Data collection explicitness as a micro-suasor: Its effect on sensitivity judgment and safety decisions

MariaVittoria Masotina[1], Patrik Pluchino[1], Francesca Freuli[1], Luciano Gamberini[1,2], and Anna Spagnolli[1,2]

[1] Human Inspired Technology Research Centre, University of Padova, Italy
[2] Dept. of General Psychology, University of Padova, Italy

mariavittoria.masotina@gmail.com
patrik.pluchino@unipd.it
francesca.freuli@studenti.unipd.it
luciano.gamberini@unipd.it
anna.spagnolli@unipd.it

Abstract. It has been found that transparency implemented in the name of the users' safety can instead encourage the users to trust the system and to disclose their personal data. In this work we consider whether the transparency of the data collection technique can work in this way. The study (N = 40) compares an explicit technique (questionnaires) with an implicit one (eye-tracker). The actual sensitivity of the data collected was also varied, sensitive (popularity) vs. non-sensitive (usability). The results suggest that, when judging general data sensitivity, the transparency of the data collection procedure tends to work as a heuristic; this is not the case when more specific judgments or decisions are asked.

1 Introduction

Improving the transparency of a system or service collecting personal data is recommended as a means to increases users' ability to protect their identity and increase their safety (e.g. EU GDPR 2016/679). Paradoxically, however, transparency might backfire: users are reported to take transparency as a cue to quickly decide that a system can be trusted [1, 3] and can then disclose their personal data. Transparency might then work as a micro-suasor [2]. In the present study we focus on the transparency of the data collection technique, comparing an explicit technique (questionnaires) with an implicit one (eye-tracker), and measure its effect on sensitivity judgment (perceived data sensitivity) and safety decisions (waving data anonymity). We also varied the actual sensitivity of the data collected, which was either sensitive (usability) or non-sensitive (popularity). We then had a 2x2 between-participant design. Condition assignment was randomized.

2 Methods

Sample and procedure. 40 university students enrolled in the Psychology School of the University of Padova participated in this study (mean age 23.61, SD = 1.71, men = 9, women = 31). They visited four of their teachers' websites and, according to the study condition, evaluated their usability or popularity via a questionnaire or via an eye tracker detecting their visual behavior. Then they were sent to a Google Form questionnaire asking on 5-point scale: (a) if they consented to *wave the anonymity* of their opinions ("Would you be willing to let us process your data renouncing anonymity, so we can associate your name and surname to the data and responses collected during the whole experiment?") and (b) to *evaluate the perceived sensitivity* of the data provided ("Do you think that the data collected during this experiment is sensitive, namely that it could identify you in a counterproductive manner?"; "Do you think that the information derived from such data could be embarrassing for you?"; "Do you think that the data provided could be offensive to the teachers?"). The whole procedure was automated via Atom software. The informed consent to participate was signed before the session, while the consent to use the data was signed after debriefing.

3 Results

The dependent variables were the frequency with participants accepted to wave the anonymity of their data, and the perceived sensitivity of the data; the values are reported in Figure 1.

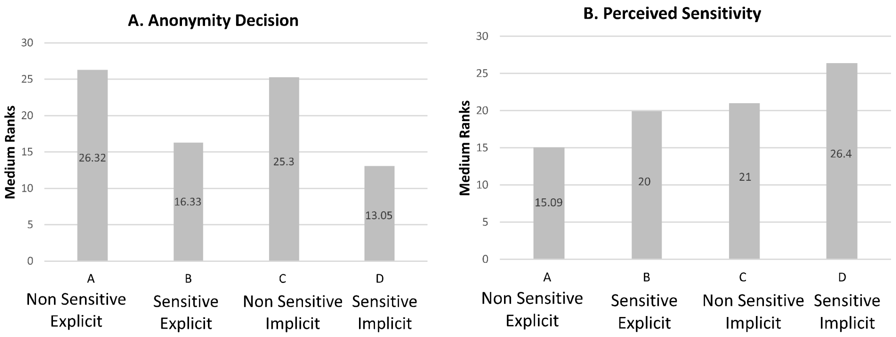

Fig. 1a, 1b. Medium ranks of the willingness to wave anonymity (a, left) and of the perceived sensitivity of the data (b, right) broken down by condition.

To assess the effect of the two manipulated factors (type of data collected and explicitness of the collection technique) a Mann-Whitney test was run (Table 1).

Table 1. Results of the Mann-Whitney test measuring the effects of the two factors, the type of data (usability vs. popularity) and the explicitness of the data collection technique.

	TYPE OF DATA		EXPLICITNESS	
	W	*p*	W	*p*
Anonymity decision	311.5	0.002	226.5	0.47
Perceived sensitivity (construct)	145	0.12	136	0.07
Item 1: Generic sensitivity	198	0.97	144	0.08
Item 2: Embarrassment	156	0.10	164	0.17
Item 3: Offensiveness	123	0.01	208	0.81

The results reported in Table 1 suggest that, when judging general data sensitivity (Item1), the explicitness of the data collection procedure tends to work as a heuristic, decreasing the perceived sensitivity of the data regardless of its actual content. Instead, safety decisions such as waving anonymity as well as more specific sensitivity judgments such as the one expressed by Item 3 are more influenced by the actual content of the data. In other words, what can make the difference is the clarity of the scenario in which the user is able to figure the possible risks. This hypothesis will be pursued in further studies and is surely to be taken into account when measuring sensitivity with self-reported methods.

References

1. Acquisti, A., Adjerid, I., Brandimarte, L: Gone in 15 seconds: The limits of privacy transparency and control. IEEE Security & Privacy, 11(4), 72-74 (2013)
2. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. San Francisco: Morgan Kaufmann (2003)
3. Oulasvirta, A., Suomalainen, T., Hamari, J., Lampinen, A., Karvonen, K.: Transparency of intentions decreases privacy concerns in ubiquitous surveillance. Cyberpsychology, Behavior, and Social Networking, 17(10), 633-638 (2014)