Does Context Influence Responses to Firewall Warnings?

Muhammad Mahmoud
Department of Systems and Computer Engineering
Carleton University
Ottawa, ON, K1S 5B6, Canada
Email: muhammad@ieee.org

Sonia Chiasson
School of Computer Science
Carleton University
Ottawa, ON, K1S 5B6, Canada
Email: chiasson@scs.carleton.ca

Ashraf Matrawy
Department of Systems and Computer Engineering
Carleton University
Ottawa, ON, K1S 5B6, Canada
Email: amatrawy@sce.carleton.ca

***Abstract*—Firewall warnings are only effective if users can respond to them in a secure and effective manner. In this paper, we define what the user was doing while receiving the warning message as *the context*. We evaluate whether either the *context* in which the user receives the warning or the *content* of the warning message affects users' response to the warning message. We ran a user study with 56 participants via an online survey. Our results show that the context in which the warning message appears has no influence on the users' responses. We further confirm users base their responses to a warning on the content of the warning message itself. We also show that users understood the need for such warning messages and wanted to be involved in the decision-making process as long as they were not interrupted too frequently.**

I. Introduction

While automated security decisions are considered more accurate than those made by humans [1], and while humans are considered to be "the weakest link in the security chain" [2] [3], there are many reasons and limitations why we cannot eliminate the human input completely [4]. These limitations include the situational and social dependencies where factors external to the security system must be considered, accommodation of end-user values which vary with different individuals, and automation failures where users must resolve cases where the system makes an error or is unable to determine the appropriate course of action. [4].

Given these circumstances, computer system designers often resort to displaying security warnings asking users to choose a course of action. Warnings are "*communication that alert users to take immediate action to avoid a hazard*" [5]. Users' reaction to warnings has been the focus of ongoing research. This is due to the important role this response plays in fighting a number of cyber security threats and perhaps in dealing with problems such as online fraud.

In earlier work [6] Mahmoud and Matrawy proposed an approach to mitigate the effect of IRC botnets by disrupting their command and control (C&C) communications. Disrupting such botnets is important in the prevention of e-crime because botnets are responsible for most of email spam, identity theft, online phishing, online fraud, adware, spyware, and DDoS attacks [7]. The approach relies on intercepting C&C protocol communication (e.g. IRC) and engaging the computer's user by showing a warning and asking the user if (s)he initiated/trusted that communication. The success of this approach relies, among other things, on the correct interpretation of warning messages by users and on the users selecting a proper response. Relying on the user to properly understand the warnings and make the right decision all the time cannot be guaranteed. In addition to many sources of decision errors, users' understanding of the warnings and responses could be affected when the users are busy with other tasks.

978-1-4673-2543-1/12/$31.00 © 2012 IEEE

In this paper, we evaluate two research question related to users' responses to warning messages. First we define what the user was doing when (s)he got the warning message as *the context* and then we evaluate whether either the *context* in which the user receives the warning or the *content* of the warning message affects users' response to the warning message. Second, we re-examine the level of users' understanding of the warning messages and the risks associated with the warning and their responses to them.

Contribution: Based on the statistical analysis of results from our user study that involved 56 users, we demonstrate for the first time (to the best of our knowledge) that the context in which the warning message appears has no influence on the users' responses. We further confirm that users base their response to a warning on the content of the warning message itself. In other words, we show that users respond differently to different security warning messages regardless of the context. We show that the majority of the users read and understood the warnings. Furthermore, the majority of the users reported that they wanted at least some involvement in the decision making process.

Paper Outline: This paper is outlined as follows: Section II is a brief related work on security warnings. In section III, we discuss the study design, methodology, questionnaire scenarios and requirements. While sections IV includes the results of the study and statistical analysis of the effect of the warning message and the effect of the context on the users' responses, section V includes results and analysis on users'

understanding and their feedback on the warnings. Finally, in section VI, we discuss the study results and limitations, and provide concluding remarks.

II. Related Work on Security Warnings

Security warnings should grab and maintain the user's attention, persuade the user to take an action, and provide clear instruction to the correct course of action [8] [9] [10].

According to Egelman et al. [9]'s user study of phishing warnings, a security warning needs to satisfy the following criteria to be effective:

- It needs to capture the user's attention by interrupting the primary task.
- It must capture the user's attention long enough to allow them to attempt comprehension.
- It should "fail safely". i.e. if the user did not read or understand the warning and clicked the default, he/she should be protected.
- It must give the user indication of danger and clearly suggest action(s).
- It should not be similar to less serious warnings to prevent habituation.

Morgan et al. [11] devised a decision-analytic framework based on mental models for warning recipients. Warning systems need to communicate risk information to the warning recipients. This communication should add missing information, tune the recipient's knowledge to ensure that it is not too narrow or too general, and remove misconceptions. Though this mental model approach was successful in many fields, such as in environmental and medical risk communication, it did not improve computer security because in computer security risk communication is typically based on the mental models of experts, which are different from the mental models of the majority users [12].

In related usable security work, Cranor [5] proposed a conceptual security framework (human-in-the-loop) to help analyze the human role in security systems. It is intended to help understand non-malicious human behavior in performing security-related tasks and user responses to warnings. This framework is based on a communication-processing model to facilitate security-related communication to a non-expert human.

