were listed in Table XI. Consequently, 87 participants from the *Control*, 101 from the *Gain Framing*, 150 from the *DP*, and 127 from the *LDP*, were included in the data analysis. The demographic distributions were similar between conditions. See Table I for descriptive statistics.

We measured the selected option of each question for each participant. Decisions were coded as *Opt out* when participants chose "Neither used by the app locally nor the server" or "I prefer not to answer". Choices of *Local only* ("Only used by the app locally") and *Both* ("Used by the app locally and the server"), as well as *Opt out* decisions were determined of each question for each participant.

*Opt out* decision, selection for *Local only* option, and choice of *Both* option collapsed across participants (see the first column of Fig. 3) were entered into a 2 (question sensitivity: *low-sensitive*, *high-sensitive*) × 4 (condition: *Control*, *Gain Framing*, *DP*, *LDP*) chi-squared tests with a significance level of .05, respectively. Post-hoc tests with Bonferroni corrections [6] were performed, testing all pairwise comparisons with corrected $p$-values for possible inflation. We report the

statistics only for the significant effects in the text. Please refer to Table VII for the full results of statistical tests.

TABLE I: Demographics of participants in each experiment. The number of participants of each experiment was listed in the brackets on the top row. EXP. means Experiment.

| Item | Options | EXP.1 (465) | EXP.2A (581) | EXP.2B (468) | EXP.3 (278) | EXP.4 (540) |
|---|---|---|---|---|---|---|
| Gender | Male | 56.8% | 50.4% | 47.9% | 55.0% | 52.6% |
| | Female | 43.0% | 49.4% | 51.3% | 44.6% | 46.7% |
| | Other | 0.2% | 0.2% | 0.6% | 0.4% | 0.2% |
| | Not to answer | 0% | 0% | 0.2% | 0% | 0.6% |
| Age | 18-24 | 8.2% | 7.1% | 6.6% | 10.1% | 8.9% |
| | 25-34 | 47.7% | 42.7% | 44.9% | 35.3% | 34.4% |
| | 35-44 | 24.9% | 28.7% | 29.1% | 26.6% | 23.0% |
| | 45-54 | 10.3% | 13.3% | 11.1% | 15.5% | 16.5% |
| | 55 or older | 8.6% | 7.9% | 8.3% | 12.6% | 16.9% |
| | Not to answer | 0.2% | 0.3% | 0% | 0% | 0.4% |
| Education | No high school | 0% | 0.2% | 0.2% | 0.4% | 0.4% |
| | High School | 27.7% | 20.5% | 23.7% | 25.9% | 24.6% |
| | College/Bachelor | 60.20% | 65.9% | 62.4% | 59% | 59.1% |
| | Masters/Ph.D. | 10.8% | 11.9% | 12.0% | 10.8% | 13.9% |
| | Medical degree | 0.6% | 0.9% | 1.7% | 1.8% | 0.4% |
| | Not to answer | 0.6% | 0.7% | 0% | 2.2% | 0.7% |
| CS Back | Yes | 22.8% | 19.4% | 28.8% | 20.5% | 20.0% |
| | No | 75.9% | 79.3% | 70.3% | 76.7% | 78.5% |
| | Not to answer | 1.3% | 1.2% | 0.9% | 2.9% | 1.5% |

*1) RQ1. Effect of Differential Privacy Descriptions:* With extra *DP* and *LDP* descriptions, participants' overall decision rates were similar to that of the *Gain Framing* condition, but they showed more willingness to share high-sensitive information.

**Opt out rate.** Participants opted out more for the high-sensitive questions (20.2%) than for the low-sensitive questions (6.0%), $\chi^2_{(1)} = 290$, $p < .001$. Neither the main effect of condition (*Control* vs. *Gain Framing* vs. *LDP* vs. *DP*: 12.9% vs. 14.3% vs. 11.8% vs. 13.5%), nor its interaction with question sensitivity was significant, suggesting little framing effect or the effect of differential privacy communication.

**Local only selection rate.** The selection rate was larger for the high-sensitive questions (41.8%) than for the low-sensitive questions (32.8%), $\chi^2_{(1)} = 43.38, p < .001$. The selection rates differed across conditions (*Control*: 42.9%, *Gain Framing* 35.5%, *LDP* 38.5%, *DP*: 32.4%), $\chi^2_{(3)} = 40.31, p < .001$. Post-hoc analyses revealed a significant difference between *Control* and *Gain Framing*, $p_{adj} < .001$, suggesting a framing effect. However, there were no differences among *Gain Framing*, *DP*, and *LDP* conditions.

The two-way interaction of question sensitivity $\times$ condition was significant, $\chi^2_{(3)} = 25.19, p < .001$. Post-hoc comparisons showed that participants in the *Gain Framing* and *Control* conditions selected more local options for the high-sensitive questions than for the low-sensitive questions, $p_{adjs} < .001$, but such pattern was not evident in the *DP* or the *LDP* condition. Thus, participants preferred high-sensitive personal information to be used by the app locally, but such preference disappeared when they were informed of DP or LDP.

**Both selection rate.** Similar to the *Local only* option results, main effects of question sensitivity, $\chi^2_{(1)} = 325.65, p < .001$, condition, $\chi^2_{(3)} = 30.31, p < .001$, as well as their interaction, $\chi^2_{(3)} = 22.45, p < .001$, were all significant. Specifically, the selection rate for the high-sensitive questions (37.9%) was smaller than that for the low-sensitive questions (61.1%). For the average selection rate of each condition (*Control*: 44.2%; *Gain Framing*: 50.2%; *LDP*: 49.7%; *DP*: 54.0%), pairwise comparisons were all significantly different, $p_{adjs} \leq .045$, except for *Gain Framing* vs. *LDP* and *Gain Framing* vs. *DP*.

Although the effect of question sensitivity was significant for all conditions, $p_{adjs} < .001$, results of the high- and low-sensitive questions showed different patterns across conditions. For the low-sensitive questions, the selection rate of the *Gain Framing* condition was larger than that of *Control* and *LDP*, $p_{adjs} \leq .013$, but not *DP*. For the high-sensitive questions, the selection rates for the *DP* and *LDP* conditions were higher than that of *Gain Framing*, $p_{adjs} \leq .048$, indicating the effect of differential privacy communication.

*2) RQ2. DP vs. LDP:* Participants showed more willingness to share their information with the *DP* description than for the *LDP* description.

**Opt out rate.** Results of *LDP* (13.6%) and *DP* (14.3%) conditions showed no significant difference. Similar results were obtained for both the low- and high-sensitive questions.

**Local only selection rate.** Post-hoc analyses revealed that participants' selection rate of the *LDP* condition (38.5%) was larger than that of the *DP* condition (32.4%), $p_{adj} < .001$. Also, such difference was evident for questions of high sensitivity, $p_{adj} = .024$, and of low sensitivity, $p_{adj} = .047$.

**Both selection rate.** Post-hoc pairwise comparisons revealed that the average selection rate of *LDP* (49.7%) was smaller than that of *DP* (54.0%), $p_{adj} = .045$. Nevertheless, the selection rate of the *LDP* condition showed no significant difference from that of the *DP* condition either for the low-sensitive or high-sensitive questions.

*3) Correct Rate of Check Questions:* The correct rate of the *DP* condition was higher than that of the *LDP* condition. Better results were evident for the second check question than for the first check question regardless of conditions.

Correct answers for check questions collapsed across participants were entered into a 2 (check: *first*, *second*) $\times$ 2 (condition: *DP*, *LDP*) chi-squared tests. The correct rate for the *DP* condition (71.6%) was higher than that of the *LDP* condition (61.8%), $\chi^2_{(1)} = 7.60, p = .006$, suggesting that the concept of DP was easier to recognize than that of LDP. The correct rate of the *second* question (74.7%) was higher than that of the *first* one (58.7%), $\chi^2_{(1)} = 20.53, p < .001$, indicating the effect of an extra presentation. The interaction of check $\times$ condition was not significant. Thus the effect of extra presentation played a similar role between the two conditions.

### D. Discussion

Consistent with the results of Pilot Study 1, participants showed more privacy concerns for the high-sensitive questions than for the low-sensitive questions. The gain framing showed little effect on participants' opt-out decisions, but it increased participants' data disclosure compared to the *Control*. Moreover, such increase was only evident for the low-

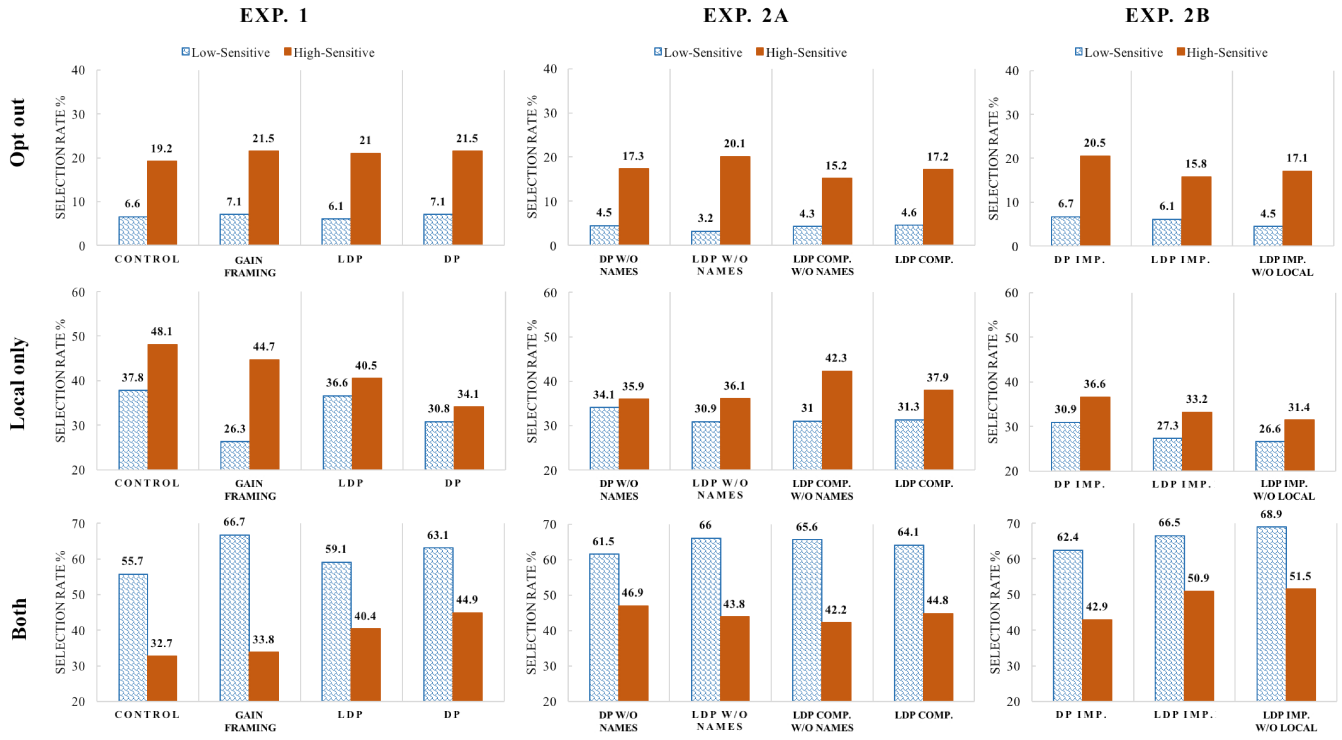


Fig. 3: Selection rates of *Opt-out* decisions (top row), *Local only* (middle row), and *Both* (bottom row) options for the low-sensitive and the high-sensitive questions in different conditions of Experiments 1, 2A, and 2B.

sensitive questions, indicating a risk aversion dependent on the question's sensitivity.

