Towards Improving the Efficacy of Windows Security Notifier for Apps from Unknown Publishers: The Role of Rhetoric

Ankit Shrestha(✉), Rizu Paudel, Prakriti Dumaru, and Mahdi Nasrullah Al-Ameen

Department of Computer Science, Utah State University, Logan, USA
{ankit.shrestha,rizu.paudel,prakriti.dumaru,mahdi.al-ameen}@usu.edu

Abstract. With over 1.4 billion users of Windows 10, it is the most widely used operating system in the world. In Windows, applications from unknown publishers are popular due to mass availability and ease of access. Installing such applications can lead to malware infection, including viruses and ransomware. Therefore, we explored the design of interventions to prevent the users from installing applications from unknown publishers. To this end, we conducted a lab study with nine participants to understand the perceptions and behavior of users toward the designed interventions. Then, we conducted an online study with 256 participants to evaluate the impact of reflection, contextualization, and persuasion used in the finalized interventions. In summary, our findings provide valuable insights into understanding the needs and expectations of the users for usable and effective interventions against applications from unknown publishers. Based on our findings, we provide guidelines for future research.

Keywords: Reflective Design · Contextualization · Persuasion · Security Warnings · Windows

1 Introduction

The prior study [68] on security warning points to the lack of comprehension, where technical jargons [4,68], and habituation [3,5,50,61] lead users to ignore a security notifier. In these contexts, little study, to date, focused on the Windows notifier presented to users while installing an application from an unknown publisher. However, installing such applications can lead to malware infection [25,26,68]. The Windows operating system accounts for over 76% of global desktop operating systems [54][1] with over 1.4 billion devices of Windows 10 alone[2]; we believe that it is high time to focus on improving the design of Windows security notifiers to help users with better comprehension and informed decision making.

[1] https://www.statista.com/statistics/218089/global-market-share-of-windows-7.
[2] https://news.microsoft.com/bythenumbers/en/windowsdevices.

A. Moallem (Ed.): HCII 2023, LNCS 14045, pp. 101–121, 2023.
https://doi.org/10.1007/978-3-031-35822-7_8

To address this challenge, we designed a security notifier where we leveraged reflective design [42] with multiple persuasion techniques, including ethos, pathos, and logos [12,15]. We then investigated the following research questions, where we evaluated the designed security notifier (treatment) and compared that with the existing one (control): (**RQ1**): *What are the user perceptions about the existing security notifier presented to them while installing an application from an unknown publisher?* (**RQ2**): *How can we help users better understand the security risks of ignoring such notifiers and making an informed decision in the process?*

To answer these questions, we conducted a lab and an online study in North America (USA and Canada). In the lab study, we conducted semi-structured interviews with nine participants. The findings from our lab study reveal the participants' perceptions towards the existing and the designed notifier. We further improved our designs based on the feedback from the lab study and evaluated the updated designs through an online study with 256 participants on Amazon Mechanical Turk (Mturk).

Our findings from the online study show that reflection with persuasion in security warnings can be helpful while supporting the users to understand and combat the risks associated with applications from unknown publishers. Overall, our study contributes to advancing the HCI and Security community's understanding of end users' needs and expectations in helping them make an informed decision while installing the application from an unknown publisher in the Windows operating system. We provide recommendations based on our findings, including moving towards more reflective and contextualized interventions in future designs.

2 Related Work

Prior research [3–5,48,50,68] showed that users often ignore security warnings due to lack of comprehension, past experiences without consequences, optimism bias and the habituation to the warning. However, a little study focused on the Windows notifier presented to users while installing an application from an unknown publisher. Installing such applications can lead to malware infection [25,26,68]. Moreover, with over 76% of the global market share, the Windows operating system is by large the most widely used desktop operating system. The mass availability of applications from unknown publishers in Windows situates its users in a vulnerable position where they may face malware infections from installing such applications. Our study focuses on improving this existing security notifier to address the users' behavior and motivations behind ignoring security warnings.

2.1 Lack of Comprehension

Prior literature showed that users need help understanding the security warnings [4]. The study of Sharek et al. [55] reported that users needed to learn to differentiate between fake and real internet popup warnings. The study of Sunshine et al.

[57] further reported that users struggled to understand the SSL warnings in the browsers as they lacked knowledge about the situation and the harm related to man-in-the-middle attacks. Prior works [21,70] have further reported that users struggle to understand the context of the warning, which leads to poor comprehension and risky behavior. Further, a set of literature [14,68,70] reported the use of technical jargon as one of the major factors leading to difficulty for users in understanding the security warnings. The study of Bravo-Lillo et al. [14] also reported that novice users need help understanding technical wordings even when they have heard about it. Therefore, our study avoids technical terms, like ransomware and malware, to create user-friendly notifiers.

2.2 Past Experience

Prior works point towards the non-consequential experience of ignoring warnings as a significant factor for ignoring the same or similar security warnings [50,61]. In cases of informing, warning, or notifying users about consequences through security notifiers, most of the users tend to disregard those security warnings passing on the same message when they do not face any negative consequences, which inevitably leads to habituation [3,5,48,61]. The study of Amran et al. [3] reported that the habituation mechanism becomes universal if there is no adverse effect when a user ignores security dialog. Moreover, habituation to frequent non-security related notifications does carry over to a one-time security warning [61]. Windows provides similar notifications for installing applications from both verified and unknown publishers, which may magnify habituation to the latter.

According to Brustoloni and Villamarin-Salomon [16], habituation occurs as users learn to avoid context-sensitive guidance (CSG). As a consequence, CSG's purpose is to prompt the user to provide them with appropriate background information in order to help them make better security decisions. Based on the latest investigation and assessments, polymorphic alerts and iterative design are a few methods used to enhance security warnings to overcome habituation [5,16]. Therefore, our study uses multiple variations of the warning, created in an iterative design process through user feedback (see Sect. 5).

2.3 Optimism Bias

Prior literature from psychology [63] showed that individuals routinely overestimate their abilities and underestimate the risk they face compared to others, termed optimism bias. The study of Cho et al. [18] reported that individuals display a strong optimistic bias about online privacy risks, judging themselves to be significantly less vulnerable than others to these risks. Further, people tend to believe that specific security software like antivirus would protect them from any security threats [50]. There is also a common misconception among the participants regarding malware having an instantly visible effect [50]. Users want to believe that they cannot be the target, assuming that they have nothing valuable on their computer [50]. The study of Wu et al. [65] also reported that users

ignored the warnings when they believed the web content seemed legitimate. Therefore, our study considers optimism bias as one of the primary reasons why users ignore security warnings.

