Measuring the Effectiveness of Spinner-based Randomized-Response Differential Privacy Communication for Sensitive Data Sharing

Seo Young Ko, Sriram Viswanathan, Alan Esquenazi, Swadhin Routray, Jatan Loya, Tianshi Li, and Lorrie Cranor

Carnegie Mellon University

Abstract

With the world moving in a direction that is heavily dependent on data, it is essential to ensure that this exchange of data does not lead to harm to humans in terms of privacy. To protect data privacy, corporations often adopt differential privacy, which is one of the Privacy Enhancing technologies (PETs), to enable privacy-preserving data sharing. In this work, we measure the usability of Differential Privacy (DP) communication, especially the Randomized Response Technique (RRT) in terms of users capability to add valid noises based on their comprehension and comfort level. We measure the effectiveness of a spinner-based RRT interface by conducting an exploratory between-group online survey which collects sensitive information in a hypothetical scenario. We found that many participants did not follow instructions on the spinner interface to add noise correctly, and did not understand the purpose of the spinner. Based on this finding, we provide suggestions for future work to improve the RRT interface for better DP communication.

1 Introduction

With the prevalence of data collection, it has become more crucial than ever to protect data privacy in our daily lives. It is particularly challenging to collect and analyze sensitive information such as medical or sexual information without privacy concerns. To enable privacy-preserving data sharing, there is a wide range of privacy-enhancing technologies (PETs), applicable to different privacy threat vectors. One of the commonly adopted PETs is Differential Privacy (DP) [4], which gives a dataset owner the ability to release a given dataset in such a way that sensitive data is not leaked, thus offering strong privacy guarantees using a precise mathematical model.

Given the technical and mathematical complexity of a technique like DP, the usability of such a technique by non-expert users is a challenge. There is existing literature that points out the lack of usability of DP and examines ways to enhance the description and interface for better user comprehension. [2, 3, 5, 8, 12, 13]. Karegar et al. [5] utilizes pictorial metaphors based approaches for better explanation of DP and Bullek et al. [2] introduces spinner based interface to enable transparent user-led local DP with a higher trust level. However, little work studied the effectiveness of the process of applying DP empirically, such as whether users are able to understand the mechanism and add a valid noise, which may affect the data analysis for future usefulness.

This work aims to improve the transparency of the process of applying DP and improve user agency in the process by designing interfaces that guide users in adding noise to their data. We focused on the Randomized Response Technique (RRT), a local DP mechanism, in which random noise is added at the individual level before sending the data to the server or administrator. Inspired by prior work that studied using spinner interfaces to explain the results of the RRT mechanism [2], we designed spinner interfaces to provide users with instructions for adding noise.

We conducted a between-group survey with one control group and two experimental groups to test the effectiveness of our spinner interfaces for user-led RRT communication grounded in a survey-taking scenario. The study uses both simple definitions and the visual spinner tool for the survey takers to understand the mechanism of adding noise when a sensitive question is asked in the hypothetical online survey. Upon gathering data, we perform both qualitative and quantitative analysis on the collected data. Our analysis results identified some limitations of the spinner RRT interface proposed by existing literature [2] in terms of the understandability and the capability to guide users to add noise correctly. Also, we propose a few suggestions to improve the RRT interface as a future research direction.

The rest of this paper is organized as follows. In Section 2, we present related work in this domain, with a focus on initial work in the area of explainable differential privacy, how privacy is communicated to users, and a short introduction to the randomized response mechanism and differential privacy. In Section 3, we describe our study methodology and define our recruitment strategy, our survey design, and data analysis. In Section 4, we outline the results we could draw from our data. Our limitations are acknowledged in Section 5. We present our recommendations for future work and conclusion in Section 6 and Section 7 respectively. Our codebook and our survey questions can be found in Section 8.

2 Related Work

2.1 Randomized Response Technique

Randomized Response Technique (RRT) is initially introduced by Warner [11] to enhance privacy and eliminate the bias as a survey technique. Nowadays, RRT is often used to facilitate collecting binary sensitive information with privacy protection with the DP guarantee by plausible deniability as a local differential privacy technique [10]. This method uses a certain randomization method, such as a coin flip, card, or spinner, to introduce random noise rather than collecting the raw answers. The method conceals individual responses and protects respondent privacy with such noise addition mechanism [1].

As an example, the randomized response mechanism can be simulated by using a visual spinner, as seen in Figure 1. The spinner will land on one of three options: “Answer Truthfully”, “Answer Yes” or “Answer no” with different probabilities. Since all respondents use the spinner to randomize their sensitive answers by following the guide correctly, “Yes” responses cannot be interpreted as true "Yes". However, the interviewers can statistically deduce the approximate frequency of the sensitive answers by examining the results in aggregate based on the known probabilities of three options in the spinner [2]. This protects respondents’ privacy by giving them plausible deniability while allowing for the computation of accurate measurable statistics, such as counts in populations or, in the case of our survey, answers to sensitive questions.