When participants were further informed of DP or LDP protection, their overall data sharing did not increase. Nevertheless, they increased their data disclosure for the high-sensitive questions, suggesting a positive effect of communicating DP and LDP. Moreover, the overall data disclosure rate of the *DP* condition was larger than that of the *LDP* condition. Together with the better results of check questions, those results suggest that DP seems to be easier for participants to understand, and thus resulted in more data sharing.

We conjecture that the better results for the *DP* condition may be due to the specific descriptions that we presented. In particular, data perturbation before data collection providing stronger privacy guarantee for LDP was not clearly described. Also, we included organization names when introducing DP and LDP techniques. Participants' trust of DP and LDP may depend on their trust of the associated organizations.

## V. Experiments 2A & 2B

To examine factors that affect DP and LDP descriptions in users' data disclosure decisions, we conducted two sub-experiments. In Experiment 2A, we emphasized the data perturbation processes of LDP and examined whether participants would understand the better privacy protection and thus increase their data disclosure. We also removed the company names associated with DP and LDP to understand their influence on participants' data disclosure decisions. Considering the difficulty for laypersons to relate data perturbation with privacy protection, we examined the effect of communicating the implications of DP and LDP techniques in Experiment 2B.

### A. Experiment 2A

*1) Participants, Stimuli, and Procedure:* Another 781 Amazon MTurk workers were recruited. To understand the effect of company name, we proposed *DP w/o Names* and *LDP w/o Names* conditions, which were the same as the *DP* and *LDP* conditions of Experiment 1 except company names associated with DP and LDP were removed. To improve users' comprehension of the better privacy protection provided by LDP, we included a *LDP Comp.* condition, in which the description differed from that of Experiment 1 by emphasizing that data perturbation (noise) was added before sending users' responses to the server. To further understand the combined influence of the company names and the new description, we included a *LDP Comp. w/o Names* condition, which was the same as *LDP Comp.* but the company names were removed. The detailed descriptions for all conditions are listed in Table XII. The procedure was identical to that of Experiment 1.

*2) Results:* We excluded participants from data analysis using the same criterion as Experiment 1 (see details in Table XI). Consequently, 125 participants from the *DP w/o Names*, 149 from the *LDP w/o Names*, 161 from the *LDP Comp. w/o Names*, and 146 from *LDP Comp.* were included in the data analyses. Results of each option are shown in the second column of Fig. 3, and were entered into chi-squared tests similar to the prior experiment. Post-hoc comparisons were also performed in a similar way (see results in Table VIII).

*a) RQ2. Effects of Company Names and Emphasizing Data Perturbation:* Without company names, participants' data sharing decisions became similar between DP and LDP descriptions. In the local setting, question sensitivity became significant for the *LDP w/o Names* description but not for the *DP w/o Names* description.

Opt-out rate of the high-sensitive questions was less for the *LDP Comp. w/o Names* condition than that for the *LDP w/o Names* condition. However, participants in the *LDP Comp. w/o Names* condition preferred to share more high-sensitive information in the local setting than participants in the *LDP w/o Names* and the *DP w/o Names* conditions.

**Opt out rate.** Like the prior experiment, participants showed more willing to opt out for the high-sensitive questions (17.5%) than for the low-sensitive questions (4.1%), $\chi^2_{(1)} = 372.68, p < .001$. The two-way interaction of question sensitivity $\times$ condition was at .05 significance level, $\chi^2_{(3)} = 7.88, p = .048$. No difference was evident among all conditions for the low-sensitive questions. For the high-sensitive questions, the selection rate of the *LDP w/o Names* condition (20.1%) was larger than that of the *LDP Comp. w/o Names* condition (15.3%), $p_{adj} = .020$.

**Local only selection rate.** Participants selected more *Local only* option for the high-sensitive (38.1%) than for the low-sensitive questions (31.6%), $\chi^2_{(1)} = 41.54, p < .001$. The interaction of question sensitivity $\times$ condition was significant, $\chi^2_{(3)} = 12.74, p = .005$. The effect of sensitivity was significant for all conditions, $p_{adjs} \leq .019$, except for the *DP w/o Names* condition. Thus, when company names were not mentioned, same results as Experiment 1 were evident for DP but not LDP, suggesting the effect of company names in the *LDP* description. Moreover, post-hoc analysis revealed that selection results across conditions showed no difference for the low-sensitive questions, but for the high-sensitive questions selection rate of the *LDP Comp. w/o Names* was greater than that of *LDP w/o Names* and *DP w/o Names*, $p_{adjs} \leq .017$.

**Both selection rate.** Similar to the other two options, the main effect of question sensitivity, $\chi^2_{(1)} = 331.56, p < .001$, and its interaction with condition, $\chi^2_{(3)} = 9.06, p = .029$, were significant. Participants selected more *Both* option for the low-sensitive questions (64.3%) than for the high-sensitive questions (44.4%). Post-hoc analysis showed that the effect of sensitivity was significant for all conditions, $p_{adjs} < .001$. Although pairwise comparisons showed no significant differences across conditions for both sensitive levels, different patterns were revealed: For the low-sensitive questions, the selection rate of the *DP w/o Names* condition was numerically smallest. In contrast, for the high-sensitive questions, its selection rate was numerically largest (see Fig. 3 second column).

*b) Correct rate of check questions.:* Better correct rates were obtained for the two *LDP Comp.* conditions. Same as the prior experiment, participants' correct answer rate for the second check question was better than that for the first check question regardless of conditions.

Check questions were analyzed in a similar way as Experiment 1. Overall, the correct rates differed across conditions, $\chi^2_{(3)} = 24.75, p < .001$. Post-hoc comparisons revealed that the difference was mainly due to better results of the *LDP Comp. w/o Names* (77.5%) and *LDP Comp.* (73.9%) conditions than that of the *DP w/o Names* condition (61.7%), $p_{adjs} > .004$, indicating the effect of emphasizing the data perturbation processes. Same as Experiment 1, the correct rate of the *second* question (77.8%) was higher than that of the *first* question (63.5%), $\chi^2_{(1)} = 36.25, p < .001$.

*3) Discussion:* With an emphasis on data perturbation processes, the opt-out rate of the high-sensitive questions was smaller for the *LDP Comp. w/o Names* condition than for the *LDP w/o Names* condition. The check question results of the two *LDP Comp.* conditions were also better than that of the *DP w/o Names* condition. Instead of increasing their data sharing, participants in the *LDP Comp. w/o Names* condition selected more *Local only* option for the high-sensitive questions than participants in the *LDP w/o Names* and *DP w/o Names* conditions. Thus, an emphasis on data perturbation processes helped participants recognize the strong privacy premise of LDP, but it seemed to make them misbelieve that such protection is local and thus showed more willingness to share the high-sensitive information locally. The effect of "local" word was also implied by more selection of *Local only* option in the *LDP* condition than in the *DP* condition in Experiment 1.

After removing the company names, the data disclosure differences between the *DP* and *LDP* conditions in Experiment 1 were not evident. For the *Local only* option, the non-significant effect of question sensitivity for both *DP* and *LDP* conditions in Experiment 1 became significant for the *LDP w/o Names* condition only. Altogether, those results suggest the company names contributed to the differences obtained between the *DP* and *LDP* conditions in Experiment 1, and the company names associated with LDP seemed to have more impact than company names associated with DP.

### B. Experiment 2B

As suggested by the results of Experiment 2A, participants had difficulty in understanding what do data perturbation processes mean for privacy protection. Thus, we conducted Experiment 2B examining whether communicating the privacy implication [31] of data perturbation will help participants understand the stronger privacy promise of LDP.

*1) Participants, Stimuli, and Procedure:* Extra 600 Amazon MTurk workers were recruited. Materials and procedures of Experiments 2B were identical to Experiment 2A except *DP Imp.* and *LDP Imp.* descriptions (i.e., whether the privacy protection provided by DP or LDP relies on the trustworthiness of the company or the server) were implemented. We also included one *LDP Imp. w/o Local* condition to examine any impact of the word "local" on participants' data sharing decisions. The detailed descriptions are listed in Table XII.

*2) Results:* After excluding participants using the same criteria as prior experiments (see details in Table XI), there were 162 participants in the *DP Imp.* condition, 149 in the *LDP*

*Imp.* condition, and 157 in the *LDP Imp. w/o Local* condition. Selections of each option collapsed across participants (see Fig. 3 last column) were analyzed similarly as prior experiments (see Table IX for the full results of statistical tests).

*a) RQ2. Effects of Implication Descriptions and Word "Local":* More data sharing and fewer opt out were evident for the two *LDP Imp.* conditions than for the *DP Imp.* condition regardless of questions' sensitivity. With the implication descriptions, no significant difference was evident between *LDP Imp.* and *LDP Imp. w/o Local*.

**Opt out rate.** Participants opted out more for the high-sensitive questions (17.8%) than for the low-sensitive questions (5.8%), $\chi^2_{(1)} = 229.49, p < .001$. Opt-out rates also differed across conditions, $\chi^2_{(2)} = 10.77, p = .005$, with the obtained result of the *DP Imp.* condition (13.6%) being larger than those of the *LDP Imp.* (10.9%) and the *LDP Imp. w/o Local* (10.8%) conditions, $p_{adjs} \leq .026$. The two-way interaction of sensitivity $\times$ condition was not significant. Thus, across experiments, for the first time, we obtained a smaller opt-out rate for LDP than for DP regardless of question sensitivity, suggesting the effect of implication communication.