3 Design Principles

Prior works [21,34,48,65] found that users routinely ignore contextual warnings – such as banners or pop-ups. However, they notice interstitial interventions that interrupt their primary task. Therefore, we design and study multiple variations of interstitial interventions. These interventions start by shifting the primary task of the users from installing an application to self-reflection, where they are urged to reflect and understand why they want to ignore the warning (see Sect. 3.1). We then contextualize the information presented by the notifier where we focus on challenging the user's particular reason for ignoring the notifier (see Sect. 3.2). Finally, we leverage persuasion methods to present the contextualized information to motivate the users to avoid installing applications from unknown publishers (see Sect. 3.3).

3.1 Reflection on Rationales

Reflection refers to people's self examination of their own actions, understanding, and monitoring of progress [42]. Reflective designs have shown to promote conscious thought and decision making and help the users take a moment to realize their actions [23,40–42]. Moreover, prior literature [6,29,38,47] from psychology, marketing and human computer interaction showed that reflective designs are useful in increasing engagement and thoughtful decision making. Therefore, we translate and deploy reflective design in this study where users are urged to reflect on their own potential actions and understand their rationales behind it. To achieve that, we use the reasons behind ignoring security warnings (see Sect. 2) to create the reflective design (see the central interface in Fig. 2). In the reflective design, we intervene the task of installing the application from an unknown publisher and ask them to identify their reason for ignoring the notifier.

3.2 Contextualization

Contextualization in design is the process of understanding the underlying context, rationale or intention (e.g., why do users ignore security warnings?) and designing the required artifact based on the identified context [28,66]. Works [7,8,24,45] from education and human computer interaction used contextualization in designing education content and web warnings respectively. The findings from these studies points towards the importance of contextualizing the information provided to the users. Further, prior literature [31,37,50] showed that users ignore warnings that provide generic information which they perceive as distant harm. Studies from psychology [49,60] suggest conveying negative impacts as it is more effective than citing advantages. Studies from Xu et al. [67], and Kaiser et al. [34] also showed that the conveyance of specific harm to the users is an

effective deterrent in convincing them to avoid risky activities. Therefore, we use the rationale selected by the users in our reflective intervention to contextualize the information in our warnings.

3.3 Persuasion Modes

We contextualized the harm based on their rationale for ignoring the warning. However, prior works pointed to the benefits of persuasion in order to motivate users to comply with the warnings [33,52]. The objective of the warning is not only to inform but also to persuade users to avoid risky activities without hindering their freedom of choice [33]. Hence, we use Aristotle's Rhetoric [15] (ethos, pathos, logos) to illustrate the contextualized harm to persuade users to avoid installing applications from unknown publishers. Ethos is persuasion using authority or credibility of character [15]. Pathos is an appeal to emotion of the user [15]. Logos is an appeal to logic by using statistics, facts, and figures [15]. Prior works [12,19,30,39] from psychology and political science used Aristotle's Rhetoric to understand persuasive communication. In our study, we use these rhetorics to persuade the users by appealing to authority, emotion or logic.

4 Lab Study Methodology

We used the existing Windows notifier as the control condition (see Fig. 1). We then created warning designs (see Fig. 2) adapting design recommendations from prior literature [12,39,42,46,48,49,60] which we call treatment condition. Using these designs, we conducted the lab study.

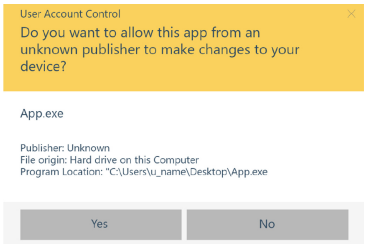

Fig. 1. Control condition for both lab and online study

In the lab study, we conducted semi-structured interviews (see Sect. 4.1) with nine participants online through Zoom/Skype between March and April 2021. The participants for the study were recruited using snowball sampling via email. A participant had to be at least 18 years old to participate in this study. Details of the participants are available in Table 1. The Institutional Review Board approved the study at our university.

4.1 Study Procedure

When a participant showed interest in participating in our study, we emailed them the informed consent document (ICD). Once they agreed to the ICD, we scheduled an online interview through Zoom or Skype. In the interview, the participants were presented with the same scenario for the control and treatment conditions in which the notifier occurred. Then, the participant interacted with the notifier and answered interview questions focused on understanding their perceptions and behavior. At the end of the interview, the participants were asked to complete a demographics survey. After completing the interview, each participant was sent an email thanking them for participating in this study.

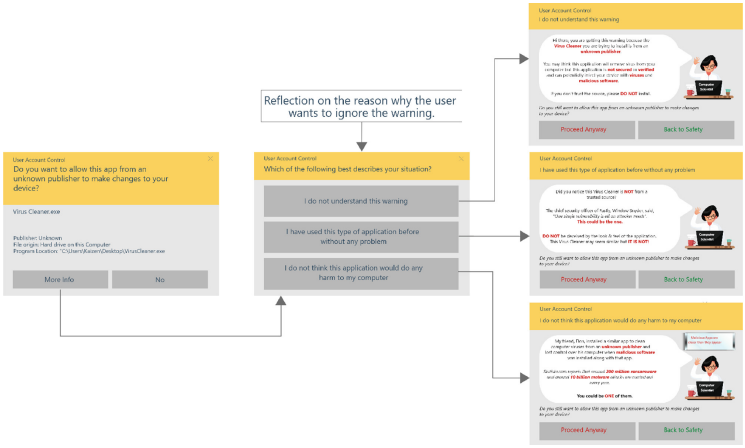

Fig. 2. Treatment condition and the flow of interaction in the lab study. (The flow is the same in online study.)

