The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/2056-4961.htm

How do non experts think about cyber attack consequences?

Keith S. Jones, Natalie R. Lodinger and Benjamin P. Widlus
Department of Psychological Sciences, Texas Tech University, Lubbock, Texas, USA

Akbar Siami Namin
Department of Computer Science, Texas Tech University, Lubbock, Texas, USA, and

Emily Maw and Miriam E. Armstrong
Department of Psychological Sciences, Texas Tech University, Lubbock, Texas, USA

Received 9 November 2020
Revised 13 October 2021
7 January 2022
Accepted 9 January 2022

Abstract

Purpose – Nonexperts do not always follow the advice in cybersecurity warning messages. To increase compliance, it is recommended that warning messages use nontechnical language, describe how the cyberattack will affect the user personally and do so in a way that aligns with how the user thinks about cyberattacks. Implementing those recommendations requires an understanding of how nonexperts think about cyberattack consequences. Unfortunately, research has yet to reveal nonexperts' thinking about cyberattack consequences. Toward that end, the purpose of this study was to examine how nonexperts think about cyberattack consequences.

Design/methodology/approach – Nonexperts sorted cyberattack consequences based on perceived similarity and labeled each group based on the reason those grouped consequences were perceived to be similar. Participants' labels were analyzed to understand the general themes and the specific features that are present in nonexperts' thinking.

Findings – The results suggested participants mainly thought about cyberattack consequences in terms of what the attacker is doing and what will be affected. Further, the results suggested participants thought about certain aspects of the consequences in concrete terms and other aspects of the consequences in general terms.

Originality/value – This research illuminates how nonexperts think about cyberattack consequences. This paper also reveals what aspects of nonexperts' thinking are more or less concrete and identifies specific terminology that can be used to describe aspects that fall into each case. Such information allows one to align warning messages to nonexperts' thinking in more nuanced ways than would otherwise be possible.

Keywords Mental models, Cyberattack consequences, Warning message design, Cybersecurity

Paper type Research paper

1. Introduction

Users often receive messages warning them about potential cyberattacks and encouraging them to behave in certain ways to avoid them (Amran *et al.*, 2017; Zaaba *et al.*, 2016). For example, users are warned when they view a potentially fraudulent website and discouraged from entering personal information into it (Agrawal *et al.*, 2020; Akhawe and Felt, 2013; Egelman *et al.*, 2008).

This research was supported in part by the U.S. National Science Foundation (Award : 1564293). Opinions, findings, and conclusions are those of the authors and do not necessarily reflect the views of the NSF.

Information & Computer Security
© Emerald Publishing Limited
2056-4961
DOI 10.1108/ICS-11-2020-0184

Some warning messages also attempt to make users fear the cyberattack and its consequences (Sasse, 2015). However, such fear appeals are ineffective (Dupuis *et al.*, 2021; Sasse, 2015) and could cause those who fall victim to attacks to feel guilt and shame, which will likely degrade their well-being (Renaud *et al.*, 2021).

1.1 Nonexperts do not always comply with warning messages

Nonexperts do not always do what warning messages advise (for a review, see Jones *et al.*, 2021). For example, nonexperts decide to not guard their passwords (Weirich and Sasse, 2001), e-mail (Renaud *et al.*, 2014) or online privacy/cybersecurity (Kang *et al.*, 2015; Theofanos *et al.*, 2017), to not install operating system or application updates (Ion *et al.*, 2015; Vaniea *et al.*, 2014) and to not use encryption (Wu and Zappalla, 2018; Sombatruang *et al.*, 2020) or two-factor authentication (Ion *et al.*, 2015). Such decisions have deleterious effects on cybersecurity.

Why do nonexperts fail to comply with warning messages? Research suggests that there are many interrelated reasons. The following describes three key reasons.

1.1.1 Nonexperts do not fully understand warning messages. Nonexperts often do not fully understand warning messages (Bartsch *et al.*, 2013; Bauer *et al.*, 2013; Egelman *et al.*, 2008; Kauer *et al.*, 2012; Modic and Anderson, 2014). This is especially true when warning messages use technical terms (Bauer *et al.*, 2013). For example, most nonexperts would not understand:

> The server you are connected to is using a security certificate that cannot be verified. A certificate chain processed, but terminated in a root certificate which is not trusted by the trust provider

which is an actual warning message that was displayed to Windows users (Bauer *et al.*, 2013). Accordingly, nonexperts often do not understand what type of threat they may be experiencing, or what to do to avoid it (Egelman *et al.*, 2008). Without that knowledge, they cannot adequately understand the costs associated with noncompliance (Bartsch *et al.*, 2013; Bartsch and Volkamer, 2013; Kauer *et al.*, 2012) and likely do whatever allows them to get back to the task at hand.

