



**HAL**  
open science

# Transparency Heuristic: Effect of Implicitness of Online Data Acquisition on Sensitivity Perception

Mariavittoria Masotina, Patrik Pluchino, Francesca Freuli, Luciano Gamberini, Anna Spagnolli

► **To cite this version:**

Mariavittoria Masotina, Patrik Pluchino, Francesca Freuli, Luciano Gamberini, Anna Spagnolli. Transparency Heuristic: Effect of Implicitness of Online Data Acquisition on Sensitivity Perception. 17th IFIP Conference on Human-Computer Interaction (INTERACT), Sep 2019, Paphos, Cyprus. pp.676-679, 10.1007/978-3-030-29390-1\_61 . hal-02878652

**HAL Id: hal-02878652**

**<https://inria.hal.science/hal-02878652>**

Submitted on 23 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Transparency heuristic: Effect of implicitness of online data acquisition on sensitivity perception

Mariavittoria Masotina<sup>1</sup>, Patrik Pluchino<sup>1</sup>, Francesca Freuli<sup>1</sup>, Luciano Gamberini<sup>1,2</sup>  
and Anna Spagnolli<sup>1,2</sup>✉

<sup>1</sup> Human Inspired Technology Research Centre, University of Padova, Italy

<sup>2</sup> Dept. of General Psychology, University of Padova, Italy  
anna.spagnolli@unipd.it

**Abstract.** We present a study that investigates whether the transparency of the data acquisition technique can work as a heuristic when making evaluations about data protection and sensitivity. The study (N = 40) compares an explicit data acquisition technique (questionnaires) with an implicit one (eye-tracker) and varies also the actual sensitivity of the data collected (popularity evaluation vs. usability evaluation). The results suggest that, when judging general data sensitivity, the transparency of the data collection procedure might work as a heuristic; instead if more specific judgments or decisions are asked this effect is not observed. Implications are discussed.

**Keywords:** Sensitivity Perception, Anonymity, Transparency.

## 1 Introduction

Improving the transparency of a system or service collecting personal data is recommended as a way to increase users' ability to protect their identity (e.g. EU GDPR 2016/679). Paradoxically, however, transparency might backfire: users can interpret transparency as a cue to quickly decide that a system can be trusted [1, 2] and then disclose their personal data. Transparency would then work as a heuristic. In the present study we focus on the transparency of the data acquisition method, and check whether it would affect users' perception and decisions in the domain of data protection regardless of the actual sensitivity of the data acquired. To this goal, we had a sample of students participating in our study; their data were acquired either with a transparent technique (questionnaires), in which the act of providing data is a voluntary and conscious one in the users' side for each datum provided, or with an implicit one (eye-tracker), in which after the initial consent from the user the system takes care of the data collection and the user is unaware of each datum collected. The data collected could be either sensitive or non-sensitive, namely potentially damaging to the user or not; in the former case students had to rate their professors' popularity, in the latter case they had to rate the usability of the professors' institutional webpage. The effect of this manipulation on their judgment (perceived sensitivity of the collected data) and decisions (willingness to keep their anonymity) was then measured.

## 2 Method

The experiment followed a 2 X 2 between-participants design; the independent variables manipulated were the sensitivity of the data acquired (non-sensitive vs. sensitive) and the transparency of the data acquisition method (transparent vs. implicit). Our research questions were whether the two manipulated variables had any effect on participants' decision to keep anonymity and on the perceived sensitivity of the data.

The procedure was double-blind; neither the participants nor the member of the research team meeting the participants were aware of the goal of the study. Participants were randomly assigned to the different conditions. The whole procedure was automated via Atom software (<https://atom.io/>) so that there was no need for the researcher to provide instructions or intervene till the end of the session. A Tobii 1750 eye-tracker was used in the implicit conditions, while a tablet was used in the transparent conditions. Google Form was used to collect the participants' responses.

Before the beginning of the experimental session, every participant read and signed an informed consent to participate in the study. The study was described as one investigating professors' webpages usability or professors' popularity, depending on the sensitivity condition. In the implicit conditions the technique used to acquire their usability/popularity rates was described as based on the detection of fatigue/repulsion via pupillary response collected with an eyetracker. In the transparent condition the technique to acquire usability/popularity rates was a questionnaire. It was explained them that the acquired data would be used in aggregate form to preserve participants' anonymity. However, participants' e-mail address was asked in the questionnaire administered at the end of the experiment, justified by the need to contact them at the end of the study to obtain their permission to use the collected data.