**Local only selection rate.** Same as the opt-out decisions, only the two main effects were significant. Participants selected more *Local only* options for the high-sensitive questions (33.7%) than for the low-sensitive questions (28.3%), $\chi^2_{(1)} = 22.58, p < .001$. Selection rates varied across conditions, $\chi^2_{(2)} = 12.36, p = .002$. Post-hoc pairwise comparisons revealed that the selection rate of *DP Imp.* (33.7%) was larger than those of *LDP Imp.* (30.3%), $p_{adj} = .049$, and *LDP Imp. w/o Local* (29%), $p_{adj} = .002$, respectively.

**Both selection rate.** Participants decided to share more low-sensitive information (65.9%) than high-sensitive information (48.4%), $\chi^2_{(1)} = 206.02, p < .001$. The main effect of condition was significant, $\chi^2_{(2)} = 29.2, p < .001$. Post-hoc comparisons revealed that the selection rate of *DP Imp.* (52.6%) was less than those of *LDP Imp.* (58.7%), $p_{adj} = .037$, and *LDP Imp. w/o Local* (60.2%), $p_{adj} < .001$, respectively.

*b) Correct rate of check questions:* A larger correct rate was obtained for the *DP Imp.* condition than for the two *LDP Imp.* conditions. Same as the prior experiments, participants' correct answer rate for the second check question was better than for the first check question regardless of condition.

Check-questions' results were analyzed similarly as Experiment 2A. The correct rates differed across conditions, $\chi^2_{(2)} = 13.41, p = .001$. Post-hoc comparisons revealed that the difference was mainly due to better results of *DP Imp.* (81.6%) than those of *LDP Imp.* (71.6%) and *LDP Imp. w/o Local* (71.6%), $p_{adjs} \leq .004$. However, the two *LDP Imp.* conditions did not differ. The correct rate of the *second* check question (81.3%) was higher than the *first* one (68.6%), $\chi^2_{(1)} = 23.95, p < .001$.

*3) Discussion:* In Experiment 2B, we made it clear to participants that DP but not LDP relies on the trustworthiness of the company or the server for privacy protection. When such implications were communicated, more data sharing and few opt out were obtained for the two *LDP Imp.* conditions than for the *DP Imp.* condition regardless of question sensitivity. However, the results of check questions revealed that participants did better for answering the DP concept. Altogether, those results indicate that participants understood better protection provided by LDP, but the concept of DP might still be easier for them to recognize. We did not obtain any difference between the two *LDP Imp.* conditions, indicating little impact of the word "local" with implication communication.

### C. Summary of Experiments 1 and 2

Using the health app data collection setting and hypothetical willingness to disclose sensitive personal information as the testbed, we evaluated participants' data disclosure rates as a function of differential privacy description. When definitions of DP and LDP were communicated (Experiment 1), participants increased their data disclosure for high-sensitive information, suggesting a positive effect of communicating differential privacy to laypeople. However, the overall data sharing was better for DP than for LDP, though the latter provides better privacy guarantee.

When we emphasized the data perturbation processes of LDP (Experiment 2A), participants showed more willingness to share high-sensitive information with the app locally. However, when the implications of DP or LDP (i.e., whether the privacy protection relies on the trustworthiness of the company) was presented (Experiment 2B), participants showed the willingness to opt out less and to share more with LDP than with DP. Altogether, those results suggest laypeople have difficulty in understanding the definitions, especially the perturbation processes of differential privacy. But communication of implications is effective, especially in helping them understand which technique provides better privacy protection.

## VI. Experiment 3

We conducted an open-question survey to understand why participants decided to share or not share their personal information given the differential privacy protection (**RQ3**), and how easy it is for them to understand the descriptions subjectively (**RQ4**).

### A. Participants, Stimuli, and Procedure

We recruited extra 280 Amazon MTurk workers. Besides the differential privacy descriptions from prior experiments, we investigated the descriptions published by companies and organizations that implemented DP or LDP, including Apple, Google, Microsoft, Uber, and US Census Bureau. We made minor changes to those descriptions to make them fit into our study (see descriptions in Table XII from Appendix D).

At the beginning of the survey, participants were instructed that the study is (1) to evaluate their data disclosure decision given one privacy protection technique, and (2) to understand why they decide to do so. Then, we described the three steps of role play as prior experiments. Participants were randomly assigned to one of the descriptions. After viewing the description, participants made one data disclosure decision for high-sensitive information only (see details in Appendix A).

Then they answered an open question about the reason for their choice. We also asked them to indicate their agreement on whether the description was easy to comprehend on a 7-point Likert scale (1: strongly disagree, 7: strongly agree). For participants who gave a rating less than 4, we asked them to highlight the words or sentences that they thought were difficult to understand. In the end, participants answered the same demographic questions as prior experiments.

### B. Results

Only duplicate IP address was used for data exclusion due to short survey time (see Table XI). Another participant who failed to complete the study was also excluded. Table I lists the summary of participants' demographics.

*1) Data Disclosure Decision and Difficult-to-Comprehend Rates:* Given descriptions of differential privacy protection, $47.8\%$ of the participants chose to share their high-sensitive information on average (see Table II). Across all conditions, the numerically largest sharing rate was obtained in the *LDP Imp.* condition ($65.2\%$), in agreement with the results obtained in Experiment 2B. On average, $13.3\%$ of the participants gave a rating less than $4$, indicating that they had difficulty in understanding the presented descriptions. For conditions included terms, such as "noise" or "random responses" (e.g., *LDP*, *Google*, *US Census Bureau*), about $18\%$ of the participants rated the descriptions as difficult (see Table II). In contrast, for conditions without those terms and mentioned benefits or implications of the techniques (e.g., *DP Imp.*, *Uber*), around $5\%$ of the participants gave ratings less than $4$. No participants in the *DP Imp.* condition rated the description as hard to understand, in agreement with better check questions results obtained in Experiment 2B.

TABLE II: Difficult-to-comprehension and sharing decision rates for all descriptions in Experiment 3. The number of participants in each condition is listed in the brackets of the first column.

| Condition | Difficult-to-Comprehend Rate | Sharing Decision Rate |
|---|---|---|
| DP Imp. (22) | 0.0% | 36.4% |
| LDP Comp. (22) | 4.5% | 40.9% |
| Microsoft (19) | 5.3% | 52.6% |
| LDP Imp. w/o Local (18) | 5.6% | 44.4% |
| Uber (18) | 5.6% | 55.6% |
| LDP Imp. (23) | 8.7% | 65.2% |
| LDP Comp. w/o Names (18) | 11.1% | 50.0% |
| Apple (21) | 14.3% | 57.1% |
| LDP w/o Names (19) | 15.8% | 52.6% |
| DP (21) | 19.0% | 33.3% |
| Google (19) | 21.1% | 42.1% |
| LDP (17) | 23.5% | 41.2% |
| US Census Bureau (21) | 23.8% | 47.6% |
| DP w/o Names (20) | 30.0% | 50.0% |

*2) Answers to Open Questions:* We analyzed the results of open questions by identifying themes and generating codes using an inductive approach [7]. The first two authors independently coded the answers for open questions based on the data disclosure decisions and easy-to-comprehend measures, and then cleaned up the codes to generate new ones. We then re-coded the results using the new codes and added emerging codes when necessary. Lastly, the research team discussed the codes and grouped them into different themes. We assigned random sequential numbers to participants for the analysis.

*a) Why share?:* One hundred and thirty-three participants choosing to share personal information. Their explanations were grouped into three main themes:

- **Trust of DP and LDP techniques.** About $62\%$ ($82$) of the participants decided to share their information because of or partially because of the described privacy protection. Their replies indicate trust for privacy protection techniques generally, such as "*it sounds like a viable and trustworthy type of protection technique, and I don't feel wary about trusting it.*" (P124). Moreover, $28$ of them demonstrated somewhat understanding of differential privacy in their replies, e.g., "*I think an additional random data set might throw off how certain information can tie to you*" (P46).
- **Utility consideration.** About $26\%$ ($34$) of the participants made the decisions due to or partly due to their data would be useful or beneficial for the app, the service they got, or the other people. Examples include "*I feel that I would be able to get more accurate information if it collects my data...*" (P48), and "*If it helps to provide data to make a more accurate algorithm or helps with someone's research I'm willing to provide it*" (P117). Also, $14$ participants' answers revealed their considerations for both utility and privacy, e.g., "*I am comfortable to share this information for the benefit that will be served to me. The privacy technique sounds like it will keep all users equally obfuscated*" (P119).
- **Little privacy concern for asked or any information, learned helpless, and no fear of loss.** About $22\%$ ($29$) participants explained that they made the decisions because (1) they did not care about the privacy for the asked information or any information, e.g., "*...personally I don't currently have a significant history of medical problems, substance abuse, family history, etc...*" (P113); (2) a lot of their personal information was already out, e.g., "*we share a lot of info already on social media*" (P10); or (3) there was nothing to hide or protect, e.g., "*I don't think its really that big a deal, i tell everyone my business lol*" (P39).

*b) Why not share?:* We also performed a similar analysis to understand why $142$ participants chose not to share, the results of which were grouped into the following four themes:

- **Too sensitive to share.** About $37\%$ ($53$) of the participants decided not to share because of the sensitivity level of requested information, e.g., "*Because it's personal I have a long list of medical conditions, and my family wouldn't want me to share their personal information as well*" (P144), and "*Even if the privacy policy is equivalent to an opt out which might be fine for other circumstances but when you are talking about your health records it just wouldn't be worth the risk.*" (P224).
- **Distrust differential privacy techniques.** About $33\%$ ($47$) of the participants explained that they distrust the described differential privacy techniques. Besides the general concern of privacy techniques, e.g., "*This technique does not sound*

*like it is full proof. I think a hacker could get through this system real easy.*" (P154), participants distrusted differential privacy techniques because (1) the descriptions were vague, e.g., "*I did not understand the explanation of how the protection worked*" (P200); (2) the techniques were new, e.g., "*Its not tested enough too new.*" (P167); and (3) further verification is needed, e.g., "*I'm still not fully convinced that it is trustworthy. I'd need to know more about the processes involved and how thoroughly tested it has been*" (P191).

- **Risks of data leak, breach, or hack.** About 30% (43) of the participants worried about future risk of leak, breach, or hack of their data, and thus chose not to share, e.g., "*...All of the data breaches that have occurred in recent years already prove that no privacy protection technique is 100%. I would like to limit the chances of my info being leaked as much as possible.*" (P271).
- **Distrust the app or tech companies.** Another 19% (27) participants explained that they chose not to share because they distrust the app or the tech companies, e.g., "*How good the privacy protection technique is one thing, and whether they will sell my private data to other parties is another thing. I never fully trust such kind of app/technique.*" (P227).