1.1.2 Nonexperts do not always trust warning messages. Nonexperts think about cyberattacks in certain ways (Kauer *et al.*, 2013; Volkamer and Renaud, 2013; Wash, 2010; Wash and Rader, 2015). For example, some nonexperts think that viruses are buggy software (Kauer *et al.*, 2013; Wash, 2010). They reported that it is not necessary to have or regularly use antivirus software, be careful about what websites are visited or turn your computer off when not in use (Wash, 2010). Further, some nonexperts think cyberattacks would not target them. Some think cyberattacks only target those who are wealthy or important, i.e. "big fish" (Kang *et al.*, 2015; Kauer *et al.*, 2013; Prettyman *et al.*, 2015; Renaud *et al.*, 2014; Sasse *et al.*, 2001; Theofanos *et al.*, 2017; Ur *et al.*, 2016; Wash, 2010; Wash and Rader, 2015; Weirich and Sasse, 2001). Others think cyberattacks only target organizations (Kauer *et al.*, 2013; Wash, 2010).

Nonexperts do not trust warning messages that contradict what they think about cyberattacks (Bartsch and Volkamer, 2013; Ibrahim *et al.*, 2010). For example, nonexperts who think viruses are buggy software would not trust a message that warned against browsing a given website because a virus could infect their computer. Instead, they would likely ignore the warning and continue browsing the site.

1.1.3 Nonexperts think compliance will cost them. Nonexperts often think that doing what warning messages suggest will distract them from the task at hand (Dourish *et al.*, 2003; Hardee *et al.*, 2006; Kang *et al.*, 2015; Sasse *et al.*, 2001) and negatively affect their productivity (Vaniea *et al.*, 2014). Accordingly, they behave in an unsafe manner to achieve

their primary goals more quickly (Acar *et al.*, 2016; Herley, 2009). For example, some nonexperts do not think they have enough time to check whether every file and link they receive is legitimate, so they do not inspect them (West *et al.*, 2009). Further, some nonexperts think that operating system or application updates may include interface updates, which will require them to relearn how to use the software and thus make it more difficult to complete their work (Vaniea *et al.*, 2014). Consequently, they do not install software updates (Vaniea *et al.*, 2014).

1.2 Designing warning messages that address those issues

Researchers have offered warning message design recommendations that address the issues noted in Sections 1.1.1–1.1.3. Specifically, it is recommended that warning messages use nontechnical language (Bauer *et al.*, 2013), describe how the cyberattack will affect the user personally (Bartsch *et al.*, 2013; Kauer *et al.*, 2012) and do so in a way that aligns with how the user thinks about cyberattacks (Bartsch *et al.*, 2013). Doing so should increase the likelihood that nonexperts understand the warning message, understand the costs associated with noncompliance relative to the perceived costs of compliance and trust the warning message, which should increase compliance.

Implementing those recommendations requires an understanding of how nonexperts think about cyberattack consequences. Unfortunately, we do not currently understand that very well. To date, only one study has examined how nonexperts think about cyberattack consequences (Bartsch *et al.*, 2013) and that study concerned consequences of only one type of attack, i.e. what would happen if an attacker accessed information the participant had stored on a website. Bartsch *et al.* (2013) reported that nonexperts did not understand attack consequences, describing them abstractly, e.g. "comparatively bad" (Bartsch *et al.*, 2013, p. 5).

Bartsch *et al.* (2013) is a great start to understanding how nonexperts think about cyberattack consequences. However, more research is needed if we are to design warning messages that describe cyberattack consequences to nonexperts in a way that aligns with how they think about those consequences.

1.3 Present study

Toward that end, the present study examined how nonexperts think about cyberattack consequences. To do so, participants organized 50 cyberattack consequences into groups based on perceived similarity. Those consequences stemmed from a wide variety of cyberattack types. Participants then labeled each group to reflect the essence of that perceived similarity. Several studies have used similar techniques to understand how users think about various cybersecurity topics (Asgharpour *et al.*, 2007; Bartsch *et al.*, 2013; Jeong and Chiasson, 2020).