Table 1. Demographic Information of the Lab Study Participants

PID	Gender	Age Range	Education
P1	Male	18–24 years old	Graduate Degree
P2	Male	30–34 years old	Graduate Degree
P3	Female	25–29 years old	Graduate Degree
P4	Female	25–29 years old	Four-year College Degree
P5	Male	25–29 years old	Four-year College Degree
P6	Prefer not to answer	18–24 years old	Four-year College Degree
P7	Male	25–29 years old	Graduate Degree
P8	Female	25–29 years old	Four-year College Degree
P9	Female	18–24 years old	Two-year College Degree

4.2 Analysis

The audio recordings from the interview were transcribed. Then, we performed thematic analysis on our transcriptions [9,11,13,56]. Two independent researchers coded each transcript, where they read through the transcripts of the first few interviews, developed codes, compared them, and then iterated again until we had developed a consistent codebook. After the codebook was finalized, two researchers independently coded the remaining interviews. Both researchers spot-checked the other's coded transcripts and found no inconsistencies. Finally, we organized and taxonomized our codes into higher-level categories.

5 Design Evolution

In this section, we will present the qualitative feedback from the participants on our designs and the changes we have made to address the issues raised by

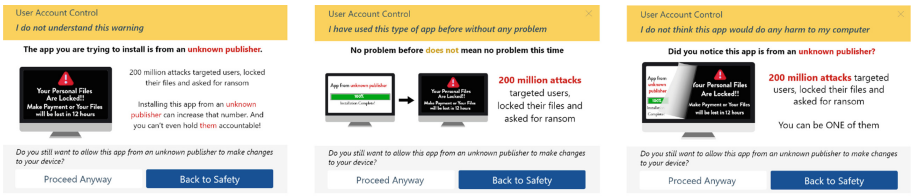

Fig. 3. Logos treatment condition in the online study

them. For consistency, we use these terms throughout the manuscript based on the frequency of comments in participants' responses: *a few* (0–10%), *several* (10–25%), *some* (25–40%), *about half* (40–60%), *most* (60–80%), and *almost all* (80–100%).

In the lab study, most participants reported that the control condition (see Fig. 1) needed to be more specific and clear as the notifier was unable to provide sufficient context for them to make an informed decision. In contrast, they found the treatment condition (see Fig. 2) to be informative; one of them stated, *"This [treatment] version of notifier was really like something that I was looking forward to that really solved my problem that I was facing in the previous notifier with clearly identifying what might be the issue that you are facing or what might be the consequences of you trying to access this [application from unknown publisher]." (P7).*

Our participants also reported satisfaction with the presentation of options that account for the reasons behind a user's intention to ignore a warning. One of them mentioned, *"It also showed options that I don't understand this warning, or that I have already used this application before ... so that I know beforehand that, these are certain things that I will have to keep in mind when I try to access this application, ... so I can use this application fully prepared." (P4).* The effectiveness of the thought-provoking questions can be attributed to the reflective design that we discussed in Sect. 3. This motivated us to retain the reflective design in our interventions for the online study.

However, the persuasion-based designs also needed improvements as presented below which we addressed through focus group discussions between the authors.

5.1 Inducing Focus

For the designs used in our lab study, we combined the three rhetorics for the treatment condition, which resulted in increased information (see Fig. 2). Some participants in the lab study found the amount of text and information in the treatment condition overwhelming. One of them reported, *"... it [treatment condition] was more clustered than I wanted to. There are certain points that are highlighted, but I would also suggest that it be more visual than more textual. So just by looking at it, we can understand that there are certain issues there that we might come across." (P1).* Therefore, to reduce the cognitive burden

Fig. 4. Pathos treatment condition in the online study

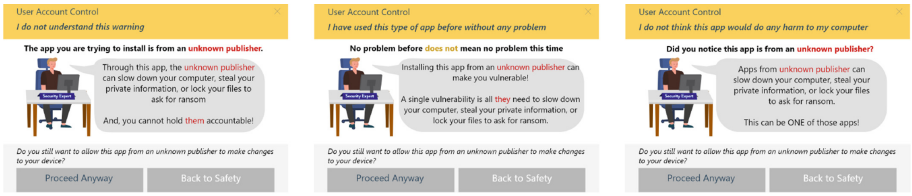

Fig. 5. Ethos treatment condition in the online study

[36,58,59], we focused on creating multiple variations of the designs focused on a particular rhetoric (see Fig. 3, 4, and 5). Moreover, as the amount of information was reduced with increased focus, we could replace texts in the design with graphical components. These changes were also motivated by prior works [25,27,43,44], which use graphics to increase perception speed and memorability of the information.

5.2 Design Identity

In the lab study, some participants found it challenging to differentiate the designs for the three reflective options (scenarios). One of them reported, *"... the three instructions were on a similar fashion. Only on the bubble of the computer representative in the instruction was changed. So, what I could suggest is you have three instructions on like different graphical format or different visual format, so that they can be separated distinctly." (P1).* To help participants avoid mistaking the different designs as the same, we imbued each design with different graphics to create their identity. Since graphics are more memorable and perceived faster [43,44], we believed the changes would help the participants identify the designs for the different options.

5.3 Overcoming Experience Bias

A few of our participants reported on their experiences installing applications from unknown publishers where they faced no problems and argued against the warning we had presented. They mentioned that there are many applications from unknown publishers that are from unverified publishers. In such cases, the notification occurs, but it is not always an infected software. To overcome this

Table 2. Demographic Information of the Participants in the Online Study (N = Number of Participants)

Demographic	Demographic Group	N
Gender	Male	146
	Female	109
	Prefer not to answer	1
Age range	18-24 years old	4
	25–29 years old	64
	30–34 years old	36
	35–39 years old	58
	40–44 years old	40
	45–49 years old	15
	50–54 years old	18
	55–59 years old	7
	60–64 years old	8
	Above 65 years old	4
	Prefer not to answer	2
Race	White	183
	Asian	50
	Black/African American	6
	Hispanic or Latino	5
	Native American	4
	Mixed Race	6
	Prefer not to answer	2
Education	High School Graduate	35
	Two-year College Degree	17
	Four-year College Degree	157
	Graduate degree	44
	Prefer not to answer	2
	Other	1
Major	Computer-Related Major	101
	Non-Computed Related Major	146
	Prefer not to answer	9

bias based on the user's experience, we changed the sentiment of the design to convey that not having problems before does not mean there will be no problems this time. We also provided scenarios depicting the severe consequences when one might face problems to dissuade the users from avoiding the warning. Prior works [34,67] have also shown that conveying relevant adverse harm can effectively deter users from risky activities.