Several studies of security warnings have focused on SSL warnings displayed in the browser. Biddle et al. [13] studied the interfaces used to display SSL certificate-related information to users. Their results suggest that there is a need to redesign these interfaces and showed some of these interfaces provide what may be viewed as misleading information.

Sotirakopoulos et al. [14] carried out a study to determine the effectiveness of SSL warnings. Interestingly, their main result has greater methodological impact than SSL implications. Their study showed that it is challenging to observe users' security behavior in the lab environment because either security sensitive participants refuse to participate, or because participants feel safe that the study has gone through some ethics approval process, and/or because they are using a lab computer, not their own.

Other work has focused on firewall warnings. Raja et al. [12] designed interactive firewall warnings that represent the action of the firewall in a physical security metaphor. They conducted a comparison study evaluating whether the warnings were understood by participants and to see if these were successful in communicating risk to the participants. They found that their physical security metaphor warnings scored higher than commercial warning systems in communicating hazard to the participants.

To explore beyond these conditions, in this paper we study the effect of the context (as defined in Section I) on users' response to warnings. This has not been studied before to the best of our knowledge.

III. Study Design

This study was reviewed and approved by our institutional ethics review board. The study was carried out through an online survey describing nine scenarios where users may see security warnings. Participants were asked to imagine that each scenario occurred and respond as they would if this occurred in real life. These scenarios presented different combinations of warning messages and contexts (what the user was doing when the warning message appeared). The only difference between scenarios was the context and the content of the warning message. The design of the warnings was kept consistent in color and appearance. An example warning is provided in Fig 1.

A. Methodology

This study aimed to answer the following two research questions:

RQ1: Does the context of what the user is doing affect his/her response to the warning messages?

RQ2: Do users understand the warning messages and the risks in ignoring them?

The study questionnaire was administered using LimeSurvey[1]. Participants completed the questionnaire remotely, using tokens provided by the experimenter. As Fig 1 shows, participants were asked to consider the following situation: while using *some application*, their Internet security software displays a warning message. A picture of the warning message is provided to the user. The warning informs the user that an application is trying to gain access to the Internet. For every scenario, participants were asked to answer six questions intended to address our research questions:

To address RQ1 and collect data on whether context or content affected user responses, we asked:

- What would be your choice? (Allow / Always Allow / Block / Always Block).

To address RQ2 and gain insight into users' understanding of the warnings and risks, we asked:

[1] http://www.limesurvey.org/

Consider a scenario where while you are using ***MS Word***, Internet security software popped up the following warning messages.

Fig. 1: Example warning message and instructions for scenario 1.

- After reading the text in the warning message, please explain in your own words what it means.
- Before making your decision, how much of the warning message did you read? (The entire message / Most of the message / About half of the message / Only a few words / None of the message).
- How much of the warning did you understand (I fully understood the warning / I partially understood the warning / I barely understood the warning / I did not understand the warning at all).
- After reading this warning message, did you believe there may be a risk in ignoring the warning? (Yes / No).
- Rate the amount of risk you feel you were warned against (1: Very low risk . . . 5: Very high risk).

After completing the scenarios, participants were asked additional questions common to all scenarios. For example, they were asked to explain what they expected to happen when choosing a specific response to the warnings. They were also asked about the ease of use and usefulness of such a warning system, and whether they would adopt such system. Finally, participants were asked some questions to determine their demographics and experience level. We did not collect any personally identifiable data.

B. Scenarios

Users were warned against four applications trying to send messages to the Internet. These warning messages contained information about the following executables:

- an application that could be mistaken for part of the OS (system.exe).
- an application with a suspicious name (sD526Fx92.exe).
- a well known application (MSN Messenger).
- an unknown, but familiar-sounding application (SkyTorrent).

The warnings appeared in one of three contextual situations. These situations stipulated that the user was currently using:

- an application that did not require Internet access (MS Word).
- an application that required Internet access (Internet Explorer).
- an Internet chat application (MSN Messenger).

In the rest of the paper these application are refereed to as Word, IE, and MSN respectively.

These warnings and contextual situations were paired to create 9 scenarios. We will use the notation *warning(context)* throughout the rest of this paper. For example, if a user received a warning message that System.exe is trying to access the Internet while (s)he is using MS Word, the notation *System(Word)* will be used.

The following nine permutations (*scenarios*) were investigated:

Scenario 1: *Sys(Word)* warns about System.exe while using MS Word.
Scenario 2: *sD5(IE)* warns about sD526Fx92.exe while using IE.
Scenario 3: *MSN(MSN)* warns about MSN Messenger while using MSN Messenger.
Scenario 4: *MSN(Word)* warns about MSN Messenger while using MS Word.
Scenario 5: *Sys(MSN)* warns about System.exe while using MSN Messenger.
Scenario 6: *MSN(IE)* warns about MSN Messenger while using IE.
Scenario 7: *sD5(Word)* warns about sD526Fx92.exe while using MS Word.
Scenario 8: *Sky(IE)* warns about SkyTorrent while using IE.
Scenario 9: *sD5(MSN)* warns about sD526Fx92.exe while using MSN Messenger.

As an example, Fig 1 shows the warning message for scenario 1.