*c) Which part(s) is hard to understand?:* Thirty-seven participants indicated that the descriptions were hard to understand. They highlighted mostly the words "*noise*" (19 times) and "*random(ly)*"(14 times), both of which are related to the perturbation processes. When answering why they thought the description was unintelligible, about half (17) of the participants' replies mentioned random noise, including "*How will the random noise protect my information?...*" (P90) and "*what is random noise?*" (P162). Nine participants also indicated that the descriptions were jargony or had technical terms.

### C. Discussion

Experiment 3 employed an open-question survey to understand factors impacting people's data sharing decisions for high-sensitive information when given descriptions of DP or LDP. We found that participants decided to share primarily because of the descriptions of differential privacy techniques and somewhat utility consideration. Participants who chose not to share had various concerns: about $1/3$ of them believed that the requested information was too sensitive to share, another $1/3$ distrusted the described differential privacy techniques, and an extra $1/3$ worried about negative consequences of sharing high-sensitive information in the future.

We also evaluated participants' subjective comprehension of descriptions to uncover the difficult parts to understand. Less than 15% of the participants rated those descriptions as hard to comprehend. They highlighted parts closely related to data perturbation processes as most difficult to understand, consistent with results obtained in Experiments 2A and 2B.

## VII. Experiment 4

While participants' self-reported comprehension was good for the descriptions, it is unclear whether the same holds for their comprehension performance. To understand participants' *objective* comprehension of DP and LDP (**RQ5**), we proposed questions evaluating their understanding of privacy and utility impacts, such as utility cost and who can see their data.

Based on the findings from prior experiments, we also proposed two new descriptions, *DP Flow* and *LDP Flow*. In each new description, we described the data flow affording privacy and utility implication inferences while simplifying the technical details such as noise and perturbation. We examined the effect of two new descriptions in improving people's understanding of DP and LDP by comparing them with descriptions from prior experiments.

To make sure the two new descriptions and objective comprehension questions are understandable, we conducted Pilot Study 2 with 20 participants. The procedure was identical to Experiment 4 except that after answering each question, participants also indicated their agreement on whether the presented question was easy to comprehend on the 7-point Likert Scale. For any question with a rating lower than 4, we asked them to describe which part or parts of the question are hard to understand and briefly explain why. Participants gave overall high ratings for both descriptions and all questions except Q4 (see details in Appendix C). Based on participants' replies, we improved the question and both descriptions.

### A. Participants, Stimuli, and Procedure

We recruited another 599 Amazon MTurk workers. Descriptions of *DP Flow* and *LDP Flow* are given in Table XII. For each description, we introduced DP or LDP data flow by listing different parties, such as external third-party companies, and explaining how those parties are involved in the data flow. We also added the description of accuracy loss, and removed technical or jargon terms included in the former proposed descriptions. Note, in the two new descriptions, the risk of data compromise for DP and LDP was not described explicitly. Descriptions of *DP Imp.*, *LDP Imp.*, *DP w/o Names*, and *LDP w/o Names* were also included for comparisons with the new descriptions.

We evaluated participants' comprehension of DP and LDP with five questions. Three of them were about privacy inferences from the perspectives of attackers (Q1), internal employees (Q2), and third-party companies (Q3). Another two were about utility inferences from the perspectives of the app company (Q4) and third-party companies (Q5) (See Appendix C).

Participants were randomly assigned to one of the six descriptions. At the beginning, we informed participants that the study was to evaluate their comprehension of one privacy protection technique based on the given description (see Appendix A). Then, we described the three steps of role play as prior experiments. After viewing one description of DP or LDP, participants answered the five questions, which were presented in a randomized order. We also randomized the options except "Unsure" and "Prefer not to answer" for each question. When answering each question, participants could see the description by placing their cursor over the text of "Hover here to see the description". Participants also indicated

whether the description of the privacy protection technique was easy to comprehend on the 7-point Likert scale. They answered questions about their demographics in the end.

### B. Results

The number of participants in each condition is listed in the first row of Table III. Participants' demographic showed a similar pattern as prior experiments (see Table I). Correct answer rate of each question for each description collapsed across participants (see Table III) were entered into 3 (description: *w/o Names*, *Imp.*, *Flow*) $\times$ 2 (technique: *DP*, *LDP*) chi-squared tests. Post-hoc comparisons were also performed as prior experiments. Participants' average easy-to-comprehend rating for each description were analyzed with analysis of variance (ANOVA) using the same two factors as chi-squared tests. Post-hoc tests were also performed with corrected $p$-values for possible inflation.

TABLE III: Correct answer rate for each question and average comprehension rating of each description in Experiment 4. The number of participants in each condition is listed in the brackets on the top row. The number in the brackets on the last row indicates the standard error of each average rating.

| Question | w/o Names | | Imp. | | Flow | |
|---|---|---|---|---|---|---|
| | DP (88) | LDP (90) | DP (86) | LDP (95) | DP (86) | LDP (95) |
| Privacy_Attackers | 19.3% | 25.6% | 87.2% | 76.8% | 51.1% | 77.9% |
| Privacy_Employees | 28.4% | 26.7% | 40.7% | 40.0% | 47.7% | 75.8% |
| Privacy_Third Party | 52.3% | 32.2% | 52.3% | 50.5% | 59.3% | 75.8% |
| Utility_Cost | 27.3% | 22.2% | 14.0% | 20.0% | 48.8% | 54.7% |
| Utility_Third Party | 55.7% | 60.0% | 81.4% | 56.8% | 89.5% | 84.2% |
| Easy-to-Comprehend Rating | 4.52 (1.71) | 3.53 (1.69) | 4.99 (1.37) | 4.57 (1.53) | 4.52 (1.59) | 5.29 (1.25) |

**RQ5: Effect of Description.** The main effect of description was significant for all comprehension questions, $\chi^2_s \geq 24.77, p_s < .001$, (see Table X for the statistical details). Across five questions, the best results were obtained for the *Flow* descriptions except Q1, privacy inference of attackers, in which the highest correct rates were evident for the *Imp.* descriptions. Compared to the *Imp.* descriptions, risk of data compromise is implicit in the *Flow* descriptions, suggesting the effect of explicitness in helping comprehension.

Correct answer rates of all privacy-related questions for the *Imp.* descriptions were larger than those for the *w/o Names* descriptions. Nevertheless, the correct rates of utility-related questions showed no significant difference between those two types of descriptions. Thus, an implication description of data breach is helpful for participants to understand the privacy protection of differential privacy.

**RQ5: Effects of Technique and Technique $\times$ Description.** The main effect of technique showed different patterns among the comprehension questions. For privacy-related questions, the overall correct rates were higher for the LDP conditions than for the DP conditions, $\chi^2_s \geq 4.09, p_s < .043$, except for Q3, privacy inference about third-party companies, which showed no significant difference. Moreover, the two-way interaction of description $\times$ technique was significant for all three privacy-related questions, $\chi^2_s \geq 10.77, p_s < .005$. Generally, the difference between DP and LDP was evident for the *w/o Names* and *Flow* descriptions, but not the *Imp.* descriptions.

For utility-related questions, better results were evident for DP than for LDP on Q5, utility inference of third-party companies, $\chi^2_{(1)} = 4.07, p = .044$. Also, such pattern was only evident with the *Imp.* descriptions, $p_{adj} = .002$, indicating the importance of utility description for LDP.

For **average easy-to-comprehend ratings**, the main effect of description was significant, $F_{(2,534)} = 27.73, p < .001, \eta^2_p = .075$. Post-hoc pairwise comparisons revealed that the average ratings of the *Flow* descriptions (4.90) were similar to those of the *Imp.* descriptions (4.78), both of which were higher than those of the *w/o Names* (3.94), $p_{adjs} < .001$. Although the main effect of technique was not significant (DP vs. LDP: 4.62 vs. 4.47), its interaction with description were significant, $F_{(2,534)} = 13.14, p < .001, \eta^2_p = .047$. Critically, with the *Flow* descriptions, participants' overall rating for LDP was higher than that for DP (5.29 vs. 4.52), whereas an opposite pattern was evident for both the *Imp.* (4.57 vs. 4.99) and the *w/o Names* (3.53 vs. 4.35) descriptions.

### C. Discussion

Contrary to the self-reported results in Experiment 3, objective comprehension results revealed that participants had difficulty in understanding the implications of DP and LDP, especially with descriptions focusing on definition (i.e., the *w/o Names* descriptions). Explicit descriptions about the trustworthiness of the company (i.e., the *Imp.* descriptions) improved the correct answer rates for privacy inference questions, especially the inference about attackers. Participants' correct answer rates for all questions were improved with the *Flow* descriptions except the inference about attackers. That is probably because the privacy inference of attackers became somewhat implicit in the *Flow* descriptions compared to the *Imp.* descriptions. Altogether, those results indicate the effects of explicitness and descriptions affording implication inferences in helping laypeople understand differential privacy.

### D. Summary of Experiments 3 and 4

Using a similar setting as prior experiments, we asked participants to explain why they decided to share or not share their personal information given the descriptions of differential privacy. Participants chose to share data mainly because of the described privacy protection, but those who chose not to disclose their personal information revealed different concerns, including the requested information was too sensitive to share, they distrusted the described privacy technique, and they worried about the risk of data breach. Less than 15% of participants rated the descriptions as hard to comprehend, and they mainly highlighted the parts related to data perturbation processes as difficult to understand. Experiment 4 was conducted to understand how participants comprehend DP and LDP objectively. Based on the findings from prior experiments, we proposed the *Flow* descriptions which afford privacy and utility implication inferences. Compared to the *Imp.* and the *w/o Names* descriptions, we obtained better

comprehension results for the *Flow* descriptions generally. However, best privacy inference results were obtained when the privacy implications were described explicitly. Overall, these results revealed the complexity of people's data disclosure decision-making, and the importance of implication communication to help people understand DP and LDP.

## VIII. General Discussion

The present study reports four experiments that were motivated by communicating differential privacy to facilitate users' data disclosure decisions. In Experiments 1 and 2, we proposed different ways of describing differential privacy techniques and evaluated the effects of those descriptions on participants' data sharing decisions (**RQ1**-**RQ2**). In Experiments 3 and 4, reasons behind participants' data sharing decisions (**RQ3**), as well as their subjective (**RQ4**) and objective (**RQ5**) comprehensions of DP and LDP were examined, respectively.