6 Online Study Methodology

We changed the treatment conditions' design based on the lab study findings (see Sect. 5). Then, we used them in an online study conducted through Amazon Mechanical Turk (MTurk) with 256 participants. We created our system for data collection, as we had multiple variations of interactive designs, which were not feasible for existing survey systems. We selected the widely used User Experience Questionnaire plus (UEQ+) scale[3] [51] to understand the user experience and the effectiveness of the warnings. We presented the questions in random order in the survey, with some reversed to avoid bias [20,64]. Additionally, we used nine attention-check questions in random order, following procedures suggested by prior works [32,35].

6.1 Participant Recruitment

We recruited participants using Amazon Mechanical Turk (Mturk). While imperfect, MTurk can provide data of at least the same quality as methods tradition-

[3] https://www.ueq-online.org/.

ally used in research, as long as the experiment is designed carefully [10,17]. Participants had to be 18 or older and live in the United States or Canada to participate in our study. We compensated the participants with USD 2.5 for the study, which took approximately 15 min, even if they failed the attention check questions. In our analysis, we only used responses from the participants who correctly answered all nine of our attention check questions. The summary of the participants' demographics is available in Table 2.

6.2 Procedure

Participants interested in our study would first accept the task in Mturk and review the ICD provided in the survey. Clicking the link to our online study system meant that the participants agreed to the ICD. The participants were greeted with information about the survey in our system. Then, the participants interacted with one of the four conditions (Control, Ethos, Pathos, and Logos). Moreover, the three treatment conditions had designs for the reflective rationales that the users could select. A survey including open-ended questions followed each design. Finally, the participants answered questions about their demographics and prior knowledge about applications from unknown publishers. At the end of the study, we provided the participants with a seven-digit code, which they entered into the Mturk Survey to complete the study.

6.3 Analysis

We use statistical tests to analyze our quantitative results. We consider results to be significant when we find p ¡ .05, but further highlight results with lower p values. When comparing two conditions, we use a Wilcoxon signed rank test for the matched pairs of subjects and a Wilcoxon Mann-Whitney test for unpaired results. Wilcoxon tests are similar to t-tests but do not assume the distributions of the compared samples, which is appropriate for our collected data.

For the qualitative results from the open-ended questions, we performed thematic analysis, where two independent researchers coded the responses and later discussed and resolved the discrepancies in the codes.

7 Online Study

After making changes based on the suggestions from the lab study, we created a survey system to conduct an online study in Amazon Mechanical Turk (see Sect. 4). Each user was either provided with the control condition (see Fig. 1) or one of the three treatment conditions (see Fig. 3, 4, and 5). The three scenarios (see Sect. 4) in treatment conditions were presented randomly to mitigate order effects. We observed that the randomization was successful, as there is a lack of significant order effects between the three conditions (see Table 3).

Table 3. Order effects between the different scenarios of treatment condition

Scenarios		Wilcoxon-Signed Rank Test	
First	Second	W	p
Understanding	Experience	1430829.0	0.137
Understanding	Optimism	444029.0	0.404
Experience	Optimism	446506.5	0.959

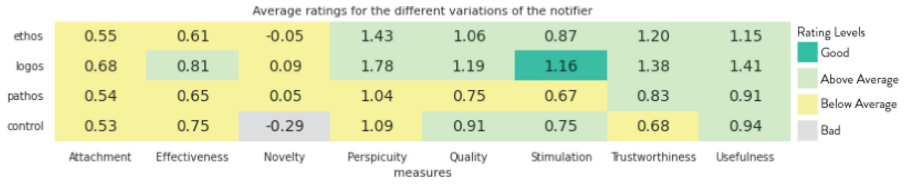

Fig. 6. Average ratings for different notifier variations

7.1 User Ratings: Sensemaking in the Context of Warning Design

Figure 6 provides the average scores along with UEQ recommended category (color-coded) for all the 24 variations of the warnings we have used in the online study.

In light of the UEQ benchmark[4], we observed that all the warnings had above-average scores for usefulness. That implies most users consider it important to be notified about applications from unknown publishers. One of the participants reported, *"This alert is useful when you want to ensure the security of your PC and avoid accidental changes to important settings."* However, only Logos was rated above average in terms of effectiveness. The high scores in the effectiveness measure may be due to the factual information presented in Logos, which some participants reported as their primary reason for liking the warning. One of them said, *"I like that it provides information about the number of attacks that have happened, and it makes me really think if it is worth it to download the app."*

We also observed that all treatment condition warnings were considered above average in trustworthiness whereas the control condition was not. Most participants preferred the contextualized information about the application (see Sect. 3), which increased their trust in the warning. Further, the reflective nature of the treatment condition (see Sect. 3) helped increase the users' trust in the warning. Most participants liked the specific scenarios addressed by the warning to persuade them to avoid installing the application. One of them mentioned, *"I like the fact that the notifier will tell you exactly some of the issues you will experience if the unknown publisher has a virus that will infect your system later."* Similarly, the participants reported on particular scenarios and how the treatment condition works, convincing them to avoid the application installation. One of them said, *"It addresses a common misconception that if you have*

[4] https://www.ueq-online.org/Material/Handbook.pdf.

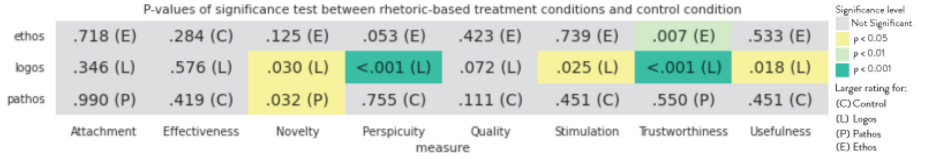

Fig. 7. P-values from the significance tests between the control condition and variations of treatment condition

downloaded software from an unknown publisher before and didn't get a virus or malware that it is OK to do so this time."

Finally, we observed that both Ethos and Logos performed above average in terms of perspicuity, quality, and stimulation. We had mixed responses for these measures, which we explore in detail in Sect. 7.3.