2.2 Usable Differential Privacy

To trust the algorithm used or deployed as part of the differential privacy practice of a corporation, the users must gain adequate understanding of the algorithm. Bullek et al. [2] found that allowing users to know the amount of obfuscation applied to their answers increases their trust in differential privacy. In their study, the authors designed a method that communicates how a specific RRT algorithm works. They checked if users, upon understanding obfuscation by the algorithm, trust it to answer sensitive information. It is shown that if the amount of noise is transparent, people are more likely to trust, and users vastly preferred the most anonymous spinner [2].

To test the users’ understanding of privacy, Smart et al. [9] designed a survey study that tested different explanations of differential privacy with different levels of privacy setting disclosures. They first aimed to develop good explanations for privacy and then measured how those explanations influenced a participant’s behavior in sharing their browser history data. The results suggest that certain aspects of differential privacy remain challenging for people to understand and explanations of DP does not significantly affect users in their decision-making for data sharing.

On the other hand, Karegar et al. [5] utilized metaphors to explain both local and central differential privacy by generating pictorial metaphors elaborated with short and simple text. They identified that although the metaphorical explanation effectively conveys the trade-off between privacy and accuracy, it has an inherent limitation that can lead to incorrect expectations and understanding. They also utilize the spinner for metaphors differently from Bullek et al. [2].

3 Methodology

3.1 Study Design

We conducted an online survey for our exploratory and experimental study. We used Qualtrics to implement the survey and prototypes, and we used Prolific to distribute to 60 participants. All participants were given a hypothetical scenario regarding the use of recreational drugs. Specifically, we asked them to imagine that they were taking an online survey asking about their experiences of using recreational drugs. We asked them to imagine that they had used recreational drugs in the past year so we know what answers they should provide if they correctly followed the spinner’s guidance to apply RRT. We do not ask for data about the users’ actual experiences. Then, users are divided into three groups of 20 participants each according to the between-group study design.

We designed to have one control group and two experimental groups with different approaches to add random noise in order to demonstrate the effectiveness of the spinner interface. The first group is the `Control` group, whose survey does not include DP explanation and mechanism. The second and third experimental groups use DP and receive explanations about it. The second one, called `AutomaticNoise`, uses randomized-response DP with machine-added noise, so users will only be informed that their answer was anonymized through the use of DP without user interaction. The third group, referred to as `SpinnerNoise`, uses randomized-response DP with user-added noise. To implement the technique for guiding users to add noise, we used a spinner, as has been previously done in studies such as Bullek et al. [2], Karegar et al. [5] and Blair et al. [1]. The code to implement our spinner was modified from publicly available code on Github [7]. Our user-led noise addition prototype currently looks as shown in Figure 1.

1. `Control`: No DP mechanism
2. `AutomaticNoise`: Textual RRT with machine-added noise
3. `SpinnerNoise`: Spinner RRT with user-led noise

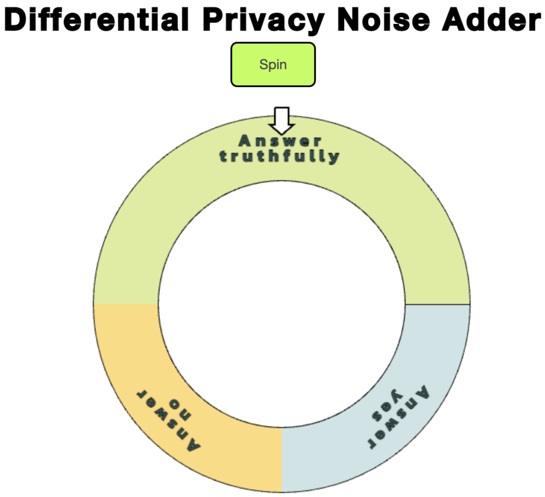


Figure 1: RRT Spinner Interface. To answer the sensitive question, a respondent first clicks on the "spin" button. If the arrow of the spinner points towards only "answer truthfully" (with 50% probability), the respondent should answer truthfully. In the other cases, when the arrow points towards "answer yes" or "answer no" (with 25% probability for each), the respondent should answer "Yes" or "No" for plausible deniability.

Users are asked demographic questions for all groups, as well as questions regarding the differential privacy mechanism (depending on which group they belong to), their understanding of it, and their trust in it. We also measured the participants' predispositions and attitudes towards technology and privacy, in general, using previous research from Malhotra [6], where they developed a model for representing information privacy concerns of online users.

3.2 Recruitment, Compensation, and Demographics

We used Prolific to distribute our survey. We have collected data from 60 participants, 20 for each of the condition groups. The survey took about 10 minutes to complete, and each participant was compensated 3 USD for completion. The demographic information including genders and ages of all participants is summarized in Figure 2.

3.3 Ethics

Our study was IRB approved. We did not collect any personally identifiable data from our participants, and the usage of a fictitious scenario helps mitigate any concerns about collecting real sensitive data from the participants.