### A. Summary of Main Results

**Difficult to Understand the Data Perturbation Processes.** When we presented DP and LDP techniques based on definition, participants increased their data disclosure of high-sensitive information. Participants' data disclosure rates were larger for DP than for LDP, despite the latter providing better privacy guarantees. Moreover, participants reduced their data sharing when the "local" aspect of LDP was emphasized. Many participants explained that they made the sharing decision because of the described techniques, but their answers of objective comprehension questions indicated that they had difficulty in understanding the privacy and utility implications and might not differentiate the benefit of differential privacy from the promise of any other privacy technology.

**Effects of Descriptions Affording Implication Inferences.** When privacy implications (i.e., whether the privacy protection relies on the trustworthiness of the company) were presented, participants opted out less and shared more with LDP relative to DP. Together with the highest correct answer rate of privacy inferences obtained with implication descriptions, it indicates that participants' data disclosure decisions were closely related to their correct understanding of privacy protection. When privacy and utility inferences were embedded within data flow descriptions, participants increased their correct answer rates for objective comprehension questions, indicating that descriptions affording implication inferences facilitate participants' comprehension of differential privacy.

**Primary Concern for Information Sensitivity.** Our results also revealed that information sensitivity is an important moderator for people's data disclosure decisions. On average, participants' data disclosure rates of the high-sensitive questions were 20% less than the low-sensitive questions in the first two experiments. Participants, especially those who have medical conditions, revealed that they worried about the negative consequences of data leakage or misuse of medical-related information, thus chose not to share.

### B. Data Disclosure Decision-Making

The effect of information sensitivity on data sharing decisions implies distribution differences between the collected low-sensitive and high-sensitive information. Such differences are informative to differential privacy algorithms or deployments in which such effect has not been accounted.

Besides information sensitivity, we also examined the framing effect to understand people's data disclosure decisions. With a loss framing presented in Pilot Study 1, participants opted out less for the low-sensitive questions. When a gain framing was presented in Experiments 1 and 2, participants only increased their data disclosure for the low-sensitive questions. Thus, our results indicate that the bounded rationality of privacy decisions [2], [25] was qualified by the sensitivity of information, providing an explanation for the differences obtained in prior studies [3], [21].

Extra factors impacting data sharing decisions were revealed in the qualitative results of Experiment 3, which we grouped into two categories: context-dependent and trustworthiness-related. When making data disclosure decisions, participants took *personal contexts* into consideration, e.g., whether they have medical conditions, have been hacked before, or heard about reports of user data breach. Some participants' decisions also factored in the *trustworthiness of the technique or the company*, e.g., whether they believe differential privacy is tested enough or the company's intention to collect users' data.

Thus, privacy-related decisions are multi-faceted [40]. Towards effectively communicating differential privacy to facilitate users' data disclosure, both general factors and users' specific concerns should be understood and addressed.

### C. Differential Privacy Communication

Our results revealed implication descriptions as one effective way to communicate differential privacy to laypeople. However, compared to the privacy aspect, participants still had difficulty in understanding the utility costs of DP and LDP. We conjecture that this is mainly because the privacy implications correspond to people's privacy concerns (e.g., data breach from attackers). Thus, the implication descriptions meet their expectation of privacy protection techniques [31]. Since people consider utility when making data sharing decisions, future work should examine other formats, e.g., graphs, which can intuitively illustrate accuracy loss in helping people understand utility cost. Also, to understand possible trustworthiness gap between DP and LDP, future work could include evaluation results or third-party reports about DP and LDP.

With definition descriptions but not implication descriptions, participants' data disclosure decisions were impacted by specific wording, e.g., "local", suggesting it may be as a result of comprehension. That participants who did not understand privacy implications were susceptible to extraneous factors (e.g., company names) and considered those factors when making data sharing decisions.

### D. Limitations

We note that proper caution should be taken to generalize our finding to other settings. First, we examined the effect of descriptions in data sharing decisions mainly based on the stronger privacy promise of LDP than DP. Other factors, such as utility cost, might render DP and LDP not strict alternatives during preference decisions. However, comprehension of utility cost showed no significant difference between DP and LDP across *w/o Names*, *Imp.*, and *Flow* descriptions. Thus, any impact of utility cost should have a similar effect on the obtained results.

Second, instead of answering the survey questions directly, participants indicated their willingness of data sharing in a hypothetical setting, limiting the ecological validity of the current experimental design. Note that we decided to use this role-play method to protect participants' privacy. Prior studies showed that people's stated intentions and actual behavior sometimes differ [37]. Yet, we replicated the well-known framing effects using the health app setting and the hypothetical questions. Thus, we are confident about our findings. Also, we were mainly interested in the comparison between conditions, so any effect of the role-play can cancel out.

Third, we recruited MTurk workers who tended to be younger, better educated, and put more value on information privacy [26]. Thus, our results may represent population having more privacy concerns than the broader U.S. public.

## IX. Conclusions

Differential privacy techniques are currently being transitioned from academic to industry. Across different approaches of textual descriptions, our study shows that descriptions affording privacy and utility implications can facilitate people's data sharing decisions and their comprehension of DP and LDP techniques. We also found that people's data sharing decisions are multi-faceted and impacted by various factors. Thus, our work highlights the importance of the user-centered deployment of differential privacy but also sheds light on the challenges for usability research and studies.

## X. Acknowledgements

This work was funded in part by the NSF awards #1640374 and #1931443. We would also like to thank Andreas Haeberlen for shepherding this paper, Joesph Calandrino for serving as our point of contact, and other reviewers for their helpful comments which guided us revise and improve the paper.

## References

[1] J. M. Abowd. Protecting the confidentiality of america's statistics: Adopting modern disclosure avoidance methods at the census bureau. https://www.census.gov/newsroom/blogs/research-matters/2018/08/protecting_the_confi.html, 2018.

[2] A. Acquisti and J. Grossklags. Privacy and rationality in individual decision making. IEEE Security & Privacy, 3(1):26–33, 2005.

[3] I. Adjerid, A. Acquisti, L. Brandimarte, and G. Loewenstein. Sleights of privacy: Framing, disclosures, and the limits of transparency. In SOUPS, 2013.

[4] Apple. Apple differential privacy team, learning with privacy at scale, 2017. Available at https://machinelearning.apple.com/docs/learning-with-privacy-at-scale/appledifferentialprivacysystem.pdf.

[5] I. Bilogrevic and M. Ortlieb. If you put all the pieces together...: Attitudes towards data combination and sharing across services and companies. In CHI, pages 5215–5227. ACM, 2016.

[6] J. M. Bland and D. G. Altman. Multiple significance tests: the bonferroni method. BMJ, 310(6973):170, 1995.

[7] V. Braun and V. Clarke. Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2):77–101, 2006.

[8] B. Bullek, S. Garboski, D. J. Mir, and E. M. Peck. Towards understanding differential privacy: When do people trust randomized response technique? In CHI, pages 3833–3837. ACM, 2017.

[9] C. Cadwalladr and E. Graham-Harrison. Revealed: 50 million facebook profiles harvested for cambridge analytica in major data breach. The Guardian, 17:22, 2018.

[10] T.-H. H. Chan, E. Shi, and D. Song. Optimal lower bound for differentially private multi-party aggregation. In ESA, 2012.

[11] Y. Chen and H. Xu. Privacy management in dynamic groups: understanding information privacy in medical practices. In CSCW, 2013.

[12] T. Dehling, F. Gao, S. Schneider, and A. Sunyaev. Exploring the far side of mobile health: information security and privacy of mobile health apps on ios and android. JMIR mHealth and uHealth, 3(1):e8, 2015.

[13] B. Ding, J. Kulkarni, and S. Yekhanin. Collecting telemetry data privately. In NIPS, pages 3571–3580, 2017.

[14] C. Dwork. Differential privacy. In ICALP, pages 1–12, 2006.

[15] C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In TCC, pages 265–284, 2006.

[16] R. M. Entman. Framing us coverage of international news: Contrasts in narratives of the kal and iran air incidents. Journal of Communication, 41(4):6–27, 1991.

[17] Ú. Erlingsson, V. Pihur, and A. Korolova. Rappor: Randomized aggregatable privacy-preserving ordinal response. In CCS, 2014.

[18] A. P. Felt, E. Ha, S. Egelman, A. Haney, E. Chin, and D. Wagner. Android permissions: User attention, comprehension, and behavior. In SOUPS, 2012.

[19] M. Gaboardi, J. Honaker, G. King, J. Murtagh, K. Nissim, J. Ullman, and S. Vadhan. Psi: A private data sharing interface. arXiv:1609.04340.

[20] C. S. Gates, J. Chen, N. Li, and R. W. Proctor. Effective risk communication for android apps. IEEE Transactions on Dependable and Secure Computing, 11(3):252–265, 2014.

[21] J. Gluck, F. Schaub, A. Friedman, H. Habib, N. Sadeh, L. F. Cranor, and Y. Agarwal. How short is too short? implications of length and framing on the effectiveness of privacy notices. In SOUPS, pages 321–340, 2016.

[22] D. J. Hauser and N. Schwarz. Attentive turkers: Mturk participants perform better on online attention checks than do subject pool participants. Behavior Research Methods, 48(1):400–407, 2016.

[23] K. Huckvale, J. T. Prieto, M. Tilney, P.-J. Benghozi, and J. Car. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. BMC Medicine, 13(1):214, 2015.

[24] N. Johnson, J. P. Near, J. M. Hellerstein, and D. Song. Chorus: Differential privacy via query rewriting. arXiv:1809.07750, 2018.

[25] D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. Journal of the Econometric Society, 1979.

[26] R. Kang, S. Brown, L. Dabbish, and S. Kiesler. Privacy attitudes of mechanical turk workers and the us public. In SOUPS, 2014.

[27] P. G. Kelley, J. Bresee, L. F. Cranor, and R. W. Reeder. A nutrition label for privacy. In SOUPS, pages 4–15, 2009.

[28] P. G. Kelley, L. Cesca, J. Bresee, and L. F. Cranor. Standardizing privacy notices: an online study of the nutrition label approach. In CHI, pages 1573–1582. ACM, 2010.

[29] K. Kenthapadi and T. T. Tran. Pripearl: A framework for privacy-preserving analytics and reporting at linkedin. In Proceedings of the 27th ACM International Conference on Information and Knowledge Management, pages 2183–2191. ACM, 2018.

[30] P. Krebs and D. T. Duncan. Health app use among us mobile phone owners: a national survey. JMIR mHealth and uHealth, 3(4):e101, 2015.

[31] J. Lin, S. Amini, J. I. Hong, N. Sadeh, J. Lindqvist, and J. Zhang. Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing, pages 501–510. ACM, 2012.