7.2 Control vs. Treatment Conditions

As we discussed the average ratings of the warnings, next, we compared the three variations of our treatment condition with the control condition (see Fig. 7). We observed that Logos and Ethos performed significantly better than the Control in terms of trustworthiness. That could be due to the factual nature of Logos and the portrayal of a credible source in Ethos, which are both lacking in the control condition [12,19,30,39]. Some of our participants also mentioned these traits of the designs; where one of the participants talking about Ethos reported, *"I like how it seems credible based on the name tag next to the man."*

The added useful information and the thought-provoking nature of the warnings mentioned by some of our participants could have resulted in significantly higher scores in perspicuity, stimulation, and usefulness for Logos. One participant, when mentioning Logos, said, *"It's relevant and timely: The user's behavior, location, or preference triggers the notification. It's personal: The content of the push appeals to the user as an individual. It's actionable: The push makes it clear what the user should do next."*

7.3 Comparison Between the Treatment Conditions

We compared the three variations of the treatment conditions with each other (see Fig. 8). We observed that Logos performed significantly better than Pathos

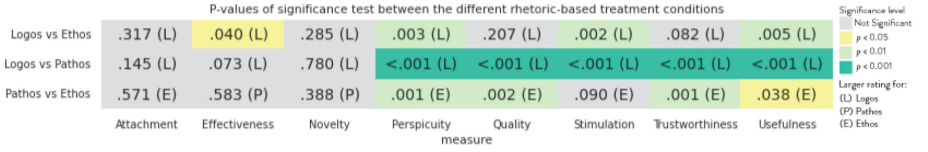

Fig. 8. P-values from the significance tests between different variations of treatment condition

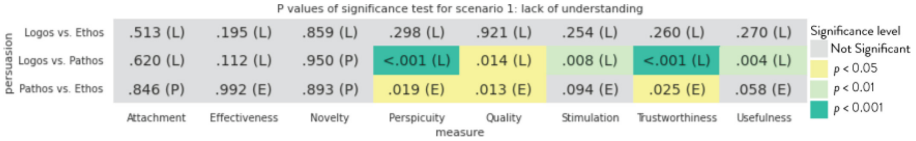

Fig. 9. P-values from the significance tests between different rhetoric used in scenario 1 of treatment condition

and Ethos in terms of perspicuity, stimulation, and usefulness. Qualitative responses from about half of the participants indicate that they liked the factual information presented in Logos, which immediately motivated them to avoid the warning. One of them said, *"This notice is very clear that there is a serious issue with this app. If these stats are true then I would never download something like this."*

Further, Logos was rated significantly higher than Pathos for information quality and trustworthiness. As we discussed above, participants found the factual information in Logos helpful which could have also increased their perceptions of trustworthiness and quality of information. Similarly, Ethos was also rated significantly higher than Pathos for information quality and trustworthiness. Ethos uses credible and authoritative sources to provide relevant information to the users. Some users reported that such a delivery helped them make an informed decision. One of them reported. *"That [security expert] gives me specific information 'unknown publisher' so if I know the publisher and feel comfortable I can feel safe to install it."*

7.4 Scenario-Based Evaluation: Rhetoric Behind the Interventions

In this section, we focus on each of the three scenarios we addressed as part of our reflective design and understand the rhetoric that can be useful for these scenarios.

Scenario I: Lack of Comprehension. Figure 9 summarizes the significance tests performed between the persuasion principles for scenario 1.

In this scenario where users did not understand the warning, we observed that both Logos and Ethos performed significantly better than Pathos regarding perspicuity, information quality, and trustworthiness. Comments from some of our participants revealed that they liked the easy-to-comprehend Logos and Ethos warnings. One of them said, *"It warns you in a clear and concise way what could happen by installing unknown apps and programs. It is also easy to read, and the colors are easy on the eyes."* Moreover, participants found the idea of helping the users by first understanding their level of knowledge preferable which could have resulted in higher scores for information quality and trustworthiness. One of them said, *"I like that it goes in-depth about what it means only after you said you don't understand. Good for people who aren't familiar with technology that much."*

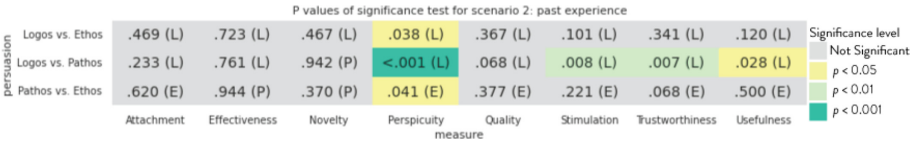

Fig. 10. P-values from the significance tests between different rhetoric used in scenario 2 of treatment condition

On the other hand, some participants found the storytelling in Pathos challenging to understand. One of them said, *"I like that it is trying to be fun and interesting, it just isn't very understandable because of it. I also like the colors and pictures used."* However, some participants thought Pathos was playful and exciting. One of them said, *"I like the way the images look, I also like it shows the hacker guy, and then you having your files locked so kind of hits harder and just like the look. Also, it tells you what could happen, like one of the worst cases of what could happen but does it in a way that's more playful"*

In conclusion, for the scenario, both Logos and Ethos performed significantly better than Pathos and should be considered in future designs to increase the understanding of the users.

Scenario II: Past Experience. In the second scenario of the user's past experience, we observed that Logos performed significantly better than Pathos in terms of stimulation, trustworthiness, and usefulness (see Fig. 10). Some participants found Logos to be thought-provoking considering how it challenges our primary task to understand and decide in an informed manner. One of them said, *"I feel like sometimes we get too busy to care about things and just accept whatever notifications when we are for instance trying to install a video game and our friends are waiting on us to complete the install. This actually happened just last night."* Some participants found the facts and statistics helpful, whereas a few found the graphics in Logos, particularly representative. One of them reported, *"I like the detailed pictorial representation in the notifier. I like it because it clearly indicates the possibility of an app not being safe even if it has been previously tested to be safe due to a past user experience."*

Moreover, Logos performed significantly better than both Pathos and Ethos in terms of perspicuity, where some participants mentioned the ease of understanding the logical reasoning provided in the Logos. In conclusion, for scenario II, Logos performed the best, but there was a significant difference between Logos and Ethos in only one measure.

Scenario III: Optimism Bias. In the final scenario of optimism bias, we observed that Logos performed significantly better than both Pathos and Ethos in terms of Perspicuity, Stimulation and Usefulness (see Fig. 11). Qualitative responses revealed that some participants found the image used in the Logos design interesting, which could have resulted in higher scores for stimulation.

