Securing the smart home environment: an experiment on the impact of explainable warnings

Angela Martone
University of Turin
Italy
angela.martone@edu.unito.it

Federica Cena
Dept. of Computer Science, University of Turin
Italy
federica.cena@unito.it

Cristina Gena
Dept. of Computer Science, University of Turin
Italy
cristina.gena@unito.it

Fabiana Vernero
Dept. of Computer Science, University of Turin
Italy
fabiana.vernero@unito.it

ABSTRACT

In this paper, we present an experiment where we study the impact of different types of explainable security warnings in a smart home environment. Results show that detailed informal-style explanations are evaluated more positively and are more effective in promoting safe home automation.

CCS CONCEPTS

• **Security and privacy → Usability in security and privacy**; • **Human-centered computing → Empirical studies in HCI**.

KEYWORDS

End-user development, trigger-action rules, privacy and security, explainability

ACM Reference Format:
Angela Martone, Federica Cena, Cristina Gena, and Fabiana Vernero. 2024. Securing the smart home environment: an experiment on the impact of explainable warnings. In *International Conference on Advanced Visual Interfaces 2024 (AVI 2024), June 03–07, 2024, Arenzano, Genoa, Italy*. ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3656650.3656721

1 INTRODUCTION

The Internet of Things has significantly impacted home automation [8], allowing users to monitor and remotely control several aspects of their domestic environment, such as the heating or lighting [11]. Research in the End-User Development (EUD) domain has shown that Event-Condition-Action (ECA) rules [4, 9] (consisting of a triggering event, the conditions to be met when the event occurs, and an action to be carried out as a consequence) represent a useful approach to allow non-expert users to specify the desired behaviour of a smart environment. In fact, they can be expressed at a high level of abstraction as conditional statements (e.g., *If the room temperature falls below 19°C, turn on the heating system*), with no need to delve into complex technical details [7]. Trigger–action Platforms (TAPs), such as If-This-Then-That (IFTTT)[1], provide an intuitive interface to specify such rules.

However, users lacking technical knowledge may inadvertently compose or re-use rules which expose them to potentially harmful scenarios [1, 3, 5, 6]. For example, if a smart home is programmed to automatically turn off lights when the last resident exits, it might be easier for an attacker to understand when the house is empty and plan an intrusion. More specifically, Surbatovich et al. [10] identified four categories of potential damage: *innocuous* (no harm), *personal* (loss of sensitive data), *physical* (damage to physical health or goods), and *cybersecurity* (interruptions in online services or distribution of malware). A promising approach to protect the security of a smart home environment and the privacy of its users is to automatically identify potentially dangerous ECA rules (see e.g., [2]) and to either warn users against their application or simply prevent their usage. However, if warnings are provided, it must be considered that their specific features can impact users' perceptions and, consequently, their acceptance of the given advice.

In this paper, we present an experiment where we study the impact of warnings which differ for their level of explainability and use of technical/informal language. In particular, we want to observe their effects on several dimensions related to users' experience and on users' willingness to apply ECA rules which might entail privacy or security risks in their smart home environment.

2 EXPERIMENT

In this experiment, we had 16 participants interact with a simulated visual TAP where they could browse 6 ECA rules (ideally, rules they had composed beforehand) and decide whether to apply them or not. The interface displayed a warning for potentially harmful rules, and participants could decide whether to confirm or change their initial choice. Rules were formulated taking inspiration from [2]: an example is *Open the shutters if the temperature is above 25°C*. For each rule, two types of warnings were available: a concise label making use of technical language (for the previous rule: "Physical damage"), and a more detailed informal-style explanation (for the previous rule: "Automatically opening windows could make it easy for an attacker to plan an intrusion").

AVI 2024, June 03–07, 2024, Arenzano, Genoa, Italy

ACM ISBN 979-8-4007-1764-2/24/06
https://doi.org/10.1145/3656650.3656721

[1] https://ifttt.com/

2.1 Methodology

Design. We used a between-subjects design, where the independent variable is the *type of warning*, with two possible values, either a concise label (group A) or a label with an explainable description (group B). The dependent variables are participants' assessments about the warnings and the number of times they changed their mind on rule application after reading the security warnings.

Hypothesis. We hypothesized that warnings consisting of a label with an explainable description would obtain higher evaluations and would have a stronger impact on participants' choices.