[32] A. M. McDonald and L. F. Cranor. The cost of reading privacy policies. ISJLP, 4:543, 2008.

[33] W. Melicher, M. Sharif, J. Tan, L. Bauer, M. Christodorescu, and P. G. Leon. (do not) track me sometimes: users' contextual preferences for web tracking. In PETS, pages 135–154. De Gruyter Open, 2016.

[34] E. Mills. Aol sued over web search data release. cnet news, 2006.

[35] P. E. Naeini, S. Bhagavatula, H. Habib, M. Degeling, L. Bauer, L. F. Cranor, and N. Sadeh. Privacy expectations and preferences in an iot world. In SOUPS 2017, pages 399–412, 2017.
[36] H. Nissenbaum. Privacy in context: Technology, policy, and the integrity of social life. Stanford University Press, Stanford, CA, 2009.
[37] P. A. Norberg, D. R. Horne, and D. A. Horne. The privacy paradox: Personal information disclosure intentions versus behaviors. Journal of Consumer Affairs, 41(1):100–126, 2007.
[38] S. Sannon, N. N. Bazarova, and D. Cosley. Privacy lies: Understanding how, when, and why people lie to protect their privacy in multiple online contexts. In CHI, 2018.
[39] F. Schaub, R. Balebako, and L. F. Cranor. Designing effective privacy notices and controls. IEEE Internet Computing, pages 70–77, 2017.
[40] D. J. Solove. Understanding Privacy. Harvard University Press, 2010.
[41] J. Torous, J. Nicholas, M. E. Larsen, J. Firth, and H. Christensen. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evidence-based Mental Health, 21(3):116–119, 2018.
[42] J. Y. Tsai, S. Egelman, L. Cranor, and A. Acquisti. The effect of online privacy information on purchasing behavior: An experimental study. Information Systems Research, 22(2):254–268, 2011.
[43] A. Tversky and D. Kahneman. The framing of decisions and the psychology of choice. Science, 211(4481):453–458, 1981.
[44] A. Tversky and D. Kahneman. Rational choice and the framing of decisions. Journal of Business, 59:S251–S278, 1986.
[45] S. L. Warner. Randomized response: A survey technique for eliminating evasive answer bias. JASA, 60(309):63–69, 1965.

## APPENDIX A
## SURVEY INSTRUMENTS

In this appendix, we list the instructions and questions of all experiments using the *LDP* condition as an example. Some procedures were explained in the square brackets.

### A. Experiments 1, 2A & 2B

In the current information age, everyone faces one question: Will you share your personal information in return for a product, service, or other benefits? [the gain framing]

The purpose of this study is to understand what kind of information you are willing to share with a health app, and how you would like your data to be used.

For this survey, suppose: 1) you just download a health app (Orange Health) and start to use it; 2) to ensure appropriate health suggestions and recommendations, the app asks you to provide some information, for example, your age and gender for accurate recommendation of daily calorie intake; 3) the app server also requests permission to access and collect information to provide you better user experience, for example, the information you shared will be used to train some machine learning algorithms at the server, which will then be used to provide more accurate suggestions for all the users. [the three-step role play]

To respect your personal information privacy and ensure best user experience, the data shared with the app will be collected via the local differential privacy (LDP) technique. LDP protects users' privacy by adding random noise to each response that users give such that the probability that any user's attribute is inferred is similar as he or she is opt-out for the data collection. LDP has been used by companies such as Google and Apple. [the *LDP* description]

- *Please select which of the following description is correct about local differential privacy (LDP).*
  - *A privacy protection technique that adds random noise to the aggregated data (e.g., average age) collected from groups of users, such that the users' privacy can be protected in the same way as they are opt-out for the data collection.*
  - *A privacy protection technique that adds random noise to each response that the users provided such that the users' privacy can be protected in the same way as they are opt-out for the data collection.*
  - *It has not been implemented by any organization or company yet.*
  - *None of the above is correct.*
  - *I prefer not to answer.*

[If participants did not answer the above check question correctly, the *LDP* description was presented again. Then participants answered the 14 data sharing questions.]

- *How would you like your answer to the following question being used?*
  - *Only used by the app locally*
  - *Used by the app locally and the server with LDP*
  - *Neither used by the app locally nor the server with LDP*
  - *I prefer not to answer*

[After the demographic questions, participants answered the check questions again. Then they evaluated their trust of all following items: $A = \{$the health app, the app server, LDP/DP$\}$.]

- *Please indicate your agreement with the following description: I trust $x \in A$ to protect my personal information privacy.*
  - *Answered on a 7-point Likert scale from "Strongly Disagree (1)" to "Strongly Agree (7)"*

### B. Experiment 3

The purpose of this study is to evaluate your willingness to share information given one privacy protection technique and to understand why you decide so. [The three-step role play was presented here.] To respect your personal information privacy and ensure better user experience, the data shared with the app will be collected via a privacy protection technique. Next we will present a description of the privacy protection technique. Please read it carefully. [After viewing one description, participants answered the questions below.]

- *Given the privacy protection provided by the technique, will you share your personal information (e.g., date of birth, family medical record, income level, substance use, surgery record, diagnostic record, current medication) with the app server?*
  - *Yes*
  - *No*
  - *I prefer not to answer*

[Based on participants' answers, they then replied one open question.]

- *Please briefly explain why you did not want/would like to share your personal data, given the described privacy protection technique.* _____
- *Please indicate your agreement with the following description: The prior description of the privacy protection technique was easy for me to understand.*
  - *Answered on a 7-point Likert scale from "Strongly Disagree (1)" to "Strongly Agree (7)"*

[Participants who gave a rating less than 4, answered an extra question.]

- *You indicated that the description of LDP was not easy to understand. Please highlight the words or sentences that are difficult for you to understand.* [The description was presented again with the highlight function available on Qualtrics.]

### C. Experiment 4

The purpose of this study is to evaluate your understanding of a privacy protection technique based on the given description. [The three-step role play was presented here.] To respect your personal information privacy and ensure better user experience, the data shared with the app will be collected via a privacy protection technique. Next we will present the description of the privacy protection technique. Please read it carefully and then answer several questions. [After participants viewed the description, they answered the five questions listed in Appendix C, and gave the easy-to-comprehend rating.]

## APPENDIX B
## PILOT STUDY 1

We validated the health app data collection setting as testbed by examining individuals' data disclosure decision of survey questions as a function of question sensitivity (*high*, *low*) and framing effect (*Control*, *Loss Framing*). The question sensitivity was varied within subjects, and the framing effect was varied between subjects. We predicted that participants' data disclosure for the high-sensitive questions would be less than that for the low-sensitive questions, and participants in the *Loss Framing* condition would have more privacy concerns, thus would reduce data disclosure compared to the *Control*.

### A. Participants and Stimuli

219 Amazon MTurk workers completed the online survey. The procedure for both conditions was the same as that in Experiment 1 except that description about loss due to privacy concern was presented at the beginning of the *Loss Framing* condition. The extra description of the *Loss Framing* condition is as follows:

- *Loss Framing*: *In the current information age, everyone faces one question: Will you protect your personal information in sacrifice of a product, service, or other benefits?*

### B. Results

Participants were excluded from data analysis using the same criteria as main experiments (see Table XI). Results of 103 participants from the *Control* and 101 from the *Loss Framing* were included in the data analyses.

45.6% of the participants were female, and their ages ranged from 18 to over 50 years, with 87.7% between 18 and 44 years. 73.5% were college students or professionals who had associates, bachelors, or higher degrees. 75.9% of the participants claimed that they do not have a degree or work experience in computer science or related fields. The demographic distributions were similar between conditions and showed a similar pattern as the main experiments.

TABLE IV: Each option selection results for high-sensitive and low-sensitive questions in each condition of Pilot study 1.

| Condition | Question Sensitivity | Opt Out | Local Only | Both |
|---|---|---|---|---|
| Control | low-sensitive | 5.1% | 32.2% | 62.7% |
| | high-sensitive | 16.1% | 45.5% | 38.4% |
| Loss Framing | low-sensitive | 2.8% | 37.8% | 59.4% |
| | high-sensitive | 17.0% | 51.3% | 31.7% |

*Opt out* decision, selection for *Local only* option, and choice of *Both* option collapsed across participants (see Table IV) were entered into a 2 (question sensitivity: *low-sensitive*, *high-sensitive*) $\times$ 2 (condition: *Control*, *Loss Framing*) chi-squared tests with a significance level of .05, respectively. Post-hoc tests with Bonferroni corrections [6] were performed, testing all pairwise comparisons with corrected *p*-values for possible inflation. We mainly report the statistics for the significant effects. Please refer to Table VI for all statistical tests results.

**Opt out rate.** Participants opted out more for the high-sensitive questions (16.5%) than for the low-sensitive questions (4.0%), $\chi^2_{(1)} = 120.5, p < .001$. The two-way interaction of sensitivity $\times$ condition was significant, $\chi^2_{(1)} = 4.66, p = .031$. Between two conditions, participants' opt-out rates were similar for the high-sensitive questions (*Control*: 16.1%, *Loss Framing*: 16.9%). However, for the low-sensitive questions, participants in the *Loss Framing* condition (2.8%) opted out less than those in the *Control* condition (5.1%), $p_{adj} = .031$. The results suggest that the loss framing made participants think more of "product, service, or other benefits" but limited to the low-sensitive information.

**Local only selection rate.** Participants selected more *Local only* option for the high-sensitive questions (48.4%) than for the low-sensitive questions (34.9%), $\chi^2_{(1)} = 52.55, p < .001$. The selection rates for the *Loss Framing* condition (44.6%) was higher than that for the *Control* condition (38.8%), $\chi^2_{(1)} = 9.37, p = .002$. Nevertheless, the two-way interaction of question sensitivity $\times$ condition was not significant. Thus, participants generally preferred high-sensitive information to be used by the app locally, and such preference was relatively independent from the framing effect.

**Both selection rate.** The results were in agreement with those of *Local only* selection. Participants chose more *Both* option for the low-sensitive questions (61.0%) than for the high-sensitive questions (35.1%), $\chi^2_{(1)} = 192.02, p < .001$. Relative to the *Control* (50.6%), participants in the *Loss Framing* selected less *Both* option (45.5%), $\chi^2_{(1)} = 6.97, p = .008$. Nevertheless, the main effect of question sensitivity did not interact with condition.

**Trust evaluation.** Trust evaluation collapsed across participants were entered into 2 (data use: *the local app*, *the app server*) $\times$ 2 (condition: *Control*, *Loss Framing*) chi-squared tests. Only the main effect of condition was significant, $\chi^2_{(1)} = 4.71, p = .029$. Participants in the *Control* showed more trust (62.1%) than those in the *Loss Framing* (50.9%).