The terms used to label the groups provide insights into how nonexperts think about cyberattack consequences (Barnett, 2004; Fincher and Tenenberg, 2005). Accordingly, those labels were analyzed to understand the general themes and the specific features that are present in nonexperts' thinking. Similar approaches have been used when studying folksonomies, i.e. user-created tag-based classification schemes, to better understand how users think about content (for a review, see Kakali and Papatheodorou, 2010). Together, those analyses provide important insights about how nonexperts conceptualize cyberattack consequences.

1.4 Contributions

This research makes two important contributions:

(1) It illuminates how nonexperts think about cyberattack consequences. Without such information, it would be impossible to adequately align warning messages with how users conceptualize those consequences.

(2) It reveals what aspects of nonexperts' thinking are more or less concrete and identifies specific terminology that can be used to describe aspects that fall into each case. Such information allows one to align warning messages to nonexperts' thinking in more nuanced ways than would otherwise be possible.

2. Method

2.1 Participants

In total, 33 undergraduate students (11 males) participated in the study for partial credit in their general psychology courses. Participants' ages ranged from 18 to 54 years (M = 19.82, SD = 6.24). Each participant had neither worked in a field related to computer security or privacy nor taken a college-level course in computer security. The study was approved by the Texas Tech University Institutional Review Board.

2.2 Creating the list of cyberattack consequences

To create the list of cyberattack consequences, the authors identified possible cyberattacks for each category of threats within the spoofing, tampering, repudiation, information disclosure, denial of service, and elevation of privilege (STRIDE) threat model (Kohnfelder and Praerit, 1999). Each letter in the STRIDE model represents a different category. The authors searched the internet for attacks that fell under each letter of the STRIDE model and collected two to five descriptions of each attack. This list was as comprehensive as possible; although we acknowledge it may not contain all possible attacks.

For each attack, the authors used the internet to research how that attack would impact users and wrote a description of that consequence. Consistent with warning message design recommendations, consequence descriptions did not contain any technical jargon and presented personal consequences to the user (Bartsch and Volkamer, 2013; Bartsch *et al.*, 2013; Bravo-Lillo *et al.*, 2011; Kauer *et al.*, 2012; Modic and Anderson, 2014). For example, one consequence of a log injection attack was "The cyber attacker modified your computer files in order to hide her activities." Table 1 presents the number of attacks considered and the number of consequences identified, as a function of STRIDE category.

2.3 Data collection procedures

Participants first completed informed consent. They then sat in front of a desktop computer and were instructed how to perform the card sorting task. Specifically, participants saw the card sorting application, OptimalSort (Optimal Workshop Ltd, 2020), on the computer

Table 1. Number of attacks considered and consequences created for each STRIDE category

STRIDE category	No. of attacks	Total consequences	Unique consequences
Spoofing	12	9	7
Tampering	15	16	8
Repudiation	9	7	5
Information disclosure	19	10	6
Denial of service	25	14	11
Elevation of privilege	22	11	2

screen (Figure 1). OptimalSort was used because it simplified data collection, eliminated the need for data entry, is stable and is cost-effective. The left side of the screen listed the 50 cyberattack consequences. The remainder of the screen was a blank space where they could create groups of consequences. Participants were instructed to create groups of similar consequences by dragging a consequence from the list to the blank space and then dragging another consequence from the list and placing it on top of the other consequence. Participants were informed that each consequence could only be in one group, to place as many consequences into groups as possible and to minimize the number of single-consequence groups.

When the task began, participants read each consequence out loud, which allowed the researcher to verify that participants were aware of all consequences. Participants then sorted consequences into groups, regrouping consequences as many times as they thought necessary. Once participants were satisfied with their groups, they labeled each group based on the perceived similarity between the consequences in that group.

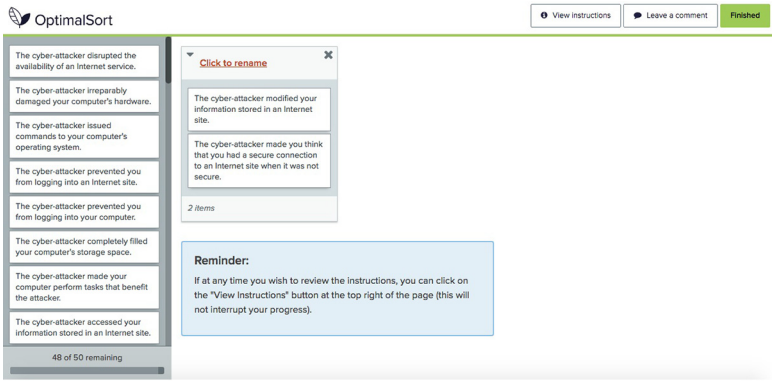
Participants then completed a postexperiment questionnaire that asked about their age, gender and computer security experience. The study lasted less than 1 h.