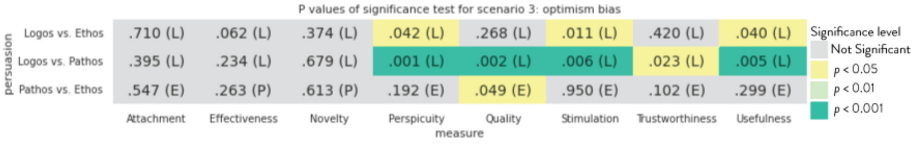

Fig. 11. P-values from the significance tests between different rhetoric used in scenario 3 of treatment condition

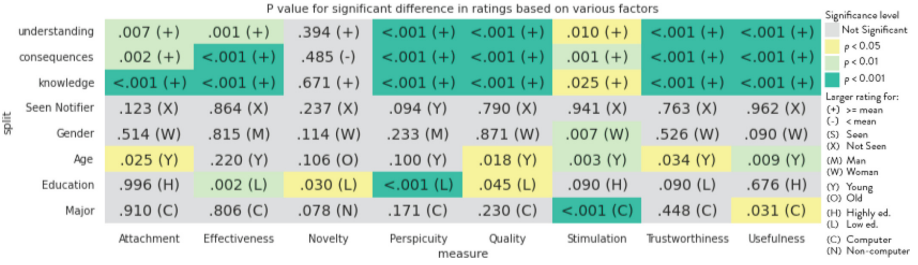

Fig. 12. P-values of significance tests showing the impact of various factors on the ratings of the warnings

One of them reported, *"Best thing is the image of the screen peeling back to reveal a possible ransomware warning. I like how you still have the choice to proceed or not though."* While about half of the participants found the logical reasoning easy to understand, some participants also expressed that the warning addressed the optimism bias appropriately making it useful. One of them said, *"I think the good thing is that it makes you think, it makes you question whether it is worth it to download the app. It gives you facts, and then states you could be one of them, because I think people believe things happen to other people, not to themselves."*

7.5 Impact of User Demographics on Warning Perceptions

We observed that the user demographics had varying impacts on the warning ratings (see Fig. 12). The ratings for the warnings are significantly higher for all measures except novelty for participants with a higher understanding and knowledge about the applications from unknown publishers. We further observed that there is no significant difference in ratings between the users who have seen the existing Windows notifier and the users who have not.

Moreover, female participants rated the warnings significantly higher than their male counterparts in terms of stimulation. Younger participants (18–39) rated the warnings significantly higher than older participants (older than 39) regarding attachment, information quality, stimulation, trustworthiness, and usefulness. In addition, less-educated participants (high school or less) rated the warnings significantly higher than highly-educated participants (2-year college degree or more) in effectiveness, novelty, perspicuity, and information quality.

Similarly, participants with computing backgrounds rated the warnings significantly higher regarding stimulation and usefulness. These findings imply that certain groups of participants may benefit more from the use of persuasion-based interventions.

8 Discussion

Our findings report on the perceptions of the users towards applications from unknown publishers and the effectiveness of the reflective rhetoric-based notifiers against them. In this section, we discuss the possible implications of our findings and provide suggestions to consider in future designs.

8.1 Moving Towards Reflective Design

Prior literature [22,69] reported the behavior and perceptions of users towards security warnings where they consider it the secondary task. In our study, the user is also primarily motivated to install the application from an unknown publisher. However, dealing with security warnings becomes a secondary task. Therefore, reflection is an essential step in the design of security warnings that intervenes the users to take a moment to identify their rationale in doing a risky activity. Our findings show that the use of reflective designs can be a practical approach in convincing users to avoid installing applications from unknown publishers (see Sect. 7.1 and 5). However, few works in computer science have used reflective designs that first aim to understand the context of the users and then present information based on the identified context. Our work provides the direction for future works to adopt and evaluate the reflective designs in various security warnings and beyond the scope of such interventions.

8.2 Addressing Habituation

Our findings highlight the importance of contextualizing the warning where participants appreciated addressing their selected rationale for installing applications from unknown publishers (see Sect. 7.4). In our designs, the contextualization of information and persuasion modes (see Sect. 3) have further resulted in polymorphic warnings. The study of Vance et al. [62] reported habituation as a significant inhibitor to the effectiveness of security warnings. However, prior works [5,16] showed that the use of polymorphic warnings could prevent habituation in the long term. Moreover, our findings show a significant impact of users' understanding of the applications from unknown publishers on the performance of the interventions (see Sect. 7.5). Therefore, understanding the reason behind the user's tendency to do a risky activity should be considered an important context in designing future security warnings. By doing that, we can address specific issues that the users face while simultaneously avoiding habituation.

8.3 Limitations and Future Work

Our study was limited to participants from the U.S. and Canada. However, recent HCI studies [1,2,53] highlight the importance of looking beyond Western contexts. Hence, future works should include participants from diverse regions to understand their perceptions and create effective interventions.

In our lab study, we interviewed nine participants by following widely-used methods for qualitative research [9,11,13,56]. We acknowledge the limitations of these studies, that a different set of samples might yield varying results. Thus, we do not draw any quantitative, generalizable conclusion from the lab study. Instead, we conduct an online study with sufficient statistical power, leveraging the findings from the lab study to reach generalizable results.

Our study focuses on a single security intervention, whereas further work is needed to understand the validity of the results for different warnings and designs. As we continuously improve designs in future iterations, we should move from just informing the user to promoting reflection where we can motivate and help them in context.

References

1. Al-Ameen, M.N., Kocabas, H.: "i cannot do anything": user's behavior and protection strategy upon losing, or identifying unauthorized access to online account. In: Symposium on Usable Privacy and Security (Poster Session) (2020)
2. Al-Ameen, M.N., Kocabas, H., Nandy, S., Tamanna, T.: We, three brothers have always known everything of each other: a cross-cultural study of sharing digital devices and online accounts. Proc. Priv. Enhancing Technol. **2021**(4), 203–224 (2021)
3. Amran, A., Zaaba, Z.F., Mahinderjit Singh, M.K.: Habituation effects in computer security warning. Inf. Secur. J. Glob. Perspect. **27**(4), 192–204 (2018)
4. Amran, A., Zaaba, Z.F., Singh, M.M., Marashdih, A.W.: Usable security: revealing end-users comprehensions on security warnings. Procedia Comput. Sci. **124**, 624–631 (2017)
5. Anderson, B.B., Kirwan, C.B., Jenkins, J.L., Eargle, D., Howard, S., Vance, A.: How polymorphic warnings reduce habituation in the brain: insights from an FMRI study. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp. 2883–2892 (2015)
6. Baek, E., Choo, H.J., Wei, X., Yoon, S.Y.: Understanding the virtual tours of retail stores: how can store brand experience promote visit intentions? Int. J. Retail Distrib. Manage. (2020)
7. Bartsch, S., Volkamer, M.: Towards the systematic development of contextualized security interventions1. In: The 26th BCS Conference on Human Computer Interaction, vol. 26, pp. 1–4 (2012)
8. Bartsch, S., Volkamer, M., Theuerling, H., Karayumak, F.: Contextualized web warnings, and how they cause distrust. In: Huth, M., Asokan, N., Čapkun, S., Flechais, I., Coles-Kemp, L. (eds.) Trust 2013. LNCS, vol. 7904, pp. 205–222. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-38908-5_16
9. Baxter, K., Courage, C., Caine, K.: Understanding Your Users: A Practical Guide to User Research Methods, 2nd edn. Morgan Kaufmann Publishers Inc., San Francisco (2015)