### C. Discussion

In Pilot Study 1, participants showed less willingness to share high-sensitive information than low-sensitive information. When the loss framing was presented, participants' data disclosure was reduced regardless of questions' sensitivity. Participants opted out less for the low-sensitive questions with the loss framing, indicating a risk seeking qualified in terms of question sensitivity. Thus, we obtained the effect of question sensitivity and the framing effect [3], [5], [21], confirming the health app data collection setting and hypothetical willingness to disclose personal information as one testbed to evaluate privacy decisions.

## APPENDIX C
## PILOT STUDY 2

To make sure the two new descriptions and five comprehension questions are understandable to the participants, we conducted a pilot study with 20 participants on Amazon MTurk. Using a between-subject design, half of the participants were randomly assigned into the *DP Flow* condition and the other half into the *LDP Flow* condition.

### A. Participants, Stimuli, and Procedure

At the beginning of the study, we made it clear to the participants that we were interested in (1) their understanding of a privacy protection technique based on the given description, and (2) whether the description is clear and the survey questions are understandable. Participants were randomly assigned to one of the descriptions. After introducing the three steps of role play similar to prior experiments, we presented the *DP Flow* or the *LDP Flow* description. Following the description, participants in each condition answered five questions evaluating their comprehension of DP or LDP implications from the perspectives of privacy and accuracy. The five questions were:

- *Q1: Suppose that you have answered truthfully the questions presented by the app, and the answers were collected using the privacy protection technique explained earlier. If an attacker gets access to the database of the health app company, will the attacker be able to see your real answer?*
- *Q2: Suppose that you have answered truthfully the questions presented by the app, and the answers were collected using the privacy protection technique explained earlier. For employees within the health app company, will they be able to see your real answer?*
- *Q3: Suppose that you have answered truthfully the questions presented by the app, and the answers were collected using the privacy protection technique explained earlier. For the third party companies with which the health app company shared data, will they be able to see the real answer that you submitted?*
- *Q4: With the modification from the privacy protection technique, the accuracy of summary results obtained by the health app company will become _____ if compared to results without the privacy protection technique (**compared to the true results [without the privacy protection technique]**).*
- *Q5: Suppose that you shared your information, such as your family medical history, with the health app. With the modification from the privacy protection technique, will the results still be useful for the third-party companies with which the health app company share data?*

The options are "Yes", "No", "Unsure" , or "Prefer not to answer" except Q4, whose options are "Better", "Worse (correct answer for DP & LDP)", "No change", "Unsure", or "Prefer not to answer". After each question, participants also rated whether the question description was easy to comprehend on the 7-point Likert scale. For participants who gave ratings smaller than 4, they were then asked to describe which part or parts of the description are hard to understand and briefly explain the reasons. The five questions were presented randomly. We also randomized the options except "Unsure" and "Prefer not to answer" for each question. After the five questions, we also asked participants to indicate their agreement on whether the description of the privacy protection technique was easy to comprehend on the 7-point Likert scale. For ratings lower than 4, we asked participants to describe which part or parts of the questions are hard to understand and briefly explain the reasons. In the end, participants answered questions about their demographics.

### B. Results

55% of the participants were female. 15% of them were less than 24 years old. 70% of them were the ages of 25 to 44. The rest 15% were between the age of 45 to 54. 65% of them have college or higher degrees. 35% of the participants had a high school degree. 65% of them indicated that they did not have a computer science background.

TABLE V: Correct answer rate and easy-to-comprehend rating for each inference question, as well as average rating of each description in Pilot Study 2.

| Question | Correct Rate | | Easy-to-Comprehend Rating | |
|---|---|---|---|---|
| | DP Flow. | LDP Flow | DP Flow | LDP Flow |
| Privacy_Attacker | 50.0% | 80.0% | 6.2 | 5.8 |
| Privacy_Employee | 60.0% | 70.0% | 5.6 | 6.4 |
| Privacy_Third Party | 50.0% | 90.0% | 5.7 | 6.1 |
| Utility_Cost | 0.0% | 10.0% | 5 | 5.2 |
| Utility_Third Party | 90.0% | 80.0% | 6.1 | 6.1 |
| Description | NA | NA | 6.2 | 6.1 |

Correct answer rate of each question for both descriptions are shown in Table V. Generally, participants could understand and answered correctly better than the chance for all questions except for the question of utility cost. Among 20 participants, four of them thought Q4 (Utility Cost) was hard to understand, e.g., "*The initial description of the privacy protection technique doesn't mention anything about comparing results with the original data, so I don't really know the answer to the question*", and "*I'm not sure if there is really enough information to know whether the answers will be better or not. The explanation is very vague.*" Participants believed both descriptions and all five questions were easy to understand. Based on the obtained results, we modified the descriptions (see details in Table XII) and Q4 (see the bold part).

## APPENDIX D
## ADDITIONAL RESULTS

We provide exclusion summary of participants for all experiments in Table XI; DP and LDP descriptions proposed in the present study and descriptions from companies are shown in Tables XII; statistical test results of each Pilot study and experiment in Tables VI, VII, VIII, IX, and X, respectively.

TABLE VI: Statistical test results of Pilot Study 1.

| Term | Local Only | | Both | | Opt Out | |
|---|---|---|---|---|---|---|
| | $\chi^2$ | $p$ | $\chi^2$ | $p$ | $\chi^2$ | $p$ |
| Question Sensitivity | **52.55** | **<.001** | **192.02** | **<.001** | **120.5** | **<.001** |
| Condition | **9.37** | **.002** | **6.97** | **.008** | <1.0 | |
| Question Sensitivity * Condition | <1.0 | | 1.04 | 0.309 | **4.66** | **.031** |
| Low-Sensitive vs. High-Sensitive | | | | | | |
| Control (Con.) | N/A | | N/A | | **<.001** | |
| Loss Framing | | | | | **<.001** | |
| Low-Sensitive | | | | | | |
| Con. vs. Loss Framing | N/A | | N/A | | **.031** | |
| High-Sensitive | | | | | | |
| Con. vs. Loss Framing | N/A | | N/A | | <1.0 | |

TABLE VII: Statistical test results of Experiment 1.

| Term | Local Only | | Both | | Opt Out | |
|---|---|---|---|---|---|---|
| | $\chi^2$ | $p$ | $\chi^2$ | $p$ | $\chi^2$ | $p$ |
| Question Sensitivity | **43.38** | **<.001** | **325.65** | **<.001** | **290.0** | **<.001** |
| Condition | **40.31** | **<.001** | **30.31** | **<.001** | 4.95 | .175 |
| Con vs. Gain Framing | **<.001** | | **.012** | | N/A | |
| Con vs. DP | **<.001** | | **<.001** | | | |
| Con vs. LDP | .101 | | **.017** | | | |
| Gain Framing vs. DP | .380 | | .165 | | | |
| Gain Framing vs. LDP | .499 | | >.999 | | | |
| DP vs. LDP | **<.001** | | **.045** | | | |
| Question Sensitivity * Condition | **21.59** | **<.001** | **22.45** | **<.001** | 3.03 | .388 |
| Low-Sensitive vs. High-Sensitive | | | | | | |
| Control | **<.001** | | **<.001** | | N/A | |
| Gain Framing | **<.001** | | **<.001** | | | |
| DP | .113 | | **<.001** | | | |
| LDP | .113 | | **<.001** | | | |
| Low-Sensitive | | | | | | |
| Con vs. Gain Framing | **<.001** | | **<.001** | | N/A | |
| Con vs. DP | **.025** | | **.018** | | | |
| Con vs. LDP | >.999 | | >.999 | | | |
| Gain Framing vs. DP | .283 | | .843 | | | |
| Gain Framing vs. LDP | **.001** | | **.013** | | | |
| DP vs. LDP | **.047** | | .409 | | | |
| High-Sensitive | | | | | | |
| Con vs. Gain Framing | >.999 | | >.999 | | N/A | |
| Con vs. DP | **<.001** | | **<.001** | | | |
| Con vs. LDP | **.021** | | **.017** | | | |
| Gain Framing vs. DP | **<.001** | | **<.001** | | | |
| Gain Framing vs. LDP | .557 | | **.048** | | | |
| DP vs. LDP | **.024** | | .258 | | | |

TABLE VIII: Statistical test results of Experiment 2A.

| Term | Local Only | | Both | | Opt Out | |
|---|---|---|---|---|---|---|
| | $\chi^2$ | $p$ | $\chi^2$ | $p$ | $\chi^2$ | $p$ |
| Question Sensitivity | **41.54** | **<.001** | **331.56** | **<.001** | **372.68** | **<.001** |
| Condition | 3.88 | .274 | <1.0 | | 3.91 | .271 |
| LDP w/o Names vs. DP w/o Names | N/A | | N/A | | N/A | |
| LDP w/o Names vs. LDP Comp. w/o Names | | | | | | |
| LDP w/o Names vs. LDP Comp. | | | | | | |
| LDP Comp. vs. LDP Comp. w/o Names | | | | | | |
| DP w/o Names vs. LDP Comp. w/o Names | | | | | | |
| DP w/o names vs. LDP Comp. | | | | | | |
| Question Sensitivity * Condition | **12.74** | **.005** | **9.06** | **.029** | **7.88** | **.048** |
| Low-Sensitive vs. High-Sensitive | | | | | | |
| DP w/o Names | .452 | | **<.001** | | **<.001** | |
| LDP w/o Names | **.019** | | **<.001** | | **<.001** | |
| LDP Comp. w/o Names | **<.001** | | **<.001** | | **<.001** | |
| LDP Comp. | **.004** | | **<.001** | | **<.001** | |
| Low-Sensitive | | | | | | |
| LDP w/o Names vs. DP w/o Names | .850 | | .270 | | .892 | |
| LDP w/o Names vs. LDP Comp. w/o Names | >.999 | | >.999 | | >.999 | |
| LDP w/o Names vs. LDP Comp. | >.999 | | >.999 | | .660 | |
| LDP Comp. vs. LDP Comp. w/o Names | >.999 | | >.999 | | >.999 | |
| DP w/o Names vs. LDP Comp. w/o Names | .356 | | >.999 | | >.999 | |
| DP w/o names vs. LDP Comp. | >.999 | | .366 | | >.999 | |
| High-Sensitive | | | | | | |
| LDP w/o Names vs. DP w/o Names | >.999 | | >.999 | | .681 | |
| LDP w/o Names vs. LDP Comp. w/o Names | **.014** | | >.999 | | **.020** | |
| LDP w/o Names vs. LDP Comp. | >.999 | | >.999 | | .544 | |
| LDP Comp. vs. LDP Comp. w/o Names | .208 | | >.999 | | >.999 | |
| DP w/o Names vs. LDP Comp. w/o Names | **.017** | | >.999 | | >.999 | |
| DP w/o names vs. LDP Comp. | >.999 | | .248 | | >.999 | |