2.4 Creating and implementing a theme coding protocol

It was necessary to code themes present in participants' labels. Accordingly, a coding scheme was devised. To do so, the authors read through the participants' labels, which revealed a total of nine themes. As many codable themes as possible were included so the coding scheme captured the essence of the participants' labels. The nine themes and their definitions are presented in Table 2.

Training began with coders learning the definitions and coding rules. Specifically, coders were instructed to assign all relevant themes to each label. If a label did not contain any of the nine themes, then raters were instructed to mark it as "other"; although the coders were instructed to minimize the number of labels marked as "other." Coders then practiced on labels created for training purposes.

Two trained coders worked independently. Each was provided a list of the 213 unique labels participants created; each coder's list was randomized separately. Each coder assigned all relevant themes to each label.



Notes: Consequences were listed on the left side of the screen. The remainder of the screen was a workspace in which groups of consequences could be created and named

Figure 1.
OptimalSort web application

Table 2. Coding scheme of themes in group labels and frequency of their use

Theme	Definition	Frequency of use
Object of attack	The object involved in or affected by the attack	122
Action	What the person attacking is doing and how they are doing it	114
Person attacking	Who is completing the attack	59
Outcome for person attacking	The results of the attack for the attacker	29
Outcome for object of attack	The results of the attack on the object of the attack	27
Person being attacked	Who is being affected by the attack	22
Outcome for person being attacked	The results of the attack on the person being attacked	17
Type of attack	Refers to a specific type of attack or a category of attacks	17
Risk of outcome	Amount of potential for a negative outcome from the attack	5
Other	The label is not informative and does not describe any code or the label suggests these consequences do not fit into any group	7

Percent agreement (Stemler, 2004) was calculated for each theme by summing the number of labels both coders coded as being either present or absent and dividing the sum by the total number of labels coded. The interrater reliability for all themes was above 80% (Lombard *et al.*, 2002). The coders then discussed and resolved all disagreements. This final set of codes was used during subsequent analysis of the themes.

2.5 Identifying specific label features

It was also necessary to extract specific features from participants' labels. To do so, the same two coders independently extracted terms or phrases that corresponded with each theme from each label. For example, for the label "messes with the Internet," the coders extracted "messes" as an "action" and "Internet" as the "object of attack."

Percent agreement (Stemler, 2004) between the two coders' data was above 80% for each list of terms (Lombard *et al.*, 2002). The coders discussed and resolved any disagreements.

The lists of label features were then simplified. Specifically, terms were stemmed, e.g. the terms "accessed" and "accesses" were stemmed to "access." Further, synonyms were removed, e.g. when the terms "change," "modify" and "alter" were all present, one term, "alter," was retained and its synonyms were removed. WordNet web (Princeton University, 2010) was used to determine whether terms were synonyms. The resultant list of label features was used during subsequent analysis of specific features.

3. Results

Participants' labels were analyzed to understand the general themes and the specific features that were present in nonexperts' thinking about cyberattack consequences. The former will be discussed in Section 3.1 and the latter in Section 3.2.

3.1 What general themes were present?

To determine what general themes were present in nonexperts' thinking, we summed the number of labels that reflected each theme for each participant and then summed across participants (Table 2). The resultant frequency indicated how many times each theme was present.

We conducted two generalized estimating equations (GEE) (Liang and Zeger, 1986) to determine whether themes occurred more or less frequently than one another. A Poisson distribution with log link function and an exchangeable correlation was used because of the within-subject design. The first GEE indicated the theme "action" occurred more frequently than all themes, logit ≤ -3.13, $Z \geq -3.30$, $p < 0.001$, except for the theme "object of attack," logit = 0.07, $Z = 0.50$, $p = 0.614$. The second GEE indicated the theme "object of attack" occurred more frequently than all themes, logit ≤ -3.19, $Z \geq -2.97$, $p < 0.001$, except the theme "action," logit = -0.07, $Z = -0.50$, $p = 0.614$. Based on these results, we concluded the themes "action" and "object of attack" occurred most frequently, and they occurred much more frequently than the other themes.

The labels reflected how participants thought the grouped consequences were similar. As such, they provide a window into how participants thought about those consequences. Based on the present analyses, it appears participants mainly thought about cyberattack consequences in terms of what the attacker is doing and how they are doing it ("action") and what the attack affects ("object of attack").