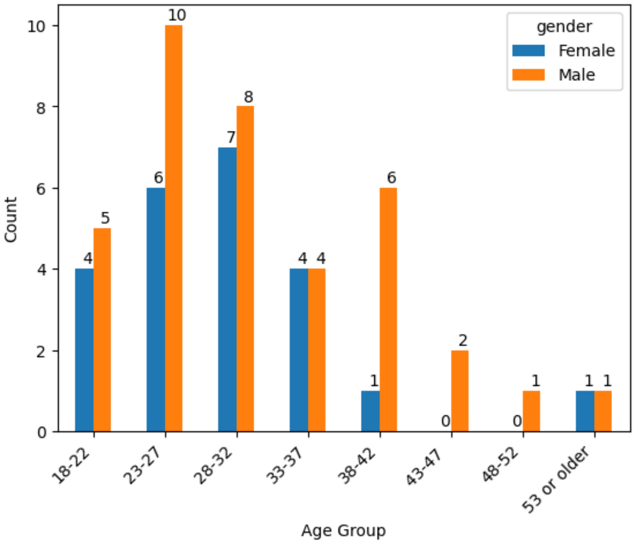


Figure 2: Participant Demographics

3.4 Data Analysis

3.4.1 Quantitative Data Analysis

From a data analysis perspective, our main independent variable is the study condition (`Control`, `AutomaticNoise`, and `SpinnerNoise`). Our dependent variables include the user's comprehension, trust level with the noise addition method, honesty level, and the comfort level regarding the sensitive questions asked of them. These variables are measured on a 5-point Likert scale.

To distinguish between actual and perceived comprehension levels, we calculated the average of participants' responses to questions that assessed their factual understanding of the noise mechanism, its guarantees in protecting sensitive information and maintaining secrecy, which represents the participants' actual comprehension level. Conversely, we used participants' self-reported level of comprehension to determine their perceived comprehension level.

3.4.2 Qualitative Data Analysis

To help us better understand the user responses, we also collected qualitative data, where we ask users to elaborate on their concerns and opinions in the form of free-text responses. For the data analysis, two researchers coded the open-ended questions in our survey independently using inductive coding. Then they resolved conflicts and reached a consensus on all answers. The results of the qualitative data analysis are discussed in the Section 4, and the codebooks and occurrences of each code can be found in the Appendix (Section 8).

4 Results

Overall, our result demonstrates that the RRT spinner interface for differential privacy communication is not as effective as expected to guide users to add noise by themselves. Our quantitative and qualitative data analysis have led to interesting observations.

4.1 User Capability To Add Noise

We measured whether users could correctly add noise with the spinner interface. Since we only tested `SpinnerNoise` group that used a user-led interface, there are only 20 responses for this measurement. Based on the recorded spinner responses, we compare their self-reported answers with the spinner responses to verify if they are able to add noise correctly. Since we are using the hypothetical scenarios, we can know the ground truth of the answer and measured whether they followed the guide successfully or not.

We found that only 13 out of 20 in `SpinnerNoise` participants were able to correctly follow the instructions and add noise, as shown in Table 1. Even with explicit "Answer Yes" and "Answer No" responses, not all participants follow the guide successfully. This suggests the ineffectiveness of the interface in communicating the essential information required by the participant to add noise correctly, as per the spinner response.

Spinner Response	The number of people who added noise correctly
"Answer Truthfully"	6 out 11 (54%)
"Answer No"	4 out of 5 (80%)
"Answer Yes"	3 out of 4 (75%)
Total	13 out of 20 (65%)

Table 1: The percentage of adding valid noise with spinner

4.2 Misunderstanding, Discomfort, Dishonesty, and Distrust

In our quantitative data analysis, we calculated the percentage of "Strongly Agree" or "Somewhat Agree" for each measure from 5-point Likert responses. Participants across all groups reported fairly high levels of honesty (86.7%) and comfort (61.7%), and relatively lower levels of trust (32.5%) and perceived comprehension (27.5%). Their actual comprehension was higher (66.3%). However, we found no significant differences between groups on any of these measures.

We also discovered that the qualitative data supports the quantitative findings. For example, the question on understanding the DP mechanism shows similar amounts of confusion, and similar concepts brought up between `AutomaticNoise` and `SpinnerNoise` groups. In this question, we have also seen "Increased Honesty" being brought up organically as a benefit of the differential privacy mechanism, which we believe to be positive, as this is what we are trying to achieve through the mechanism.

Some reasons for discomfort for `AutomaticNoise` were the lack of transparency and lack of trust in the mechanism. This is also related to users' explanations of reasons for untrustworthiness, where we found, in `AutomaticNoise`, complaints about lack of transparency:

> "Just being told that it exists doesn't mean it will work"

For the reasons for untrustworthiness, we also saw some lack of understanding of the spinner mechanism in `SpinnerNoise` group:

> "I don't fully understand how a spinner creates noise over my data"

We found significant bias that affect users' trust and comfort levels related to their previous conceptions surrounding the internet and privacy-preserving mechanisms. Some users believed they were always being tracked; Others thought that the mechanism would not be able to protect them from data leaks. On the other hand, another user considered that the noise addition mechanism itself is unnecessary on top of existing protections provided by the default survey platform (even though we never specified what these were). We also found one user mentioning unfamiliarity as a reason for untrustworthiness, which is reasonable, as differential privacy is relatively recently deployed as a user-facing mechanism.

We also found an interesting phenomenon where many users felt uncomfortable or were dishonest due to fear of law enforcement, as the question inquired about the usage of illegal drugs. This was not one of our goals with this survey, but with hindsight, we can understand why this topic could have drawn that reaction.