C. Recruitment and Participants

Participants were recruited through posters and emails. Fifty six participants (35 male, 20 female, 1 unspecified) from diverse age groups agreed to participate: 25% of participants were between 19 and 29 years of age, 38% between 30 and 39, 16% between 40 and 49, 21% were 50 years or older. Approximately 23% of our participants were undergraduates, 64% were graduate students, and 13% were employed in various fields. Participants were asked to rate their computer experience on a scale of 1 (least) to 5 (most) experienced. The average score was 3.4 with a median of 3.

This sample is not necessarily representative of the whole population. It does, however, reflect the younger, technology literate segment of the population – a group that would frequently need to deal with these types of warnings.

For our analysis, we examined several aspects of the data to address RQ1 and RQ2, our two research questions. The following two sections will focus on each research question separately.

IV. Results for Research Question 1 (RQ1)

For RQ1, our primary interest was users' choice of action for each scenario, following by analysis of whether content or context impacted user choice.

A. User Choice

For each of the 9 scenarios presented to users in the survey, users chose whether they would *always allow*, *allow* (for this instance), *block* (for this instance), or *always block* the request presented in the warning. Table I illustrates users' choice in this respect. The scenarios are numbered according to the order that they were presented to users. We found that users did vary their responses to the warning messages. In this section, we focus on a few interesting situations that arose where the results have security implications.

1) "Official-sounding application name": As scenarios 1 and 5 show, if an application with a name that could be mistaken for part of the OS (system.exe) asked for permission to access the Internet, about 22% of the users *allowed* or *always allowed* the communication. This is important because it points to a potential attack vector. If a malicious application uses an "official-sounding" name, a non-trivial number of users could be tricked into trusting it.

2) "Odd-sounding application name": On the other hand, if an application with an odd name (sD526Fx92.exe) asked to access the Internet (scenarios 2, 7 and 9) most users chose to *block* or *always block* it. However, some 9% of participants allowed this communication request if the warning came while they were accessing the Internet with another application, as in scenario 2. Though this is a small percentage, it shows that malicious applications may potentially increase their chances of getting access to the Internet if they synchronize their communication with user's activities that require access to the Internet.

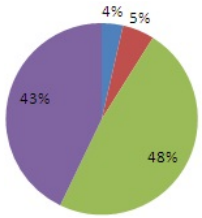
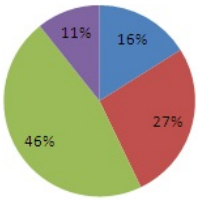
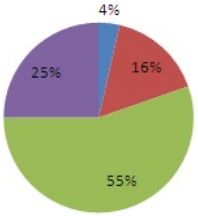
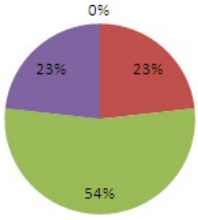
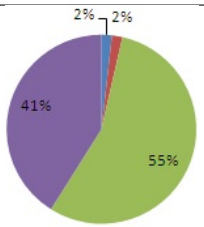
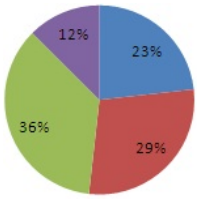
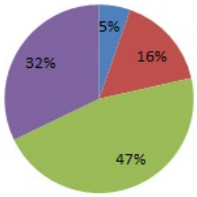
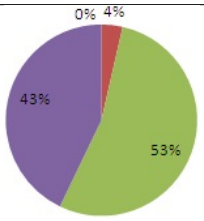
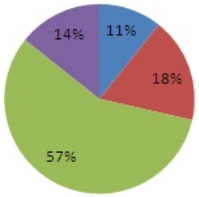
3) "Familiar-sounding application name": When an unknown application with a familiar looking name (SkyTorrent) asked to access the Internet (scenario 8), it had a 20% chance of success. Because this application has the word Torrent in its name, several users later explained that they thought it was their torrent application trying to access the Internet and *allowed* or *always allowed* access. Malicious applications may increase their chances of getting access to the Internet if they use a name that is similar to familiar applications.

4) "Known application name": In the scenarios (3, 4, 6) where a well-known application (MSN Messenger) tried to access the Internet, we have the most varied responses. Approximately 19% of users *allowed* or *always allowed* MSN communication because they said they simply trusted this application. This percentage increases to 46% if the request came while the users were already using Internet Explorer. However, if users were already using the MSN Messenger and received a warning message about it (scenario 3), 48% *blocked* or *always blocked* it. Users in this category later explained that they thought that they should not get such a warning if they were already using this same application. They explained that because they were warned, then some other application must have imitated MSN Messenger or maybe MSN Messenger was doing something illegal. This may provide support for a white list of legitimate applications to avoid such confusion if a security warning system was implemented. From a methodological perspective, this also provides evidence that users were in fact reading and considering the application context instructions for each of the scenarios.

B. Statistical Tests

To more rigorously assess whether context or content impacted users' responses to the warning system, we performed statistical analysis using the R programming language. We designed an experiment with a complete block design, where each participant saw each scenario. The non-parametric test that fits this situation is the *Friedman test*.

TABLE I: User responses to the warning messages from the 9 scenarios. Scenarios are numbered in the order that they were displayed to users.