Measures. Participants' evaluations were collected by asking them to express their level of agreement with four statements regarding various aspects of the warnings (usefulness, understandability, adequacy of their level of detail, perceived persuasiveness) using 5-point Likert scales ranging from "Disagree strongly (1)" to "Agree strongly (5)". Participants' choices about rule application were manually recorded by the experimenter.

Material. Two online questionnaires were used to collect participants' demographics and their evaluations. An interactive prototype simulating the interface of a TAP was build using Figma[2].

Participants. We recruited 16 participants, 76,5% females and 23,5% males, aged 18-74, through a convenience sampling strategy. Most of them (58,8%) were aged 18-24.

Procedure. One at a time, participants were welcomed by the experimenter and introduced to the experimental procedure. Firstly, they had to fill in a questionnaire aimed at collecting their demographics. Then, they were allowed to freely explore the TAP prototype, to familiarize themselves with the smart home scenario and with the idea of controlling the behaviour of a number of connected devices to carry out everyday tasks. After that, they were asked to access a specific section where they could browse a series of ECA rules. For each rule, they were asked to: 1) decide whether they would apply it or not; 2) ask the system to apply the rule; 3) observe the system response (which might imply the display of a safety warning); and 4) decide whether they would confirm their initial choice or not. They were also encouraged to express their impressions using the *thinking aloud* technique. Finally, they were asked to fill in a second questionnaire to evaluate the warnings.

2.2 Results and discussion

The average values for participants' evaluations of the warnings are reported in Figure 1. Consistently with our hypothesis, we can observe that warnings consisting of both a label and an explainable description (namely, those received by participants in group B) were assessed more positively for all four aspects. The largest difference can be observed for *understandability* (4,875 vs. 2,875), indicating that a concise label alone, although it identifies the type of damage participants might be exposed to, is not enough to clearly appreciate the risks entailed by a certain ECA rule. Participants' comments collected through thinking aloud confirm this impression (U3A: *The label itself is not very clear, I need an example to understand why I should not apply this rule*; U8B: *The description is very important to understand the actual risks -the label alone could be ambiguous*). Understandability is also the only aspect where warnings (in this case, group A) received an unsatisfying evaluation (i.e., lower than

[2]https://www.figma.com/

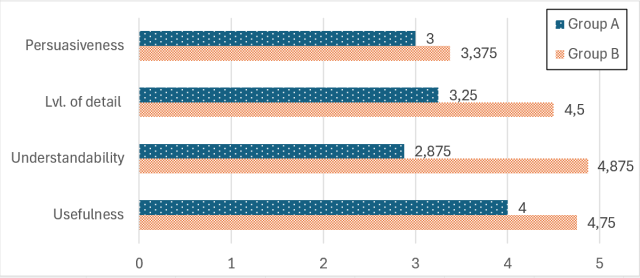

Figure 1: Evaluations of warnings: average values

the average value, 3). Large differences between the two groups can also be observed for *adequacy in the level of detail* (4,5 vs. 3,25) and *usefulness* (4,75 vs. 4), while the average evaluations are quite similar for *perceived persuasiveness*, i.e., participants' evaluations of the impact warnings had on their final decision whether to apply the rule or not.

The small difference in perceived persuasiveness seems to be in line with the number of participants who changed their mind on rule application after having received a warning: only 4 in group A and 6 in group B. These numbers, however, should be interpreted considering that most participants had correctly identified potentially dangerous rules since the beginning, even if they might not be fully aware of the actual risks they imply. If we focus on participants' behaviour with regards to specific rules, however, a more significant shift can be observed: one of the rules (*If the last family member leaves home, then turn off lights*) was deemed innocuous by all participants, who uniformly stated they would apply the rule in their smart home environment when first asked. After receiving a warning consisting of the label "Physical damage", only one participant in Group A changed their mind. In contrast, 5 participants in Group B took the final decision not to apply the rule after they had received a warning containing a more detailed description ("Turning off the lights predictably signals that the house is empty, making it easier for an attacker to plan an intrusion"). Hence, it seems that more informative and explainable warnings are crucial to convince users to make a choice which is in contrast with their initial evaluation, and that people might not be willing to simply trust privacy and security advice from the system if they do not understand the reasons behind it.

3 CONCLUSION

Results substantiate our hypothesis that more explainable warnings including informal-style explanations are more impactful and elicit more positive evaluations. While our experiment included too few participants to provide statistically significant results, it nevertheless represents a useful pilot for a larger-scale study.