3.2 What specific features were present?

To determine what specific features were present in participants' labels, we created mind maps for each of the nine themes (see Figures 2–8). To create each map, the relevant theme (e.g. "object of attack") was set as the central node. Specific terms participants used in labels that reflected that theme were then distributed around that central node. In certain cases, terms were organized hierarchically. For example, the terms "hardware," "function" and "files" are nested under the term "computer" because participants referred to "computer hardware," "computer function" and "computer files" (Figure 3). Arranging those terms hierarchically drew attention to relations between terms that otherwise might have been difficult to detect. We discuss individual mind maps in Sections 3.2.1 and 3.2.2.

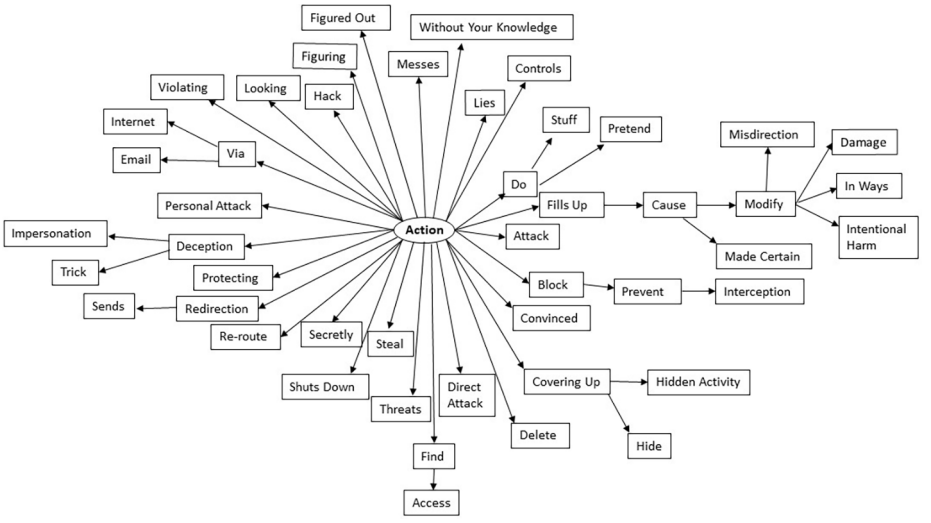


Figure 2.
Mind map for the "action" theme, which concerns what the attacker is doing or how they are doing it

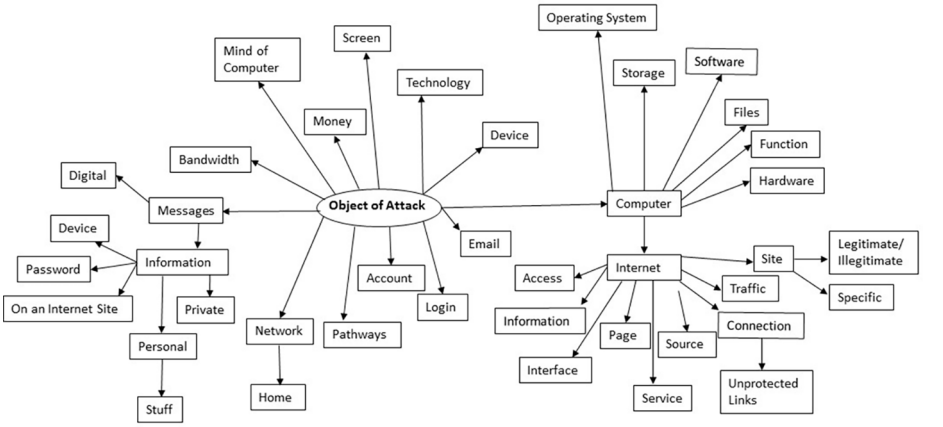


Figure 3.
Mind map for the "object of attack" theme, which concerns the object involved in or affected by the attack

Please note that we did not discuss mind maps for two themes:

(1) "person attacking"; and

(2) "person being attacked."

Participants used few terms related to those themes, and the terms they used did not reveal much about how participants thought about cyberattackers or attack victims. For example, for "person attacking," participants used either a variant of "attacker" or pronouns such as "they" or "him." Accordingly, we thought it best to exclude those mind maps.

