Protecting personal health information: The roles of context, framing and priming in privacy-related choices

Vanessa Boothroyd
Privacy Analytics, Inc.
251 Laurier Ave W
Ottawa, ON
vboothroyd at privacy-analytics dot com

Ester Moher
Children's Hospital of Eastern Ontario
401 Smyth Rd
Ottawa, ON
emoher at ehealthinformation dot ca

Khaled El Emam
Children's Hospital of Eastern Ontario; Privacy Analytics, Inc.
251 Laurier Ave W
Ottawa, ON
kelemam at privacy-analytics dot com

1. INTRODUCTION

When making privacy-based decisions, such as whether to share data or how to protect it, it has traditionally been assumed that decision making reflects fixed personal preferences. However, people do not always have well-defined preferences [1]–[4]. Instead, preferences are often constructed as they are elicited [2], which may explain inconsistencies observed in the literature.

One such method of influencing decision making involves wording of the question. For example, when asked to *choose* between two options, participants tend to select the option with the strongest positive features; however, when asked to *reject* one of two options, participants tend to select (reject) the option with the strongest negative features. The option people reject and choose could well be the same [2].

We propose these types of context-driven biases affecting simple choices (framing and priming) may also be at play in complex decision making, such as in the context of making privacy-related decisions. We examine whether participants' choices regarding anonymization of health data are influenced by a framing manipulation, a priming manipulation, and providing additional scenario-related contextual information.

2. STUDY 1: FRAMING

We recruited participants (N=403) from Crowdflower. Participants read a scenario about a hospital looking to anonymize health data for the purpose of sharing data secondarily for research. Participants were asked to decide which of two anonymization software options they would prefer to use as the hospital representative. Software options were presented either in enriched format (benefits and drawbacks of the software were described in extreme terms, such as "very expensive/inexpensive") or in neutral format ("can be costly/reasonably priced"), in counterbalanced fashion. We presented two software options, one which was objectively more secure than the other.

Half of participants were asked which software they would choose to use (choice condition), and half of participants were asked which software they would prefer not to use (reject condition). Participants were also asked to justify their responses.

Of the main sample of 403 participants, n=24 were excluded because they either did not complete the survey, were inconsistent in justifying the option they selected, or misunderstood whether a "low probability of the patient being identified" was good or bad. N=379 remained, and are included in analyses.

When software selection decision was framed as a choice, there was a significant effect of enrichment on software preference, χ^2 (1, 189) = 12.14, $p < .01$, such that the more secure software was preferred more often when it was framed neutrally. However, there was no evidence of an effect of enrichment on software preference in the Reject condition, χ^2 (1, 190) < 1, $p > .50$. Thus, we do observe effects of choice framing, but these effects are inconsistent with those in the literature [2].

Findings of this study suggest that by asking participants to choose, thereby focusing their attention on software benefits, we may be able to encourage more secure software selections. Particularly, by adding weight to negative features through enriched wording, we may encourage participants to focus less on privacy benefits, and more on potential consequences that could impact the overall goal described in the scenario, such as poor data quality for cancer research.

Participants in this study were likely to mention privacy concerns more than half of the time (63%); it was unknown as to whether this was due to the description of the task (i.e., acting as a decision maker to protect patient data), or whether this reflected a baseline of individuals who are inherently concerned with privacy. In Study 2, we explicitly include a description of privacy concerns for half of the participants to examine this hypothesis more thoroughly.

3. STUDY 2: PRIMING

We recruited participants (N=419) from Crowdflower. Again, we asked participants to make a software decision on behalf of a hospital looking to share data. We varied whether we included a description of privacy concerns in the problem scenario. That is, half of participants saw a scenario with a specific description of protecting patients within the data, whereas half of participants read the scenario without such a description.

Overall, n=21 participants who were confused about the meaning of identification were removed from analysis. N= 388 remained, and are included in analyses.

We did not observe any significant effects of priming, χ^2 (1, 381) < 1, $p > .80$. Results of Study 2 suggest that priming patient protection may not motivate more privacy-focused decision making—perhaps because protection concerns were already quite salient. Alternatively, it is possible that participants considered not just which software was more secure, but whether the software offered any protections, legal or otherwise, in the event of a breach or attack on data. We examine this hypothesis in Study 3,

along with whether adding valenced contextual information could influence software choice.