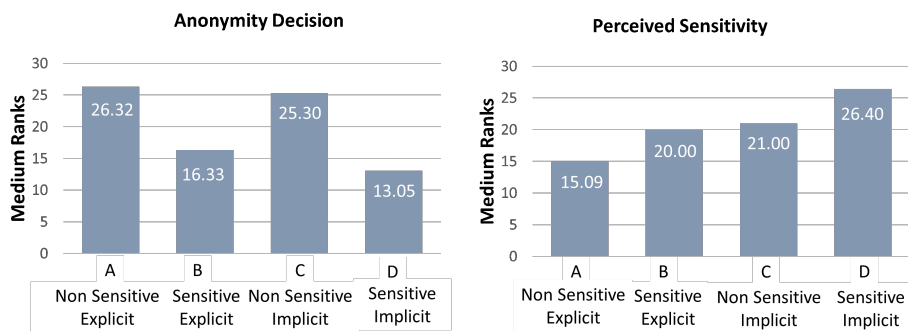
The experimental session then began. Participants were sitting at a desk on which a computer was positioned. First they were asked to provide some demographic data and the name of their master or bachelor program. In the implicit conditions, the eye-tracker was activated and calibrated. Then, the researcher launched the application administering the experimental task and left the room. Participants saw displayed on the screen a list of all teachers in their master or bachelor program, and were asked to select the one they knew better. The webpage of the selected professor would then appear and be briefly inspected by the participant. In the implicit conditions, a message appeared on the screen notifying that the eye-tracker had successfully captured their reactions. In the transparent conditions participants were asked to rate the usability of the web page or the popularity of the professor in a short questionnaire. This task was repeated four times until the webpages of four professors (two men and two women) were visited. Afterwards, participants were automatically sent to a Google Form questionnaire asking on a 5-point Likert scale: (a) if they consented to waive the anonymity of their data ("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 (Item 1- Do you think that the data collected during this experiment is sensitive, namely that it could identify you in a counterproductive manner?; Item 2 - Do you think that the information derived from such data could be embar-

rating for you?; Item 3 - Do you think that the data provided could offend the teachers?).

**Sample.** 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 were all Italian native speakers. The webpage displayed in the study were the institutional webpage of each professor.

### 3 Results

The participants' willingness to waive anonymity and their perceived sensitivity of the collected data are reported in Figure 1.



**Fig. 1 a,b.** Medium ranks of the willingness to waive 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 transparency of the collection technique) on the dependent variables, 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.

	DATA SENSITIVITY		TRANSPARENCY	
	W	<i>p</i>	W	<i>p</i>
Anonymity decision	311.5	0.002	226.5	0.47
Sensitivity Item 1 (sensitive)	198	0.97	144	0.08
Sensitivity Item 2 (embarrassing)	156	0.10	164	0.17
Sensitivity Item 3 (offensive)	123	0.01	208	0.81
Consolidated score from items 1,2,3	145	0.12	136	0.07

The results show that the willingness to waive anonymity was affected only by the actual sensitivity of the data acquired and not by the transparency of the data collection method: frequency is significantly higher in the non-sensitive conditions (usability rates) than in the sensitive ones (popularity rates). Regarding the perceived sensitivity of the data collected, there seems to be an effect of the transparency of the data collection method in the consolidated scores of the three items and in Item n. 1 with a tendency to statistical significance. The other two items did not seem affected by the data collection method. There seem also to be a cumulative effect so that the highest difference in perceived sensitivity at a Mann Whitney test was between the two most extreme conditions in which the manipulated factors are both present or both absent,  $W = 26.5$ ,  $p\text{-value} = 0.037$ .

#### 4 Conclusions and future work

The results suggest that, when judging sensitivity in generic terms the transparency of the data collection procedure might be relied upon as a heuristic, decreasing the perceived sensitivity of the data regardless of its actual content. Transparency does not seem to be relied upon, instead, when specific sensitivity judgments are at stake. Likewise, the decision whether to waive the anonymity of the data is not affected by the data collection technique. What seems to make the difference across all these results is the clarity of the scenario in which the user is able to figure the possible risks. The more specific the question, the easier it is for them to estimate the actual risks and avoid recourse to heuristics. Studies seem to confirm that sensitivity is a contextual estimation (e.g.,[3]), connected to the possible detrimental uses of the data [4]. This hypothesis will be pursued in further studies by varying the specificity of the decision and of the judgment. To scholars in human-computer interaction, these initial results renew the challenge of finding a genuine way to implement transparency in a way that avoids backlashes. At the same time, it suggests to adopt compensatory measures when using implicit/background data techniques that per se might reduce trust in users.

#### 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. 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).
3. Sun, Y., Wang, N., Shen, X. L., & Zhang, J. X. (2015). Location information disclosure in location-based social network services: Privacy calculus, benefit structure, and gender differences. *Computers in Human Behavior*, 52, 278-292.
4. Dinev, T., & Hart, P. (2004). Internet privacy concerns and their antecedents-measurement validity and a regression model. *Behaviour & Information Technology*, 23(6), 413-422.