Inspection of the remaining seven mind maps revealed that those for the "action," "object of attack" and "outcome for object of attack" themes were more detailed and contained more concrete terms than the others. Accordingly, we decided to discuss the mind maps for those three themes in Section 3.2.1 and to discuss those for the other themes in Section 3.2.2.

3.2.1 Mind maps for the "action," "object of attack" and "outcome for object of attack" themes. Figures 2–4 present the mind maps for the "action," "object of attack" and "outcome for object of attack" themes, respectively. Those themes concern what the attacker is doing or how they are doing it ("action"), the object involved in or affected by the attack ("object of attack") and the results of the attack on the object of the attack ("outcome for object of attack"), respectively.

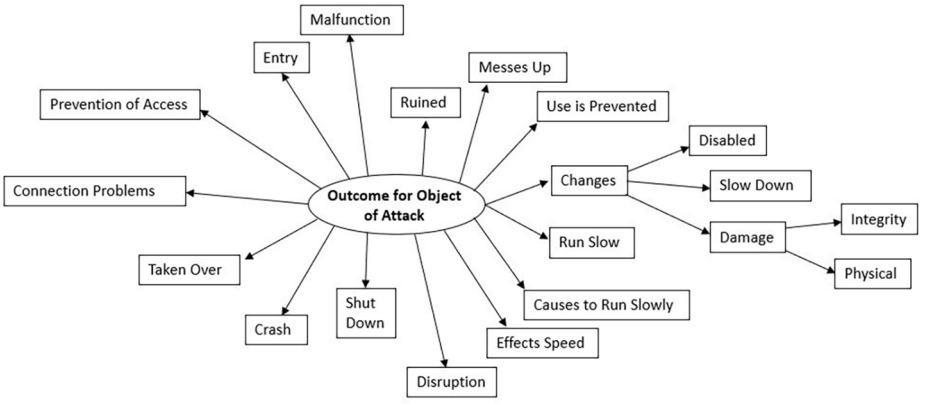


Figure 4.
Mind map for the "outcome for object of attack" theme, which concerns the results of the attack on the object of the attack

Inspection of Figures 2–4 suggest participants used concrete terms when evoking these themes. For example, participants used terms such as "shuts down" and "fills up" when evoking the "action" theme, "password information" and "Internet connection" when evoking the "object of attack" theme and "use is prevented" and "cause to run slowly" when evoking the "outcome for object of attack" theme. This suggests nonexperts think about these aspects of cyberattacks in concrete terms.

3.2.2 Mind maps for the "outcome for person attacking," "outcome for person being attacked," "risk of outcome" and "type of attack" themes. Figures 5–8 present the mind maps for the "outcome for person attacking," "outcome for person being attacked," "risk of outcome" and "type of attack" themes, respectively. Those themes concern the results of the

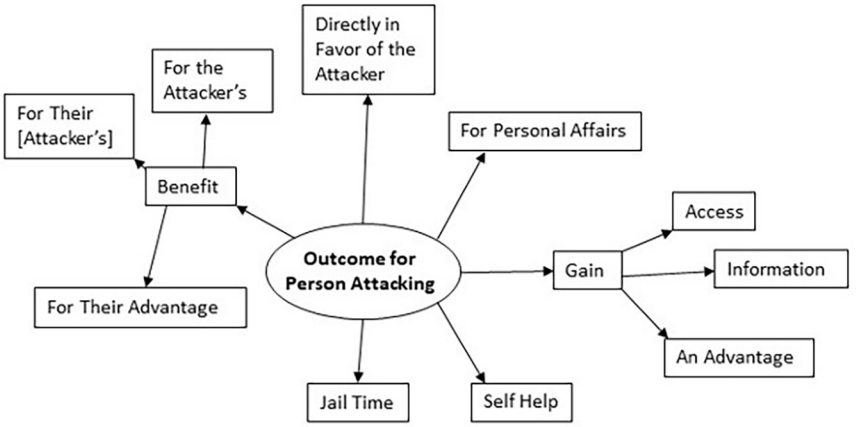


Figure 5. Mind map for the "outcome for person attacking" theme, which concerns the results of the attack for the attacker