TABLE IX: Statistical test results of Experiment 2B.

| Term | Local Only | | Both | | Opt Out | |
|---|---|---|---|---|---|---|
| | $\chi^2$ | $p$ | $\chi^2$ | $p$ | $\chi^2$ | $p$ |
| Question Sensitivity | **22.58** | **<.001** | **206.02** | **<.001** | **229.49** | **<.001** |
| Condition | **12.36** | **.002** | **29.2** | **<.001** | **10.77** | **.005** |
| LDP Imp. vs. DP Imp. | **.049** | | **.037** | | **.026** | |
| LDP Imp. vs. LDP Imp. w/o Local | >.999 | | >.999 | | >.999 | |
| DP Imp. vs. LDP Imp. w/o Local | **.002** | | **<.001** | | **.012** | |
| Question Sensitivity * Condition | <1.0 | | 1.36 | .508 | 3.61 | .164 |
| Low-Sensitive vs. High-Sensitive | | | | | | |
| DP Imp. | N/A | | N/A | | N/A | |
| LDP Imp. | | | | | | |
| LDP Imp. w/o Local | | | | | | |
| Low-Sensitive | | | | | | |
| LDP Imp. vs. DP Imp. | N/A | | N/A | | N/A | |
| LDP Imp. vs. LDP Imp. w/o Local | | | | | | |
| DP Imp.vs. LDP Imp. w/o Local | | | | | | |
| High-Sensitive | | | | | | |
| LDP Imp. vs. DP Imp. | N/A | | N/A | | N/A | |
| LDP Imp. vs. LDP Imp. w/o Local | | | | | | |
| DP Imp. vs. LDP Imp. w/o Local | | | | | | |

TABLE X: Statistical test results of Experiment 4.

| Question | Term | $\chi^2$ | $p$ |
|---|---|---|---|
| Q1. Privacy_Attackers | Technique | **9.49** | **.002** |
| | Description | **102.3** | **<.001** |
| | w/o Names vs. Imp. | **<.001** | |
| | w/o Names vs. Flow | **<.001** | |
| | Imp. Vs. Flow | **.002** | |
| | Technique * Description | **15.91** | **<.001** |
| | DP vs. LDP | | |
| | w/o Names | **.004** | |
| | Imp. | .107 | |
| | Flow | **.001** | |
| Q2. Privacy_Employees | Technique | **4.09** | **.043** |
| | Description | **45.6** | **<.001** |
| | w/o Names vs. Imp. | **.042** | |
| | w/o Names vs. Flow | **<.001** | |
| | Imp. vs. Flow | **.001** | |
| | Technique * Description | **10.77** | **.005** |
| | DP vs. LDP | | |
| | w/o Names | >.999 | |
| | Imp. | >.999 | |
| | Flow | **.001** | |
| Q3. Privacy_Third Party | Technique | <1.0 | |
| | Description | **24.77** | **<.001** |
| | w/o Names vs. Imp. | .272 | |
| | w/o Names vs. Flow | **<.001** | |
| | Imp. Vs. Flow | **.006** | |
| | Technique * Description | **12.66** | **.002** |
| | DP vs. LDP | | |
| | w/o Names | **.032** | |
| | Imp. | .926 | |
| | Flow | **.040** | |
| Q4. Utility_Cost | Technique | <1.0 | |
| | Description | **56.33** | **<.001** |
| | w/o Names vs. Imp. | .274 | |
| | w/o Names vs. Flow | **<.001** | |
| | Imp. Vs. Flow | **<.001** | |
| | Technique * Description | 2.02 | .365 |
| | DP vs. LDP | | |
| | w/o Names | NA | |
| | Imp. | NA | |
| | Flow | NA | |
| Q5. Utility_Third Party | Technique | **4.07** | **.044** |
| | Description | **37.32** | **<.001** |
| | w/o Names vs. Imp. | .115 | |
| | w/o Names vs. Flow | **<.001** | |
| | Imp. vs. Flow | **.001** | |
| | Technique * Description | **8.96** | **.011** |
| | DP vs. LDP | | |
| | w/o Names | .666 | |
| | Imp. | **.002** | |
| | Flow | .605 | |

TABLE XI: Summary of participants who were excluded from data analysis for all experiments. The number of participants before exclusion in each experiment is listed in the first column.

| **EXP.** | **Median Complete Time (sec)** | **Payment ($)** | **Same IP** | **Time < 120 sec** | **2nd Check Question Failure** | |
|---|---|---|---|---|---|---|
| | | | | | **Condition** | **Count** |
| Pilot 1 (219) | 264 | 1 | 4 | 11 | N/A | |
| EXP. 1 (598) | 289 | | 12 | 28 | DP | 42 |
| | | | | | LDP | 51 |
| EXP. 2A (781) | 328 | | 15 | 19 | DP w/o Names | 58 |
| | | | | | LDP w/o Names | 46 |
| | | | | | LDP Comp. w/o Names | 30 |
| | | | | | LDP Comp. | 32 |
| EXP. 2B (600) | 303 | | 9 | 15 | DP Imp. | 28 |
| | | | | | LDP Imp. | 43 |
| | | | | | LDP Imp. w/o Local | 37 |
| EXP. 3 (279) | 192 | 0.75 | 1 | N/A | N/A | |
| Pilot 2 (20) | 303 | 1 | 0 | 0 | N/A | |
| EXP. 4 (599) | 245 | | 7 | 52 | N/A | |

TABLE XII: Proposed descriptions of DP and LDP techniques of each condition in each experiment. Descriptions of DP w/o Names, LDP w/o Names, and LDP Comp. w/o Names are the same as DP, LDP, and LDP Comp., respectively, except the last sentences in square brackets are deleted. DP Flow and LDP Flow are updated based on results of Pilot Study 2 (in bold font).

| Condition | Description |
|---|---|
| **DP** | To respect your personal information privacy and ensure best user experience, the data shared with the app will be processed via the differential privacy (DP) technique. DP protects users' privacy by adding random noise to aggregated data, for example, average age, such that the probability that an individual's information is inferred is low. [DP has been used across academia and industry, including Harvard University, U.S. Census Bureau, and companies such as LinkedIn and Uber.] |
| **LDP** | To respect your personal information privacy and ensure best user experience, the data shared with the app will be collected via the local differential privacy (LDP) technique. LDP protects users' privacy by adding random noise to each response that users give such that the probability that any user's attribute is inferred is similar as he or she is opt-out for the data collection. [LDP has been used by companies such as Google and Apple.] |
| **LDP Comp** | To respect your personal information privacy and ensure best user experience, the data shared with the app will be collected via the local differential privacy (LDP) technique. LDP protects your privacy by adding random noise to the raw data locally BEFORE you give the data to the company (raw data never leaves your device). [LDP has been used by companies such as Google and Apple.] |
| **DP Imp.** | To respect your personal information privacy and ensure best user experience, the data shared with the app will be processed via the differential privacy (DP) technique. That is, the app company will store your data but only use the aggregated statistics with modification so that your personal information cannot be learned. However, your personal information may be leaked if the company's database is compromised. |
| **LDP Imp.** | To respect your personal information privacy and ensure best user experience, the data shared with the app will be processed via the local differential privacy (LDP) technique. That is, the app will randomly modify your data on your cellphone before sending it to the app server. Since the app server stores only the modified version of your personal information, your privacy is protected even if the app server's database is compromised. |
| **LDP Imp. w/o Local** | To respect your personal information privacy and ensure best user experience, the data shared with the app will be processed via the differential privacy (DP) technique. That is, the app will randomly modify your data on your cellphone before sending it to the app server. Since the app server stores only the modified version of your personal information, your privacy is protected even if the app server's database is compromised. |
| **DP Flow** | When differential privacy (DP) is used, the app sends the user's answers to the company. These answers are then stored in the company's databases. When the company wants to use the collected data, either internally or for sharing with third-party companies, the company sends queries to the databases, applies DP techniques to modify the returned results, and uses only the modified results. These modified results reveal limited information specific to each individual user. However, by examining the modified answers of a large number of users, the company can still get useful summary results in the user population, **even though the accuracy is reduced (compared to the case when no modification is applied)**. |
| **LDP Flow** | When local differential privacy (LDP) is used, the app modifies the answers before sending them from the user's device to the company. The company only sees and stores the modified version of each user's information, and is uncertain about each individual user's true answer. However, by examining the modified answers of a large number of users, the company can still get useful summary results in the user population, **even though the accuracy is reduced (compared to the case when no modification is applied)**. |
| **Apple** | Differential privacy transforms the information shared with the company before it ever leaves the user's device such that the company can never reproduce the true data. The differential privacy technology used is rooted in the idea that statistical noise that is slightly biased can mask a user's individual data before it is shared with the company. If many people are submitting the same data, the noise that has been added can average out over large numbers of data points, and the company can see meaningful information emerge. |
| **Google** | Building on the concept of randomized response, local differential privacy (LDP) enables learning statistics about the behavior of users' software while guaranteeing client privacy. LDP builds on the above concept, allowing software to send reports that are effectively indistinguishable from the results of random coin flips and are free of any unique identifiers. However, by aggregating the reports we can learn the common statistics that are shared by many users. |
| **Microsoft** | Differential privacy is a technology that enables researchers and analysts to extract useful answers from databases containing personal information and, at the same time, offers strong individual privacy protections. This seemingly contradictory outcome is achieved by introducing relatively small inaccuracies in the answers provided by the system. These inaccuracies are large enough that they protect privacy, but small enough that the answers provided to analysts and researchers are still useful. |
| **Uber** | Differential privacy is a formal definition of privacy and is widely recognized by industry experts as providing strong and robust privacy assurances for individuals. In short, differential privacy allows general statistical analysis without revealing information about a particular individual in the data. Results do not even reveal whether any individual appears in the data. For this reason, differential privacy provides an extra layer of protection against re-identification attacks as well as attacks using auxiliary data. |
| **US Census Bureau** | Differential privacy was developed by researchers at Microsoft and is now utilized by many leading tech firms. There are many variants of differential privacy. The one we selected introduces controlled noise into the data in a manner that preserves the accuracy at higher levels of geography. Our differential privacy methods will be designed to preserve the utility of our legally mandated data products while also ensuring that every respondents' personal information is fully protected. |