4.3 Correlations with Privacy Attitudes

We measured general privacy attitudes using Malhotra [6] to understand how the general privacy attitudes of the participants correlate with their comprehension level and trust levels in the noise-adding mechanism. We performed a Pearson-R correlation between privacy attitudes and the comprehension levels, as well as privacy attitudes and trust levels in the noise-adding mechanisms.

- Actual comprehension level: $r = 0.399$, $p = 0.0107$
- Trust level in noise addition mechanism: $r = 0.36$, $p = 0.02$

While there was no significant correlation between privacy attitudes and the perceived comprehension level, we found statistically significant and medium positive correlation in the above two cases. A possible explanation is that as people are more cautious about data collection and risks, they are more likely to understand and trust the noise addition mechanism. This gives an additional insight on how privacy-attitudes affect the usability of RRT interfaces.

5 Limitations

Our study was limited by a small sample size of participants (20 participants per group, 60 total). The small sample size led to a lack of statistical power, which may be the reason we did not find any significant correlation between the study condition and the dependent variables. We believe a larger sample size might reveal clearer patterns that we were not able to observe in this small dataset.

In addition, our study only asks about one kind of sensitive information which is the usage of recreational drugs, and the finding may not apply to other questions with different levels of perceived sensitivity.

Lastly, we tried to mimic authentic survey-taking experiences by designing the survey question around a realistic situation, but the use of a hypothetical scenario may still cause confusion and result in a lack of ecological validity.

6 Discussion and Future Work

Based on our findings, we present a few suggestions and ideas that might help future research design more effective DP communication techniques.

6.1 Complement Study Design

While our survey provided valuable data, we believe that an interview study would help with providing more in-depth insights into users' thought processes. This will help uncover what specific features of the spinner interface that require improvement in terms of understandability and usability.

We believe that reframing of the study could help to mitigate the limitation that the survey is using hypothetical scenarios to answer a sensitive question. For example, we could assign an unrelated task and ask a question based on the task. However, it can be challenging to identify a suitable unrelated task that involves sensitive questions.

6.2 Potential Improvements of Spinner Interfaces for DP Communication

We believe that the ineffectiveness of the spinner-based interface may be attributed to the lack of understanding about DP it provides or the lack of trust in the mechanism.

In our pilot tests, we observed that users cannot understand how the spinner interface works until they get a sense of the aggregation and possibility of data analysis with a noise-added database. We believe that instead of focusing on a narrow view of adding noise, we should provide a holistic view of the data life cycle to enhance understandability. We also believe that this can be regarded as an inherent limitation of spinner interface for communication because the purpose of it is to visualize the noise addition mechanism, which is the narrow view of RRT.

It is also possible that users may distrust that the spinner is truly random due to their prior experiences where its results were not genuinely random by design, such as giving coupons or winning a lottery. To address such bias, we can consider using an alternative digital interface that fosters trust in the randomness of the mechanism. Another option is to guide users to follow an offline coin-flip mechanism (in place of a spinner) to add noise, as it has been previously suggested in the foundational literature on differential privacy [4]. A few participants expressed concerns over tracking and data leakage that we think may be related to the fact that the noise-addition mechanisms are all virtual. We hypothesize that giving participants even more control over the noise-addition mechanism through the use of a coin flip, controlled by the participants themselves, might help ease these concerns.

Given that the current implementation is only pictorial and visual, using multi-modal interface would be effective. For example, we could change the explanation for the noise addition mechanism to make it more descriptive and audio-visual, in the form of a video or an infographic, to better represent the process of adding noise to a response. We believe such changes might lead to a better understanding of the mechanism.

7 Conclusion

In this work, we conducted a between-group survey to investigate the effectiveness of RRT interfaces in terms of DP communication and summarized our suggestions based on the findings. We found that the spinner mechanism has some limitations in terms of the user's capability to add noise following the instruction of the interface based on the data analysis. This implies that the spinner RRT mechanism lacks providing the correct communication of adding noise by itself. We could not find enough evidence to show the difference in the user's comprehension, honesty, and trust levels of using the spinner RRT interface from the quantitative data analysis. However, the qualitative data analysis showcases the lack of understanding even with the spinner interface.

Based on these findings, we recommend to design RRT interfaces to include a holistic view of the data processing cycle instead of a narrow view of noise addition to achieve better comprehension. More qualitative research is required in understanding the reasons for the lack of effectiveness of such an interface and ways in which we can improve them.

References

[1] Graeme Blair, Kosuke Imai, and Yang-Yang Zhou. Design and analysis of the randomized response technique. *Journal of the American Statistical Association*, 110(511):1304–1319, 2015. doi: 10.1080/01621459.2015.1050028. URL https://doi.org/10.1080/01621459.2015.1050028.

[2] Brooke Bullek, Stephanie Garboski, Darakhshan J. Mir, and Evan M. Peck. Towards Understanding Differential Privacy: When Do People Trust Randomized Response Technique? In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 3833–3837, Denver Colorado USA, May 2017. ACM. ISBN 9781450346559. doi: 10.1145/3025453.3025698. URL https://dl.acm.org/doi/10.1145/3025453.3025698.