Always Allow / Allow / Block / Always Block	Application trying to access the internet *(Content)*			
User working on *(Context)*	System.exe (Sys)	sD526Fx92.exe (Sd5)	MSN Messenger (MSN)	SkyTorrent (Sky)
Internet Explorer		 Scenario 2	 Scenario 6	 Scenario 8
MSN Messenger	 Scenario 5	 Scenario 9	 Scenario 3	
MS Word	 Scenario 1	 Scenario 7	 Scenario 4	

The Friedman test is a non-parametric test for complete block designs to detect differences between treatments (warning messages scenarios in our study). The Friedman test has the advantage that within subject dependence is considered (i.e., does not assume independence). However, the study output needs to be ordinal (e.g. scores) [15] [16]. Therefore, we used the following ranks/scores for our study with respect to users' responses to the warning messages, ordered from safest to riskiest behavior.

```
Always Block     4
Block            3
Allow            2
Always Allow     1
```

Where significant effects were found, we ran *post-hoc tests* to determine the reason for the significant effect by comparing relevant treatment pairs. The *Wilcoxon rank sum test* (also known as Mann-Whitney-Wilcoxon (MWW) test) is a paired difference test used for such post-hoc analysis [15] [16]. In all cases, we considered $p < 0.05$ as a statistically significant result indicating that the differences between conditions were unlikely to have occurred by chance.

C. Effect of Warning Message Content

We separately analyzed the effect of different warning messages for each of the three application contexts. Table II summarizes the results of the three Friedman tests. For all three application contexts, Friedman tests revealed a significant effect of warning on user responses. This indicates that users' responses varied based specifically on the content of the warning messages.

TABLE II: Significance tests comparing the effect of different warning messages for each context. Items in bold represent statistically significant results.

Application Context	Friedman Test
Word	$\boldsymbol{\chi^2(2) = 17.37, p < 0.05}$
IE	$\boldsymbol{\chi^2(2) = 23.92, p < 0.05}$
MSN	$\boldsymbol{\chi^2(2) = 20.76, p < 0.05}$

We next performed post-hoc Wilcoxon rank sum tests to identify which pairs of messages led to the significant dif-

ferences. Table III shows only the relevant pairs for different warning messages. For example, using MS Word as the context, there are three possible warning messages (system.exe, sD526Fx92.exe, MSN Messenger, as shown in Table I). This gives three possible pairings: [Sys(Word) – sD5(Word)], [Sys(Word) – MSN(Word)], and [sD5(Word) – MSN(Word)]. Table III shows results of the post-hoc significance tests for each pairing. Significant differences were found between all pairs ($p < 0.05$) except for the [MSN(Word) – Sys(Word)] pair, indicating that users were in fact reacting to the content of each warning message.

TABLE III: Post-hoc Wilcoxon tests identifying differences between pairs of warning messages for each application context. Results in bold indicate significant differences.

Warning Pairs *(Content)*	Running Application *(Context)*		
	IE	MSN	Word
sD5 – MSN	$\mathbf{p < 0.05}$	$\mathbf{p < 0.05}$	$\mathbf{p < 0.05}$
sD5 – Sys	–	$\mathbf{p < 0.05}$	$\mathbf{p < 0.05}$
sD5 –Sky	$\mathbf{p < 0.05}$	–	–
MSN – Sys	–	$\mathbf{p < 0.05}$	$p > 0.05$
MSN – Sky	$\mathbf{p < 0.05}$	–	–

D. Effect of Application Context

To investigate whether the context in which the warning messages were displayed impacted users' responses to the warnings, we compared user responses for each type of warning when shown in different contexts. Three of our warnings (Sys, Sd5, and MSN) were shown in different contexts, so we ran Friedman tests for each type of warning to assess the effect of context. Table IV shows the results of these tests. In all three cases, Friedman test revealed no significant effect of context on user responses, indicating that users have similar responses to the same warning message regardless of the current application context. Since no significant results were found, further post-hoc analysis is unnecessary.

TABLE IV: Significance tests comparing the effect of application context for three different warning messages. None of the tests revealed statistically significant results.

Warning Message Content	Friedman Test
sD5	$\chi^2(2) = 1.41, p > 0.05$
Sys	$\chi^2(1) = 0.8, p > 0.05$
MSN	$\chi^2(2) = 3.18, p > 0.05$

E. Summary of Research Question 1 Results

To summarize, we found that the content of the warning messages had a significant impact on users' decisions on whether to allow the mentioned application access to the internet. We further found that the context in which the warning message appeared did not have any impact on users' decisions.

Examining specific scenarios, we identified potential risks that security designers may want to consider when implementing a warning system. Users make decisions based the information at hand and may be misled if attackers are able to imitate familiar- or official-sounding executable names.

V. Results for Research Question 2 (RQ2)

For RQ2, our primary interest was in exploring what users understood about the risks, how they made their decisions, and their willingness to use a warning system, rather than statistically comparing the different scenarios. Users were asked the questions related to RQ2 after every scenario.