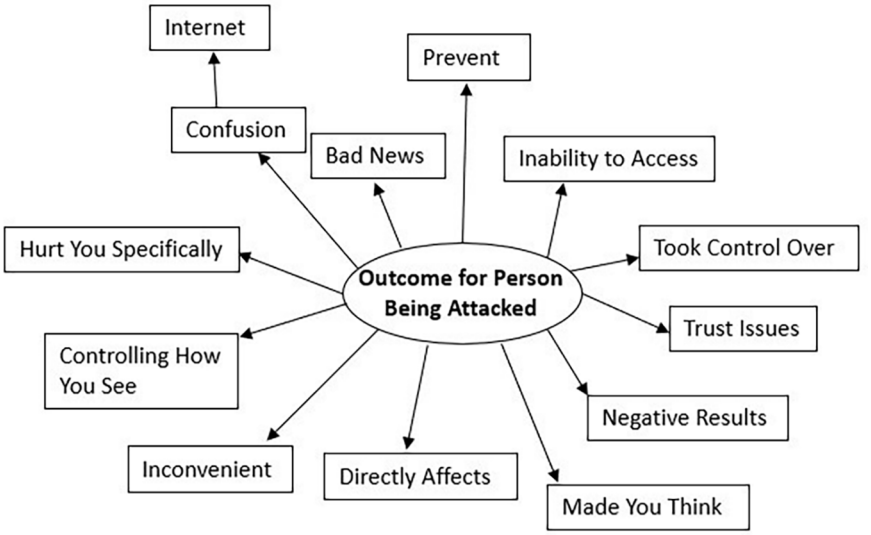


Figure 6. Mind map for the "outcome for person being attacked" theme, which concerns the results of the attack on the person being attacked

attack for the attacker ("outcomes for person attacking"), the results of the attack on the person being attacked ("outcome for person being attacked"), amount of potential for a negative outcome from the attack ("risk of outcome") and specific types or categories of attacks ("type of attack"), respectively.

Inspection of Figures 5–8 suggest participants typically used very general terms when evoking these themes. For example, participants used terms such as "for personal affairs" and "for the attacker's benefit" when evoking the "outcome for the person attacking" theme, "inconvenient" and "negative results" when evoking the "outcome for the person being attacked theme," "potential damage" and "more serious" when evoking the "risk of outcome" theme and "general computer attack" and "identity theft" when evoking the "type of attack" theme.

4. Discussion

In the following subsections, we discuss the present results' implications. Specifically, in Section 4.1, we discuss our results' general implications for our understanding of how nonexperts think about cyberattack consequences. In Section 4.2, we discuss specific

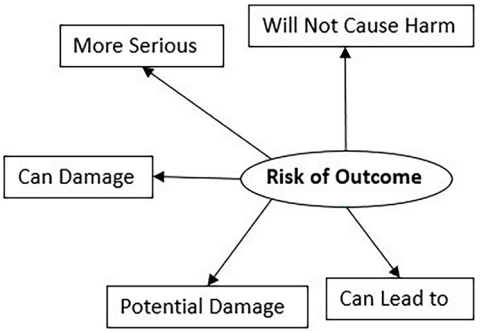


Figure 7.
Mind map for the "risk of outcome" theme, which concerns the amount of potential for a negative outcome from the attack

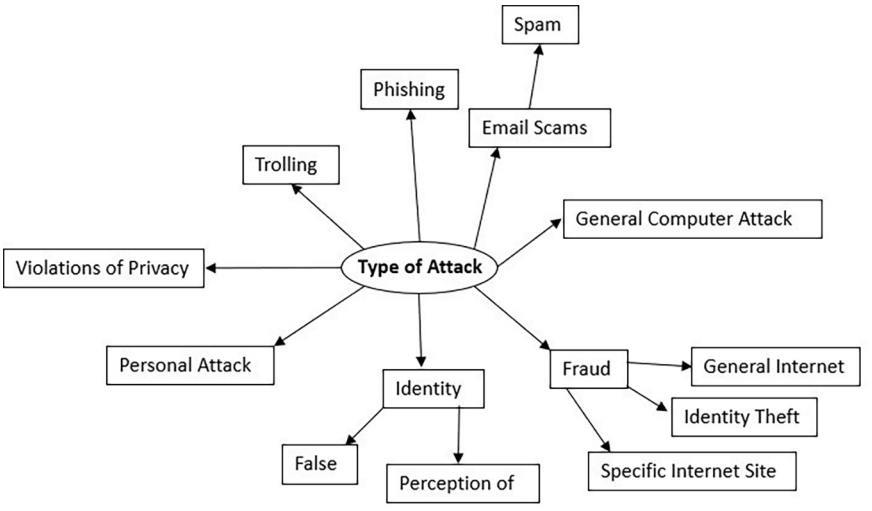


Figure 8.
Mind map for the "type of attack" theme, which concerns specific types or categories of attacks

recommendations for warning message design that stem from our results. In Section 4.3, we discuss instances in which the present results replicated past findings.