[3] Rachel Cummings, Gabriel Kaptchuk, and Elissa M. Redmiles. "i need a better description": An investigation into user expectations for differential privacy. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, CCS '21, page 3037–3052, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450384544. doi: 10.1145/3460120.3485252. URL https://doi.org/10.1145/3460120.3485252.

[4] Cynthia Dwork. Differential privacy. In Michele Bugliesi, Bart Preneel, Vladimiro Sassone, and Ingo Wegener, editors, *Automata, Languages and Programming*, pages 1–12, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg. ISBN 978-3-540-35908-1.

[5] Farzaneh Karegar, Ala Sarah Alaqra, and Simone Fischer-Hübner. Exploring user-suitable metaphors for differentially private data analyses. In *Proceedings of the Eighteenth USENIX Conference on Usable Privacy and Security*, SOUPS'22, USA, 2022. USENIX Association. ISBN 978-1-939133-30-4.

[6] Agarwal Malhotra, Kim. Internet users' information privacy concerns (iuipc): The construct, the scale, and a causal model. (CSCW2), December 2004. doi: 10.1287/isre.1040.0032. URL https://journals.sagepub.com/doi/10.1177/1094670514539730.

[7] MikeyC0340. dinnerspinner, July 2022. URL https://github.com/MikeyC3040/dinnerSpinner.

[8] Jack Murtagh, Kathryn Taylor, George Kellaris, and Salil Vadhan. Usable Differential Privacy: A Case Study with PSI, September 2018. URL http://arxiv.org/abs/1809.04103. arXiv:1809.04103 [cs].

[9] Mary Anne Smart, Dhruv Sood, and Kristen Vaccaro. Understanding Risks of Privacy Theater with Differential Privacy. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2):1–24, November 2022. ISSN 2573-0142. doi: 10.1145/3555762. URL https://dl.acm.org/doi/10.1145/3555762.

[10] Teng Wang, Xuefeng Zhang, Jingyu Feng, and Xinyu Yang. A comprehensive survey on local differential privacy toward data statistics and analysis. *Sensors*, 20 (24):7030, dec 2020. doi: 10.3390/s20247030. URL https://doi.org/10.3390%2Fs20247030.

[11] Stanley L. Warner. Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309): 63–69, 1965. doi: 10.1080/01621459.1965.10480775. URL https://www.tandfonline.com/doi/abs/10.1080/01621459.1965.10480775. PMID: 12261830.

[12] Felix Wolter and Peter Preisendörfer. Asking sensitive questions: An evaluation of the randomized response technique versus direct questioning using individual validation data. *Sociological Methods & Research*, 42(3): 321–353, 2013.

[13] Aiping Xiong, Tianhao Wang, Ninghui Li, and Somesh Jha. Towards effective differential privacy communication for users' data sharing decision and comprehension, 2020.

8 Appendix

8.1 Appendix 1: Survey Questions

Here is our survey, including the flow, as seen in Qualtrics. Notice that users are divided into three condition groups. Also, in the online version, some of the questions (specifically, some of the open-response questions) were randomized in order to avoid biasing effects.

Imagine you have used recreational drugs in the past year. One day, you saw a paid survey on an online forum for research purposes. You were interested in earning some money and wanted to fill it out. Then you noticed that there were some sensitive questions related to your experience of using recreational drugs.

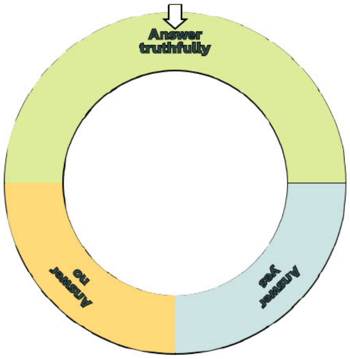
Consider the following definition of **LDP:**

Local Differential Privacy or LDP protects users' privacy by adding random noise to each response that users give and provides a statistical guarantee of privacy. So, no one can know, just by looking at the overall data, what a specific user's response is, or even if a user was part of the data collection. LDP has been used by companies such as Google and Apple to collect information from users in a privacy-preserving manner.

We are going to use the LDP technique here – you will be adding noise to the answer by using a spinner. To correctly add the noise to your answer, spin the wheel and follow the instructions.

Differential Privacy Noise Adder

Spin



Please answer the question below, imagining what you would do if you had actually used recreational drugs in the past year and were presented with a paid survey online for research purposes that uses the LDP spinner technique described above.

Have you used recreational drugs in the past 1 year?

Spin the spinner and:
If the spinner lands on "Answer Yes", then answer "Yes".
If the spinner lands on "Answer No", then answer "No"
If the spinner lands on "Answer Truthfully", then answer truthfully.

Yes No

To what extent do you agree or disagree with the below statement about the noise-adding mechanism used in the previous question:

The noise addition guarantees the secrecy of my sensitive data (whether I used recreational drugs in the past year or not).

Strongly Agree
Somewhat agree
Neither agree nor disagree
Somewhat disagree
Strongly disagree

To what extent do you agree or disagree with the below statement about the noise-adding mechanism used in the previous question:

The noise addition guarantees that my sensitive data (whether I used recreational drugs in the past year or not) will not be revealed, even if the server's data is disclosed because of an attack.