A. User Interpretation of Warning Messages

When users were asked to explain the meaning of the warning messages in their own words, they showed a great deal of understanding. Following are some sample of user explanations:

- "*It means that a program 'system.exe' is attempting to send information via the internet from my computer. The message also asks me if I trust the program and to make a decision on whether or not to let it perform this operation.*"
- "*A program, which seems like a system application, is restricted by the windows firewall to reach the Internet and needs my permission to do so. It is recommended that I not allow it to do so.*"
- "*It means that some program is trying to execute. Given that I'm in IE, I would suspect it is a trojan on a website I'm on.*"

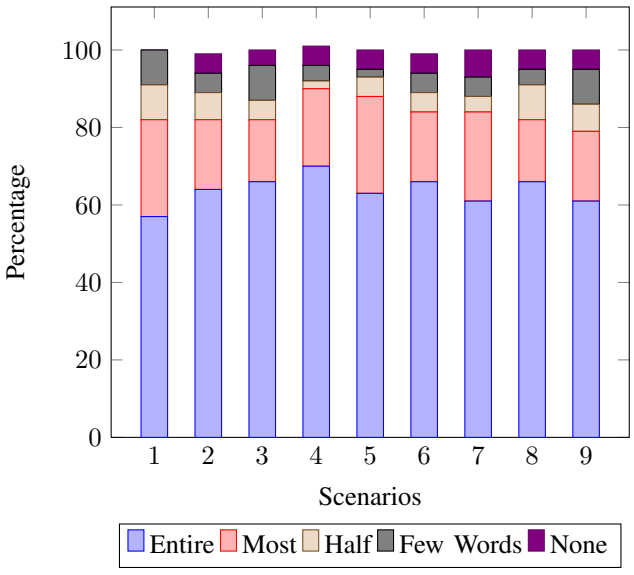


Fig. 2: How much of the warning is read?

B. How Much of the Warning Message was Read

We asked participants to report how much of each warning message they read before deciding how to respond to the

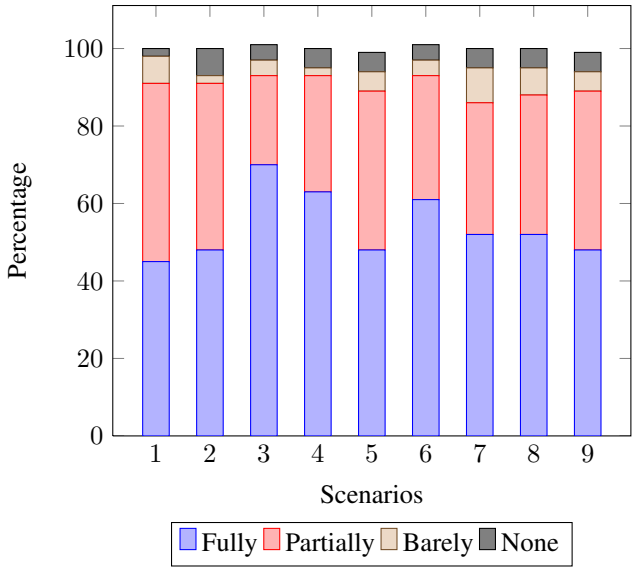


Fig. 3: How much of the warning is understood?

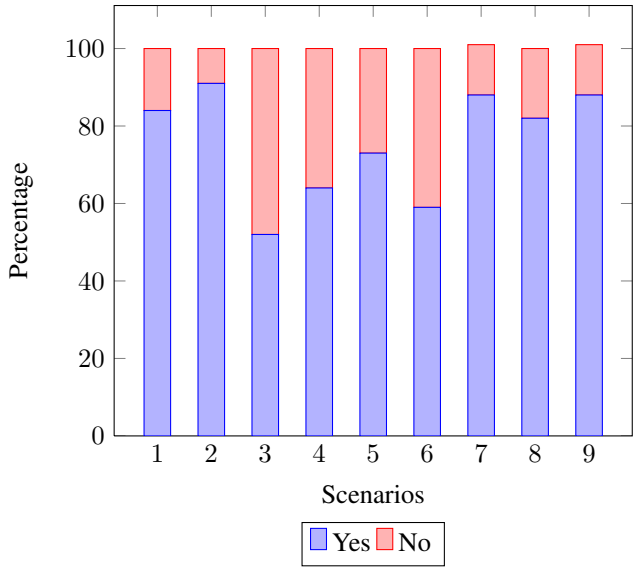


Fig. 4: Is there a risk in ignoring the warning?

warning message. As shown in Fig 2, over 80% of users reported reading at least "most" of the each warning message. Figures 2 to 5 use stacked bar graphs to show the percentage of users who responded with each potential answer.

C. How Much of the Warning was Understood

We further asked users to report how much of the warning they understood. User responses are summarized in Fig 3. We see that about half of users reported full understanding of the warning message, while approximately 90% of users reported at least partially understanding the meaning of the message. While these are self-reports that cannot be verified, users' free-form interpretation of the warnings seem to corroborate that most users did in fact read and pay attention to the text of the warnings.

D. Perceived Risk in Ignoring the Warning

Participants were asked whether there was a risk in ignoring each warning message and could reply with Yes/No. Responses are summarized in Fig 4. A response of Yes indicates that users felt there was a risk, while No indicates that users were unconcerned about this potential scenario. As show in the figure, the majority of users thought that the warning was about a legitimate threat. Users were least concerned about warnings involving MSN Messenger.

E. Perceived Level of Risk Conveyed

For each scenario, users were asked to rate the level of risk communicated by the warning message on a scale of 1 (low) to 5 (high). Their responses are included in Fig 5. We see a divergent set of answers for each scenario, with users expressing varying levels of concern. No clear pattern emerges in the responses, except that users seem more concerned about scenarios 2, 7, and 9 — indicating that users were suspicious of the sD5 executable.