10. Berinsky, A.J., Huber, G.A., Lenz, G.S.: Evaluating online labor markets for experimental research: Amazon. com's mechanical Turk. Polit. Anal. **20**(3), 351–368 (2012)
11. Boyatzis, R.E.: Transforming Qualitative Information: Thematic Analysis and Code Development. Sage Publications, Thousand Oaks (1998)
12. Braet, A.C.: Ethos, pathos and logos in Aristotle's rhetoric: a re-examination. Argumentation **6**(3), 307–320 (1992)
13. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qual. Res. Psychol. **3**(2), 77–101 (2006)
14. Bravo-Lillo, C., Cranor, L.F., Downs, J., Komanduri, S.: Bridging the gap in computer security warnings: a mental model approach. IEEE Secur. Priv. **9**(2), 18–26 (2010)
15. Brinks, M.: Ethos, pathos, logos, Kairos: the modes of persuasion and how to use them. Prep Scholar (2019). Accessed 20 Aug 2021
16. Brustoloni, J.C., Villamarín-Salomón, R.: Improving security decisions with polymorphic and audited dialogs. In: Proceedings of the 3rd Symposium on Usable Privacy and Security, pp. 76–85 (2007)
17. Buhrmester, M., Kwang, T., Gosling, S.D.: Amazon's mechanical Turk: a new source of inexpensive, yet high-quality data? (2016)
18. Cho, H., Lee, J.S., Chung, S.: Optimistic bias about online privacy risks: testing the moderating effects of perceived controllability and prior experience. Comput. Hum. Behav. **26**(5), 987–995 (2010)
19. Demirdöğen, Ü.D.: The roots of research in (political) persuasion: ethos, pathos, logos and the Yale studies of persuasive communications. Int. J. Soc. Inquiry **3**(1), 189–201 (2010)
20. DeSimone, J.A., Harms, P.D., DeSimone, A.J.: Best practice recommendations for data screening. J. Organ. Behav. **36**(2), 171–181 (2015)
21. Egelman, S., Cranor, L.F., Hong, J.: You've been warned: an empirical study of the effectiveness of web browser phishing warnings. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1065–1074 (2008)
22. Egelman, S., Schechter, S.: The importance of being earnest [in security warnings]. In: Sadeghi, A.-R. (ed.) FC 2013. LNCS, vol. 7859, pp. 52–59. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39884-1_5
23. Fang, Y.M., Chen, K.M., Huang, Y.J.: Emotional reactions of different interface formats: comparing digital and traditional board games. Adv. Mech. Eng. **8**(3), 1687814016641902 (2016)
24. Fernandes, P., Leite, C., Mouraz, A., Figueiredo, C.: Curricular contextualization: tracking the meanings of a concept. Asia Pac. Educ. Res. **22**, 417–425 (2013)
25. Good, N., et al.: Stopping spyware at the gate: a user study of privacy, notice and spyware. In: Proceedings of the 2005 Symposium on Usable Privacy and Security, pp. 43–52 (2005)
26. Good, N., Grossklags, J., Thaw, D., Perzanowski, A., Mulligan, D.K., Konstan, J.: User choices and regret: understanding users' decision process about consensually acquired spyware. I/S J. Law Policy Inf. Soc. **2**(2), 283–344 (2006)
27. Good, N.S., Grossklags, J., Mulligan, D.K., Konstan, J.A.: Noticing notice: a large-scale experiment on the timing of software license agreements. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 607–616 (2007)
28. Haleblian, K.: The problem of contextualization. Missiology **11**(1), 95–111 (1983)