Strongly Agree
Somewhat agree
Neither agree nor disagree
Somewhat disagree
Strongly disagree

How would you rate your level of understanding of the noise addition mechanism?

Very High
High
Medium
Low
Very Low

To what extent do you agree or disagree with this statement:

The prior description of the privacy protection technique was easy for me to understand.

Strongly Agree
Somewhat agree
Neither agree nor disagree
Somewhat disagree
Strongly disagree

To what extent do you agree or disagree with this statement:
I expect the noise addition technique to provide benefits.

Strongly Agree
Somewhat agree
Neither agree nor disagree
Somewhat disagree
Strongly disagree

What benefit do you expect from adding noise based on the aforementioned explanation?

What was your comfort level when taking this survey and being asked sensitive questions in a hypothetical scenario?

Very High
High
Medium
Low
Very Low

What was the reason for you to feel discomfort?

If this had been a real situation and you were asked to disclose your recreational drug use (or similar sensitive information) within the same online survey interface, how much do you agree that you would be honest in answering this question based on the noise-addition technique?

Strongly agree
Somewhat agree
Neither agree nor disagree
Somewhat disagree
Strongly disagree

What was the reason that you couldn't be honest?

What is your level of trust in the noise-adding mechanism presented before in terms of its privacy preserving capability?

Very High
High
Medium
Low
Very Low

What are the reasons for you to feel less trustworthy?

How confident are you that you correctly added noise to your answers?

Very High
High
Medium
Low
Very Low

Condition Group 2

Imagine you have used recreational drugs in the past year. One day, you saw a paid survey on an online forum for research purposes. You were interested in earning some money and wanted to fill it out. Then you noticed that there were some sensitive questions related to your experience of using recreational drugs.

Consider the following definition of **LDP:**

Local Differential Privacy or LDP protects users' privacy by adding random noise to each response that users give and provides a statistical guarantee of privacy. So, no one can know, just by looking at the overall data, what a specific user's response is, or even if a user was part of the data collection. LDP has been used by companies such as Google and Apple to collect information from users in a privacy-preserving manner.

We are going to use the LDP technique here – random noise will be added automatically once you choose an option and the noise-added answer will be stored in the server.

Please answer the question below, imagining what you would do if you had actually used recreational drugs in the past year and were presented with a paid survey online for research purposes that uses the LDP technique described above.

Have you used recreational drugs in the past 1 year?

Yes No

To what extent do you agree or disagree with the below statements about the noise-adding mechanism used in the above question:

The noise addition guarantees the secrecy of my sensitive data (whether I used recreational drugs in the past year or not).

Strongly Agree
Somewhat agree
Neither agree nor disagree
Somewhat disagree
Strongly disagree

The noise addition guarantees that my sensitive data (whether I used recreational drugs in the past year or not) will not be revealed, even if the server's data is disclosed because of an attack.

Strongly Agree
Somewhat agree
Neither agree nor disagree
Somewhat disagree
Strongly disagree

How would you rate your level of understanding of the noise addition mechanism?

Very High
High
Medium
Low
Very Low

To what extent do you agree or disagree with this statement:

The prior description of the privacy protection technique was easy for me to understand.

Strongly Agree
Somewhat agree
Neither agree nor disagree
Somewhat disagree
Strongly disagree

To what extent do you agree or disagree with this statement:

I expect the noise addition technique to provide benefits.

Strongly Agree
Somewhat agree
Neither agree nor disagree
Somewhat disagree
Strongly disagree

What benefit(s) do you expect/see from adding noise based on the explanation?

What was your comfort level when taking this survey and being asked sensitive questions in a hypothetical scenario?

Very High
High
Medium
Low
Very Low

What was the reason for you to feel discomfort?

If this had been a real situation and you were asked to disclose your recreational drug use (or similar sensitive information) within the same online survey interface, how much do you agree that you would be honest in answering this question based on the noise-addition technique?

Strongly agree
Somewhat agree
Neither agree nor disagree
Somewhat disagree
Strongly disagree

What were the reason(s) that you couldn't be honest?

What is your level of trust in the noise-adding mechanism presented before in terms of its privacy preserving capability?

Very High
High
Medium
Low
Very Low

What are the reason(s) you felt less trustworthy of the mechanism?

Condition Group 1

Imagine you have used recreational drugs in the past year. One day, you saw a paid survey on an online forum for research purposes. You were interested in earning some money and wanted to fill it out. Then you noticed that there were some sensitive questions related to your experience of using recreational drugs.

Please answer the question below, imagining what you would do if you had actually used recreational drugs in the past year and were presented with a paid survey online for research purposes.

Have you used recreational drugs in the past 1 year?

Yes No

What was your comfort level when taking a survey and being asked sensitive questions in a hypothetical scenario?

Very High
High
Medium
Low
Very Low

What were the reason(s) for you to feel discomfort?

If this had been a real situation and you were asked to disclose your recreational drug use (or similar sensitive information) within the same online survey interface, how much do you agree that you would be honest in answering this question?

Strongly agree
Somewhat agree
Neither agree nor disagree
Somewhat disagree
Strongly disagree

What were the reason(s) that you couldn't be honest?

Demographics

What is your age range?

18-22
23-27
28-32
33-37
38-42
43-47
48-52
53 or older

What is your gender?