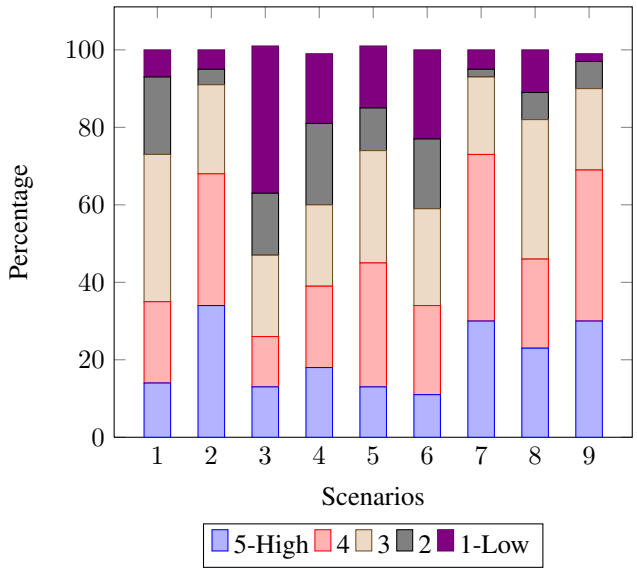


Fig. 5: Risk level communicated

F. General Feedback about the Warning System

At the end of the survey, participants were asked a series of questions pertaining to the warning system in general.

1) Level of Automation: Users were asked whether the system should: (1) show no warnings and make all decisions, (2) show some warnings and sometimes make a "best guess" decision, or (3) show all warnings. As shown in Fig 6, only 7% of users preferred a fully automated system. The remaining 93% reported that they wanted at least some involvement in the decision making process.

2) Acceptable Number of Warnings: As a follow-up question, users were asked how many warnings they would be comfortable in receiving on their system per week. Figure 7

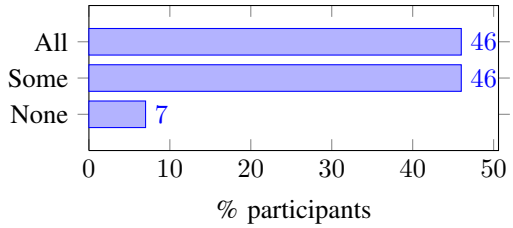


Fig. 6: Should the system show (All) warnings, (Some) warnings, or (None)?

summarizes their responses. From the figure, we see that 13% of users preferred to receive no warnings, while a further 52% would find 1-5 warnings acceptable. This means that any warning system generating more than 5 warnings per week may not be effective because users may develop coping mechanisms such as disabling the system or making arbitrary decisions.

3) Ease of Understanding the Warning System: We asked users whether they found the warning system easy to understand. Fig 8 shows their responses. Users generally had a neutral-to-favorable opinion of the system, with 74% of users giving it a rating of at least 3/5 and only 5% of users reporting that the system was "not easy" (1).

4) Perceived Increase in Security: Participants were asked if they think that such warning system would increase the security of their computer. Fig 9 illustrates their responses. The majority of users (71%) agreed that a warning system would increase the security of their computer system.

5) Overall Impressions of the System: We gave gave users a list of adjectives and asked them select those that best described the warning system. Users responses are provided in Fig 10. Responses are generally positive, with "helpful", "useful", and "informative" receiving the three highest scores, while "useless", "confusing", and "annoying" received the three lowest scores.

VI. Discussion and Conclusion

Security software must often rely on users to make security decisions. In these cases, the security of the system relies on users' correct choice among the available options when a warning is displayed. Users are typically not experts in security and are often busy with other primary tasks when they are interrupted by the security warning. At this point, users must make a decision based on the available information, which includes the content of the warning message and the context in which the warning appeared. For example, a user who is currently editing a local text file may be more suspicious of a warning about an application requiring internet access than if that same user is actively viewing web content.

In earlier studies [9][12][13][14], researchers focused on the *content* of the warning messages and finding the best ways to present potential threats to users. In this work, we investigated whether the context in which the message is received impacted their choice of action. For comparison,

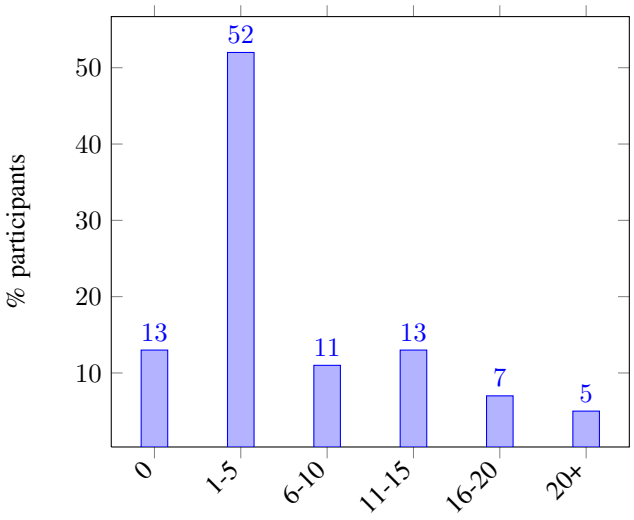


Fig. 7: In a regular week, how many times would you be willing to answer a warning such as these?