4. STUDY 3: CONTEXT

We recruited participants (N=428) from Crowdflower. We asked participants to make software decisions on behalf of a hospital looking to share data. Unique to Study 3, we offered the chance to defer the choice. Participants were presented with one of three scenarios and one of two different anonymization software presentations: The scenario included a section providing context about the benefits of sharing data (positive condition), a section providing context about the risks of data breach (negative condition), or no additions (control condition).

Overall, N = 43 participants were removed for not completing the task, completing selection with an inconsistent response, or completing the survey multiple times. N = 385 remained, and are included in analyses.

There was no significant association between the three different valence conditions (positive, negative, neutral) on software choice, $\chi^2(4, 385) = 7.30$, $p > .10$. However, in a similar study on the effects of risk framing, Botzen et al [7] found that merely providing additional information, regardless of valence, increased participants' willingness to purchase flood insurance. Analyses support this assertion: the addition of positively *or* negatively valenced contextual information encouraged choosing the more secure software option more often relative to the control, $\chi^2(2, 385) = 5.96$, $p = .05$. We hypothesize that the provision of additional contextual information included in the problem description may have driven this effect, similar to [7]. However, given the nature of this analysis, more research is necessary to support these findings.

5. DISCUSSION

Together, these studies suggest that even in a hypothetical scenario, privacy-based decision making is highly complex, and may not be subject to the same systematic heuristics and biases of more simple choices (see [2]). Our results illustrate that participants were not influenced as expected by enrichment framing effects or priming effects. We do see that overall, the more secure option was chosen more frequently, but this choice hovered between 60% and 90%, indicating a tension between privacy concerns and other important features, such as data quality. These findings demonstrate that deciding between security and data quality is not trivial, particularly when the goal of sharing data is for cancer research, which is data-quality-dependent (as many of our participants noted). However, providing added contextual information about a privacy-related scenario can increase individuals' selection of more a more secure option, assisting in this difficult decision process.

6. REFERENCES

[1] D. Ariely, G. Loewenstein, and D. Prelec, "'Coherent Arbitrariness': Stable Demand Curves Without Stable Preferences," *The Quarterly Journal of Economics*, 2003.

[2] E. Shafir, "Choosing versus rejecting: Why some options are both better and worse than others," *Memory & Cognition*, vol. 21, no. 4, pp. 546–556, 1993.

[3] P. Slovic, "The construction of preference," *American Psychologist*, vol. 50, no. 5, pp. 364–371, 1995.

[4] A. Tversky and D. Kahneman, "Rational Choice and the Framing of Decisions," *The Journal of Business*, vol. 59, no. 4, Part 2, pp. S251–S278, 1986.

[5] K. El Emam and L. Arbuckle, *Anonymizing Health Data: Case Studies and Methods to Get You Started*. O'Reilly, 2013.

[6] H. Kunreuther and P. Slovic, "Science, Values, and Risk," *Annals of the American Academy of Political and Social Science*, vol. 545, pp. 116–125, 1996.

[7] W. J. Wouter Botzen, J. de Boer, and T. Terpstra, "Framing of risk and preferences for annual and multi-year flood insurance," *Journal of Economic Psychology*, vol. 39, pp. 357–375, 2013.

[8] D. Kahneman and A. Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, vol. 47, no. 2, pp. 263–292, 1979.

[9] G. F. Loewenstein, E. U. Weber, C. K. Hsee, and N. Welch, "Risk as Feelings," *Psychological Bulletin*, vol. 127, no. 2, pp. 267–286, 2001.

[10] G. D. Logan, "Repetition priming and automaticity: Common underlying mechanisms?," *Cognitive Psychology*, vol. 22, no. 1, pp. 1–35, 1990.

[11] J. H. Neely, "Semantic priming effects in visual word recognition: A selective review of current findings and theories," in *Basic Processes in Reading: Visual Word Recognition*, D. Besner and G. W. Humphreys, Eds. Psychology Press, 1991, p. 350.