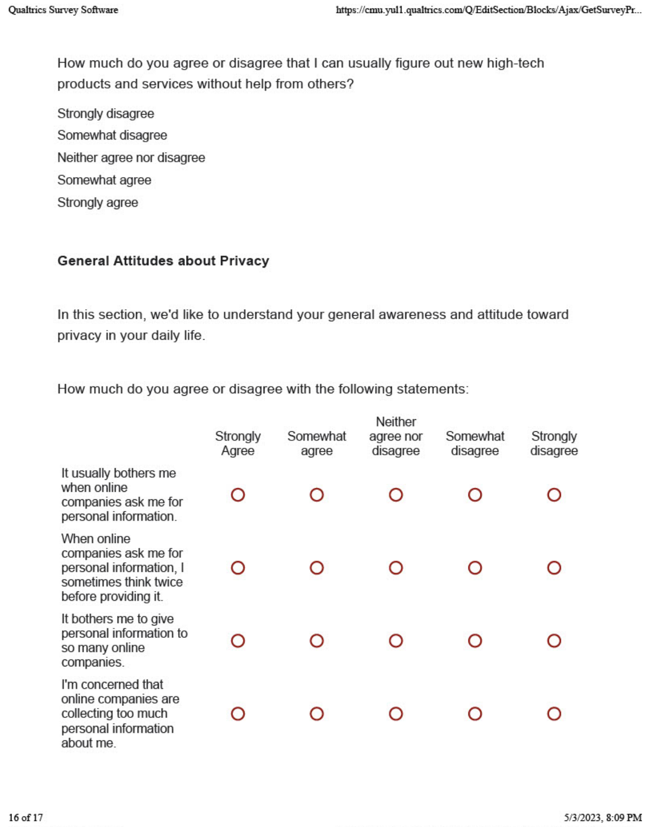
Male
Female
Prefer not to disclose
Prefer to self-describe

What is your highest education level?

Middle School or lower
Partial completion of high school
High School
Associate's Degree
Bachelor's Degree
Graduate Degree

General Attitudes about Technology

In this section, we'd like to understand your general understanding of technology in your daily life.



Qualtrics Survey Software https://cmu.yul1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPr...

How much do you agree or disagree that I can usually figure out new high-tech products and services without help from others?

Strongly disagree

Somewhat disagree

Neither agree nor disagree

Somewhat agree

Strongly agree

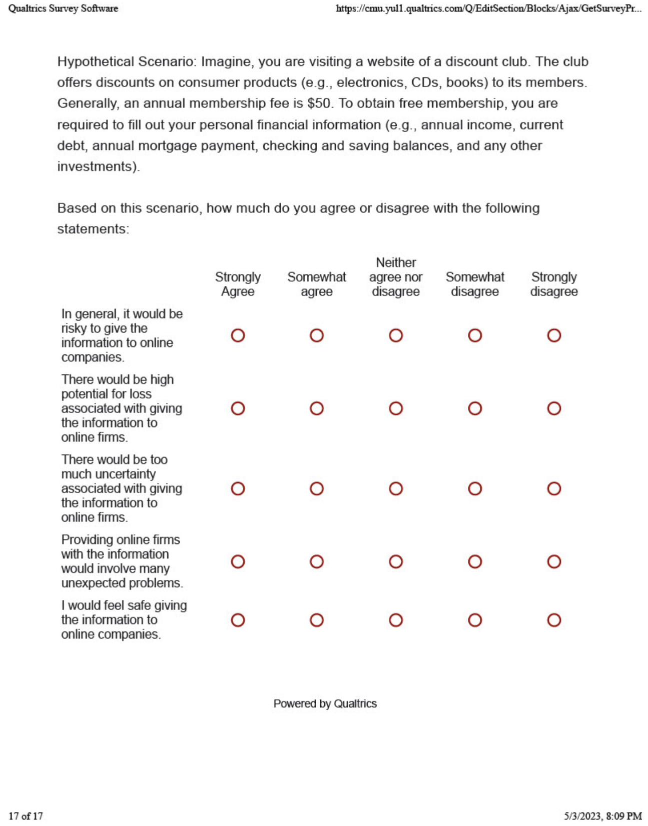
General Attitudes about Privacy

In this section, we'd like to understand your general awareness and attitude toward privacy in your daily life.

How much do you agree or disagree with the following statements:

	Strongly Agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
It usually bothers me when online companies ask me for personal information.	○	○	○	○	○
When online companies ask me for personal information, I sometimes think twice before providing it.	○	○	○	○	○
It bothers me to give personal information to so many online companies.	○	○	○	○	○
I'm concerned that online companies are collecting too much personal information about me.	○	○	○	○	○

16 of 17 5/3/2023, 8:09 PM



Qualtrics Survey Software https://cmu.yul1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPr...

Hypothetical Scenario: Imagine, you are visiting a website of a discount club. The club offers discounts on consumer products (e.g., electronics, CDs, books) to its members. Generally, an annual membership fee is $50. To obtain free membership, you are required to fill out your personal financial information (e.g., annual income, current debt, annual mortgage payment, checking and saving balances, and any other investments).