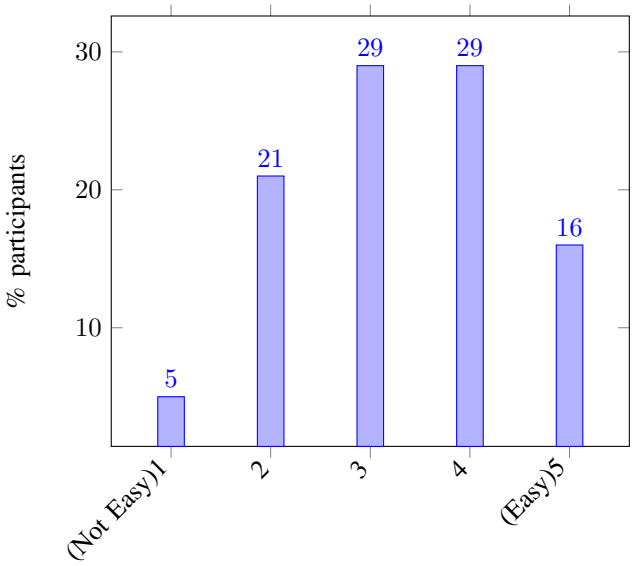


Fig. 8: How easy to understand is the warning system?

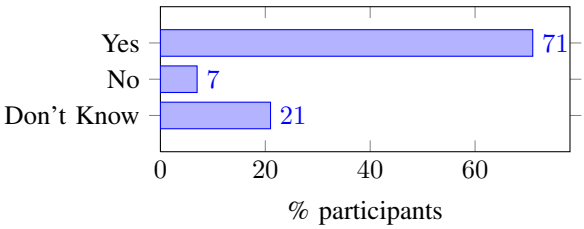


Fig. 9: Do you think that such warning system would increase the security of your computer?

we also examined whether the content of the message had any effect and questioned users about their understanding

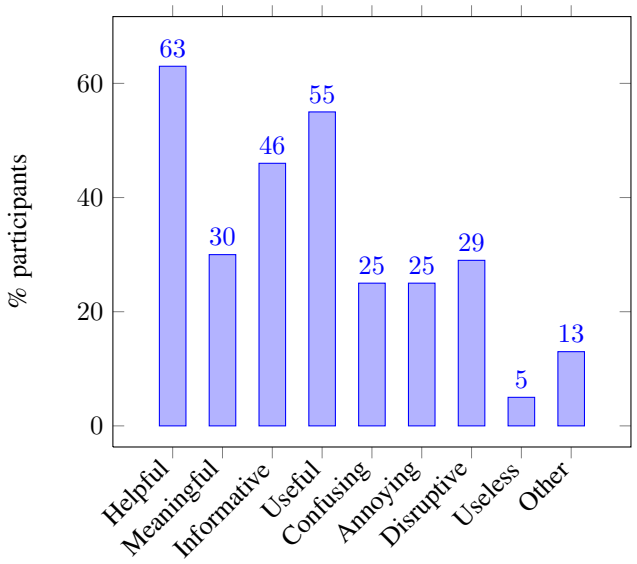


Fig. 10: User choice of words that best describe the system.

of the risks involved as well as their overall acceptance of security systems that rely on user interaction to determine what decision should be taken when faced with a threat. Keep in mind that our definition for context and the scenarios examined are simplified. Other circumstances like the user's state of mind, whether the user is in a hurry to get things done, how many other warning messages s/he has already dealt with that day, etc. could also impact user decisions. Our results may represent a "best case" scenario. However, we wanted early indications to help with the system design

We chose an online survey as our test instrument rather than a lab study for several reasons. Surveys may not completely reflect users' real life behavior, but neither do lab studies [14]. A long-term field study would be ideal, however, this requires significant development and reasonable assurance of the security properties of the system before such a study is possible. We wanted to inform the design of our future system before building it and decided that the insight gained from a survey would be beneficial. Secondly, we felt that a survey completed remotely would be less likely to influence users' responses than a lab study where users may modify their behavior because of the artificial setting. In analyzing users' free-form responses and comments, we believe that users seriously considered the questions and were aware of both the context and the warning content for each scenario. As a next step, we are planning to develop a prototype firewall application that would display warnings to users and record their responses. We intend to deploy the prototype for usage over an extended period of time.

Returning to our first research question: does the context in which the warning message appears influence users' response? Statistically, the answer is no. User decisions were not affected by context. Although, in some cases, as in section IV-A4, some of the user comments seem to indicate that they were being influenced at least to some extent by context. Further studies would be needed to more fully understand what is happening. We did find, however, that the message content impacted users' responses to the warning messages. Users were more likely to grant permission to known executables and least likely to provide access to odd or suspicious sounding executable names.

For our second research question, we were particularly interested in exploring what users understood about the risks, how they made their decisions, and their willingness to use a warning system, rather than statistically comparing the different scenarios. We found that many users were able to verbalize their understanding and showed reasonable mental models of how the system would work. Users reported reading most of the warning messages, understanding their meaning, preferred to be involved in the decision-making process rather than having the system be fully automated, and felt that a warning system would improve the security of their computer. They reported a fairly low threshold for the number of warnings they would be willing to answer in a given week, a consideration that system designers should to remember.