ACKNOWLEDGMENTS

This work is partially supported by the Italian Ministry of University and Research (MIUR) under grant PRIN 2017 "EMPATHY: EMpowering People in deAling with internet of THings ecosYstems". We thank Bernardo Breve (University of Salerno) for his consultancy on security aspects.

REFERENCES

[1] Margherita Andrao, Fabrizio Balducci, Bernardo Breve, Federica Cena, Giuseppe Desolda, Vincenzo Deufemia, Cristina Gena, Maristella Matera, Andrea Mattioli, Fabio Paternò, Carmen Santoro, Barbara Treccani, Fabiana Vernero, and Massimo Zancanaro. 2023. Understanding Concepts, Methods and Tools for End-User Control of Automations in Ecosystems of Smart Objects and Services. In *End-User Development - 9th International Symposium, IS-EUD 2023, Cagliari, Italy, June 6-8, 2023, Proceedings (Lecture Notes in Computer Science, Vol. 13917)*, Lucio Davide Spano, Albrecht Schmidt, Carmen Santoro, and Simone Stumpf (Eds.). Springer, 104–124. https://doi.org/10.1007/978-3-031-34433-6_7

[2] Bernardo Breve, Gaetano Cimino, and Vincenzo Deufemia. 2023. Identifying Security and Privacy Violation Rules in Trigger-Action IoT Platforms With NLP Models. *IEEE Internet Things J.* 10, 6, March 15 (2023), 5607–5622. https://doi.org/10.1109/JIOT.2022.3222615

[3] Bernardo Breve, Giuseppe Desolda, Vincenzo Deufemia, Francesco Greco, and Maristella Matera. 2021. An end-user development approach to secure smart environments. In *International symposium on end user development*. Springer, 36–52.

[4] Fabio Casati, Silvana Castano, Mariagrazia Fugini, Isabelle Mirbel, and Barbara Pernici. 2000. Using patterns to design rules in workflows. *IEEE Transactions on Software Engineering* 26, 8 (2000), 760–785.

[5] Federica Cena, Cristina Gena, Claudio Mattutino, Michele Mioli, and Fabiana Vernero. 2023. From Psychological Traits to Safety Warnings: Three Studies on Recommendations in a Smart Home Environment. In *Joint Proceedings of the Workshops, Work in Progress Demos and Doctoral Consortium at the IS-EUD 2023 co-located with the 9th International Symposium on End-User Development (IS-EUD 2023), Cagliari, Italy, June 6-8, 2023 (CEUR Workshop Proceedings, Vol. 3408)*, Andrea Bellucci, Luigi De Russis, Paloma Díaz, Anders I. Mørch, Daniela Fogli, and Fabio Paternò (Eds.). CEUR-WS.org. https://ceur-ws.org/Vol-3408/short-s4-05.pdf

[6] Camille Cobb, Milijana Surbatovich, Anna Kawakami, Mahmood Sharif, Lujo Bauer, Anupam Das, and Limin Jia. 2020. How Risky Are Real Users'{IFTTT} Applets?. In *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*. 505–529.

[7] Fulvio Corno, Luigi De Russis, and Alberto Monge Roffarello. 2019. A high-level semantic approach to end-user development in the Internet of Things. *International Journal of Human-Computer Studies* 125 (2019), 41–54.

[8] Tran Anh Khoa, Le Mai Bao Nhu, Hoang Hai Son, Nguyen Minh Trong, Cao Hoang Phuc, Nguyen Thi Hoang Phuong, Nguyen Van Dung, Nguyen Hoang Nam, Dong Si Thien Chau, and Dang Ngoc Minh Duc. 2020. Designing efficient smart home management with IoT smart lighting: a case study. *Wireless communications and mobile computing* 2020 (2020), 1–18.

[9] Norman W Paton. 2012. *Active rules in database systems*. Springer Science & Business Media.

[10] Milijana Surbatovich, Jassim Aljuraidan, Lujo Bauer, Anupam Das, and Limin Jia. 2017. Some Recipes Can Do More Than Spoil Your Appetite: Analyzing the Security and Privacy Risks of IFTTT Recipes. In *Proceedings of the 26th International Conference on World Wide Web* (Perth, Australia) *(WWW '17)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 1501–1510. https://doi.org/10.1145/3038912.3052709

[11] DS Vijayan, A Leema Rose, S Arvindan, J Revathy, and C Amuthadevi. 2020. Automation systems in smart buildings: a review. *Journal of Ambient Intelligence and Humanized Computing* (2020), 1–13.