Based on this scenario, how much do you agree or disagree with the following statements:

	Strongly Agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
In general, it would be risky to give the information to online companies.	○	○	○	○	○
There would be high potential for loss associated with giving the information to online firms.	○	○	○	○	○
There would be too much uncertainty associated with giving the information to online firms.	○	○	○	○	○
Providing online firms with the information would involve many unexpected problems.	○	○	○	○	○
I would feel safe giving the information to online companies.	○	○	○	○	○

Powered by Qualtrics

17 of 17 5/3/2023, 8:09 PM

8.2 Appendix 2: Codebook and results

Here is the codebook we used for the open-response questions. Each answer may be coded in more than one category.
Understanding of the benefits of Differential Privacy:
Definitions:

1. Privacy - the participant named privacy as one of the main benefits they expect from the mechanism.

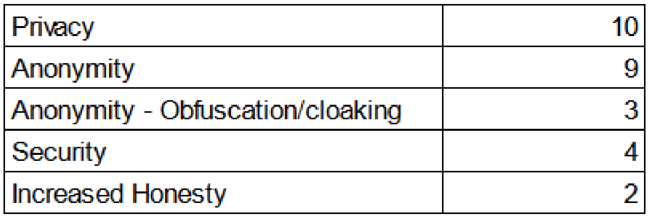
2. Anonymity - the participant named anonymity as one of the main benefits they expect from the mechanism.

3. Anonymity- Obfuscation/Cloaking - the participant named anonymity as one of the main benefits they expect from the mechanism, and specifically mentioned obfuscation or cloaking as the reason for anonymity.

4. Security - the participant named security as one of the main benefits they expect from the mechanism.

5. Increased Honesty - the participant named increased honesty when answering sensitive questions as one of the main benefits they expect from the mechanism.

Count:



Privacy	10
Anonymity	9
Anonymity - Obfuscation/cloaking	3
Security	4
Increased Honesty	2

Reason for discomfort:
Definitions:

1. Disclosure of Sensitive Data - The user named the possible disclosure of sensitive data as a reason for their discomfort.

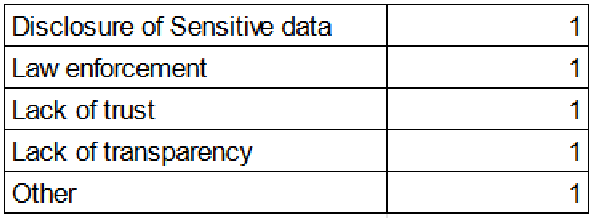
2. Law enforcement - The participant named the possible involvement of law enforcement with the survey (due to questioning about drug usage) as a reason for their discomfort.

3. Lack of trust - The participant named their lack of trust in the mechanism as a reason for their discomfort.

4. Lack of transparency - The participant named the lack of transparency of the mechanism as a reason for their discomfort.

5. Other - Response not fitting in any of the aforementioned categories.

Count:

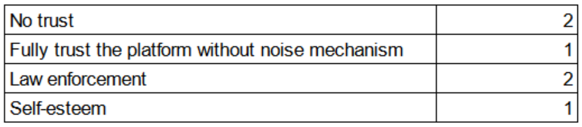


Disclosure of Sensitive data	1
Law enforcement	1
Lack of trust	1
Lack of transparency	1
Other	1

Reason for dishonesty:
Definitions:

1. No trust: The participant named their lack of trust in the mechanism as a reason for being dishonest.
2. Fully trust the platform without noise mechanism: The participant thought that the differential privacy mechanism was unnecessary in the face of other mechanisms used by the survey platform to maintain anonymity.
3. Law Enforcement: The participant was afraid of potential disclosure of data to law enforcement
4. Self-esteem: The participant did not want to admit to taking drugs for the sake of their self-esteem.

Count:

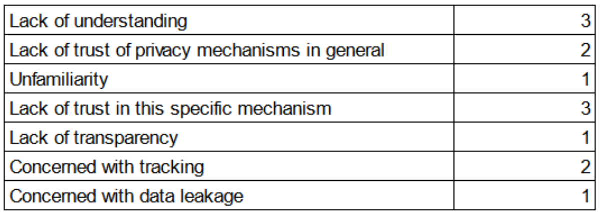


No trust	2
Fully trust the platform without noise mechanism	1
Law enforcement	2
Self-esteem	1

Reason for lack of trustworthiness:

1. Lack of understanding: The participants admitted to not fully understanding the workings of the mechanism.
2. Lack of trust in privacy mechanisms in general: The participants do not trust privacy-enhancing mechanisms in general.
3. Unfamiliarity: The participant was unfamiliar with the technique used, and therefore did not trust it.
4. Lack of trust in this specific mechanism: The participant did not trust this mechanism's capacity of providing privacy.
5. Lack of transparency: The participant did not trust this mechanism due to a lack of transparency of its inner workings.
6. Concerned with tracking: The participant was concerned with being tracked while using the internet.
7. Concerned with data leakage: The participant was concerned with their data being leaked.

Count:



Lack of understanding	3
Lack of trust of privacy mechanisms in general	2
Unfamiliarity	1
Lack of trust in this specific mechanism	3
Lack of transparency	1
Concerned with tracking	2
Concerned with data leakage	1