In conclusion, fighting e-crime is a daunting task and one that might be difficult to fully automate. Given these circumstances we will likely need to continue relying on security warnings and user intervention. This paper provides the first evidence that the context in which warning messages are displayed does not affect user decisions but that the content of the warning message does have an impact. We also show that users understood the need for such warning messages and wanted to be involved in the decision-making process as long as they were not interrupted too frequently.

Acknowledgment

Muhammad Mahmoud acknowledges King Fahd University of Petroleum & Minerals for its support. Sonia Chiasson acknowledges NSERC funding for her Canada Research Chair in Human-Oriented Computer Security and acknowledges support from NSERC ISSNet. Ashraf Matrawy acknowledges funding through an NSERC Discovery Grant (2007-2012).

References

[1] I. Flechais, J. Riegelsberger, and M. A. Sasse, "Divide and Conquer: The Role of Trust and Assurance in the Design of Secure Socio-Technical Systems," in *Proceedings of the 2005 workshop on New security paradigms*, ser. NSPW '05. New York, NY, USA: ACM, 2005, pp. 33–41. [Online]. Available: http://doi.acm.org/10.1145/1146269.1146280

[2] B. Schneier, *Secrets and Lies: Digital Security in a Networked World*. John Wiley & Sons, 2004. [Online]. Available: http://dl.acm.org/citation.cfm?id=994962

[3] M. J. Kalsher and W. J. Kevin, "Behavioral Compliance: Theory, Methodology, and Results," in *Handbook of Warnings*, M. S. Wogalter, Ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2006, pp. 313–331, book, Chapter.

[4] W. K. Edwards, E. S. Poole, and J. Stoll, "Security automation considered harmful?" in *Proceedings of the 2007 Workshop on New Security Paradigms*, ser. NSPW '07. New York, NY, USA: ACM, 2008, pp. 33–42. [Online]. Available: http://doi.acm.org/10.1145/1600176.1600182

[5] L. F. Cranor, "A Framework for Reasoning About the Human in the Loop," in *Proceedings of the 1st Conference on Usability, Psychology, and Security*, ser. UPSEC'08. Berkeley, CA, USA: USENIX Association, 2008, pp. 1:1–1:15. [Online]. Available: http://usenix.org/events/upsec08/tech/full_papers/cranor/cranor.pdf

[6] M. Y. Mahmoud and A. Matrawy, "IRC BotChallenger: Creating Botnet-Resilient Networks," *Journal of Information Assurance and Security*, vol. 6, pp. 503–511, 2011.

[7] J. R. Binkley and S. Singh, "An Algorithm for Anomaly-Based Botnet Detection," in *SRUTI'06: Proceedings of the 2nd conference on Steps to Reducing Unwanted Traffic on the Internet*. Berkeley, CA, USA: USENIX Association, 2006, pp. 7 –7.

[8] M. S. Wogalter, "Purposes and Scope of Warnings," in *Handbook of Warnings*, M. S. Wogalter, Ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2006, pp. 3–9, book, Chapter.

[9] S. Egelman, L. F. Cranor, and J. Hong, "You've Been Warned: An Empirical Study of the Effectiveness of Web Browser Phishing Warnings," in *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*. New York, NY, USA: ACM, April 2008, pp. 1065 –1074. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1357054.1357219

[10] M. S. Wogalter and W. J. Vigilate, "Attention Switch and Maintenance," in *Handbook of Warnings*, M. S. Wogalter, Ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2006, pp. 245–265, book, Chapter.

[11] G. M. Morgan, B. Fischhoff, A. Bostrom, and C. J. Atman, *Risk Communication: A Mental Models Approach*. Cambridge University Press, July 2002. [Online]. Available: http://books.google.com/books?vid=ISBN0521002567

[12] F. Raja, K. Hawkey, S. Hsu, K.-L. C. Wang, and K. Beznosov, "A Brick Wall, a Locked Door, and a Bandit: A Physical Security Metaphor For Firewall Warnings," in *Proceedings of the Seventh Symposium on Usable Privacy and Security*, ser. SOUPS '11. New York, NY, USA: ACM, 2011, pp. 1:1–1:20. [Online]. Available: http://doi.acm.org/10.1145/2078827.2078829

[13] R. Biddle, P. C. van Oorschot, A. S. Patrick, J. Sobey, and T. Whalen, "Browser Interfaces and Extended Validation SSL Certificates: An Empirical Study," in *Proceedings of the 2009 ACM workshop on Cloud computing security*, ser. CCSW '09. New York, NY, USA: ACM, 2009, pp. 19–30. [Online]. Available: http://doi.acm.org/10.1145/1655008.1655012

[14] A. Sotirakopoulos, K. Hawkey, and K. Beznosov, "On the Challenges in Usable Security Lab Studies: Lessons Learned from Replicating a Study on SSL Warnings," in *Proceedings of Symposium on Usable Privacy and Security*, ser. SOUPS '11. New York, NY, USA: ACM, June 2011, pp. 3:1–3:18. [Online]. Available: http://doi.acm.org/10.1145/2078827.2078831

[15] M. Hollander and D. A. Wolfe, *Nonparametric Statistical Methods*, 2nd ed. John Wiley & Sons, 1999. [Online]. Available: http://books.google.com/books?vid=ISBN0471190454

[16] W. J. Conover, *Practical Nonparametric Statistics*, 3rd ed. John Wiley & Sons, 1999. [Online]. Available: http://books.google.com/books?vid=ISBN0471160687