29. Heidig, S., Müller, J., Reichelt, M.: Emotional design in multimedia learning: differentiation on relevant design features and their effects on emotions and learning. Comput. Hum. Behav. **44**, 81–95 (2015)
30. Higgins, C., Walker, R.: Ethos, logos, pathos: strategies of persuasion in social/environmental reports. In: Accounting Forum, vol. 36, pp. 194–208. Elsevier (2012)
31. Hora, A., Anquetil, N., Ducasse, S., Allier, S.: Domain specific warnings: are they any better? In: 2012 28th IEEE International Conference on Software Maintenance (ICSM), pp. 441–450. IEEE (2012)
32. Ipeirotis, P.G., Provost, F., Wang, J.: Quality management on amazon mechanical Turk. In: Proceedings of the ACM SIGKDD Workshop on Human Computation, pp. 64–67 (2010)
33. Jones, C.P., Robinson, S.J., Sabadosh, N., Bishop, D., Koyani, S.: How can rhetoric and argumentation help us make the case for UCD? In: CHI 2006 Extended Abstracts on Human Factors in Computing Systems, pp. 415–418 (2006)
34. Kaiser, B., Wei, J., Lucherini, E., Lee, K., Matias, J.N., Mayer, J.: Adapting security warnings to counter online disinformation. In: 30th USENIX Security Symposium (USENIX Security 2021), pp. 1163–1180 (2021)
35. Kung, F.Y., Kwok, N., Brown, D.J.: Are attention check questions a threat to scale validity? Appl. Psychol. **67**(2), 264–283 (2018)
36. Lenzner, T., Kaczmirek, L., Lenzner, A.: Cognitive burden of survey questions and response times: a psycholinguistic experiment. Appl. Cogn. Psychol. **24**(7), 1003–1020 (2010)
37. Lesch, M.F., Powell, W.R., Horrey, W.J., Wogalter, M.S.: The use of contextual cues to improve warning symbol comprehension: making the connection for older adults. Ergonomics **56**(8), 1264–1279 (2013)
38. Lindgaard, G., Dudek, C., Sen, D., Sumegi, L., Noonan, P.: An exploration of relations between visual appeal, trustworthiness and perceived usability of homepages. ACM Trans. Comput. Hum. Interact. (TOCHI) **18**(1), 1–30 (2011)
39. Mshvenieradze, T.: Logos ethos and pathos in political discourse. Theor. Pract. Lang. Stud. **3**(11) (2013)
40. Norman, D.: The Design of Everyday Things: Revised and expanded edition. Basic books (2013)
41. Norman, D.A.: Introduction to this special section on beauty, goodness, and usability. Hum. Comput. Interact. **19**(4), 311–318 (2004)
42. Norman, D.A., Ortony, A.: Designers and users: two perspectives on emotion and design. In: Symposium on Foundations of Interaction Design, pp. 1–13 (2003)
43. Paivio, A.: Mind and Its Evolution: A Dual Coding Theoretical Approach. Psychology Press, London (2014)
44. Parkinson, M.: The power of visual communication. Billion Dollar Graphics (2012)
45. Perin, D.: Facilitating student learning through contextualization: a review of evidence. Commun. Coll. Rev. **39**(3), 268–295 (2011)
46. Petelka, J., Zou, Y., Schaub, F.: Put your warning where your link is: improving and evaluating email phishing warnings. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–15 (2019)
47. Peters, D., Calvo, R.A., Ryan, R.M.: Designing for motivation, engagement and wellbeing in digital experience. Front. Psychol. **9**, 797 (2018)
48. Reeder, R.W., Felt, A.P., Consolvo, S., Malkin, N., Thompson, C., Egelman, S.: An experience sampling study of user reactions to browser warnings in the field. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, pp. 1–13 (2018)

49. Rozin, P., Royzman, E.B.: Negativity bias, negativity dominance, and contagion. Pers. Soc. Psychol. Rev. **5**(4), 296–320 (2001)
50. Sasse, M.A., Krol, K., Moroz, M.: Don't work. can't work? why it's time to rethink security warnings. In: 2012 7th International Conference on Risks and Security of Internet and Systems (CRiSIS), pp. 1–8. IEEE Computer Society (2012)
51. Schrepp, M., Hinderks, A., Thomaschewski, J.: Applying the user experience questionnaire (UEQ) in different evaluation scenarios. In: Marcus, A. (ed.) DUXU 2014. LNCS, vol. 8517, pp. 383–392. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-07668-3_37
52. Seo, H., Xiong, A., Lee, D.: Trust it or not: effects of machine-learning warnings in helping individuals mitigate misinformation. In: Proceedings of the 10th ACM Conference on Web Science, pp. 265–274 (2019)
53. Shahid, F., Kamath, S., Sidotam, A., Jiang, V., Batino, A., Vashistha, A.: It matches my worldview: examining perceptions and attitudes around fake videos. In: CHI Conference on Human Factors in Computing Systems, pp. 1–15 (2022)
54. Share, N.M.: Operating system market share (2009). https://marketshare.hitslink.com/operating-system-market-share.aspx
55. Sharek, D., Swofford, C., Wogalter, M.: Failure to recognize fake internet popup warning messages. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 52, pp. 557–560. SAGE Publications Sage CA: Los Angeles, CA (2008)
56. Shrestha, A., Graham, D.M., Dumaru, P., Paudel, R., Searle, K.A., Al-Ameen, M.N.: Understanding the behavior, challenges, and privacy risks in digital technology use by nursing professionals. Proc. ACM Hum. Comput. Interact. **6**(CSCW2), 1–22 (2022)
57. Sunshine, J., Egelman, S., Almuhimedi, H., Atri, N., Cranor, L.F.: Crying wolf: an empirical study of SSL warning effectiveness. In: USENIX Security Symposium, pp. 399–416. Montreal (2009)
58. Sweller, J.: Cognitive load theory: recent theoretical advances (2010)
59. Sweller, J.: Cognitive load theory. In: Psychology of Learning and Motivation, vol. 55, pp. 37–76. Elsevier (2011)
60. Vaish, A., Grossmann, T., Woodward, A.: Not all emotions are created equal: the negativity bias in social-emotional development. Psychol. Bull. **134**(3), 383 (2008)
61. Vance, A.: The fog of warnings: how non-essential notifications blur with security warnings. In: Symposium on Usable Privacy and Security (SOUPS) (2019)
62. Vance, A., Kirwan, B., Bjornn, D., Jenkins, J., Anderson, B.B.: What do we really know about how habituation to warnings occurs over time? a longitudinal FMRI study of habituation and polymorphic warnings. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp. 2215–2227 (2017)
63. Warkentin, M., Xu, Z., Mutchler, L.A.: I'm safer than you: the role of optimism bias in personal it risk assessments. In: Proceedings of, pp. 1–32 (2013)
64. Weijters, B., Baumgartner, H.: Misresponse to reversed and negated items in surveys: a review. J. Mark. Res. **49**(5), 737–747 (2012)
65. Wu, M., Miller, R.C., Garfinkel, S.L.: Do security toolbars actually prevent phishing attacks? In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 601–610 (2006)
66. Wyatt, T.: Understanding the process of contextualization. Multicultural Learn. Teach. **10**(1), 111–132 (2015)
67. Xu, H., Rosson, M.B., Carroll, J.M.: Increasing the persuasiveness of it security communication: effects of fear appeals and self-view. In: Workshop on Usable

IT Security Management, Symposium on Usable Privacy and Security (SOUPS), Pittsburgh, PA (2007)
68. Zaaba, Z.F., Boon, T.K.: Examination on usability issues of security warning dialogs. Age **18**(25), 26–35 (2015)
69. Zaaba, Z.F., Lim Xin Yi, C., Amran, A., Omar, M.A.: Harnessing the challenges and solutions to improve security warnings: a review. Sensors **21**(21), 7313 (2021)
70. Zaaba, Z., Furnell, S., Dowland, P., Stengel, I.: Assessing the usability of application-level security warnings. In: Proceedings of the 11th Security Conference (Security Assurance & Privacy) (2012)