4.1 General implications of how nonexperts think about cyberattack consequences

Inspection of Figures 5–8 suggests nonexperts think about these aspects of cyberattacks in nondescript ways, which could be problematic. The following describes two examples.

A nonexpert who thinks in general terms about how an attack will benefit attackers ("outcome for person attacking") will have difficulty understanding attackers' motives. Truth-default theory (TDT) suggests people typically assume others are being honest and do not evaluate the veracity of what they are being told unless something triggers them to do so (Clare and Levine, 2019; Levine, 2014). Further, TDT suggests that becoming aware of a motive for lying can be a powerful trigger (Levine, 2014). Accordingly, a nonexpert who thinks in nondescript ways about how an attack will benefit attackers ("outcome for person attacking") and thus does not understand the attacker's motives will likely believe the attacker is being honest.

A nonexpert who thinks in general terms about the risks associated with a given attack will have difficulty understanding what that attack will cost them. Users weigh the pros and cons of doing what it takes to protect their cybersecurity (Hardee *et al.*, 2006; Herley, 2009). Further, as discussed in Section 1.1.3, nonexperts often think that doing what it takes to protect their cybersecurity will distract them from the task at hand (Dourish *et al.*, 2003; Hardee *et al.*, 2006; Kang *et al.*, 2015; Sasse *et al.*, 2001) and negatively affect their productivity (Vaniea *et al.*, 2014). Accordingly, a nonexpert who thinks in nondescript ways about the risks associated with a given attack will likely think the pros associated with protecting their cybersecurity are outweighed by the cons associated with doing so, which will likely result in them choosing to behave in an unsafe manner.

Further, inspection of Figure 7 suggests participants thought some cyberattack consequences were more severe than others. Specifically, participants used the phrase "more serious" when labeling certain groups of consequences, and the phrase "will not cause harm" when labeling other groups of consequences.

4.2 Specific recommendations for warning message design

As noted in Section 1.2, it is recommended that warning messages use nontechnical language (Bauer *et al.*, 2013), describe how the cyberattack will affect the user personally (Bartsch *et al.*, 2013; Kauer *et al.*, 2012) and do so in a way that aligns with how the user thinks about cyberattacks (Bartsch *et al.*, 2013). The present results have important implications for how to accomplish that goal.

First, the present results revealed that nonexperts mainly thought about cyberattack consequences in terms of what attackers are doing ("action") and what will be affected ("object of attack"). Therefore, warning message wording should be anchored on those two topics. For example, a warning message could state "[. . .] this website is not safe because the information entered onto the site ("object of attack") could be viewed ("action") by other people, including cyber-attackers who could gain the user's personal information ("object of attack") [. . .]."

Second, the present results indicated nonexperts thought in concrete terms about what attackers are doing ("action"), what will be affected ("object of attack") and how it will be affected ("outcome for object of attack"). Accordingly, one should write about such topics in concrete terms, preferably using specific terms found in Figures 2–4. For example, a warning message could state "[. . .] Someone is attempting to fill up ("action") your computer ("object of attack") so that it malfunctions ("outcome for object of attack") [. . .]."

Third, the present results indicated nonexperts thought in general terms about the results of the attack for the attacker ("outcomes for person attacking"), the results of the attack on the person being attacked ("outcome for person being attacked"), amount of potential for a negative outcome from the attack ("risk of outcome") and specific types or categories of attacks ("type of attack"). Thus, when one writes about those topics, they should do so in general terms, preferably using the terms found in Figures 5–8. For example, a warning message could frame outcomes for the cyberattacker in terms of "[. . .] the attack will benefit the attacker [. . .]" or "[. . .] the attacker will gain information [. . .]."

Admittedly, this is counter intuitive. Typically, warning messages aim to describe the nature of the attack as concretely as possible and risks in ways that impress upon users the seriousness of the situation. The present results suggest such wording likely will not align with how nonexperts think about these aspects of cyberattack consequences. Accordingly, nonexperts may distrust the warning message (Bartsch *et al.*, 2013; Bartsch and Volkamer, 2013; Ibrahim *et al.*, 2010).

Fourth, the present results revealed that nonexperts thought in technical terms in one instance. Specifically, they thought in terms of "phishing" when evoking the "type of attack" theme. In general, warning messages should not include technical language (Bauer *et al.*, 2013). However, the present results suggest that it may be acceptable to use the technical term "phishing" in a warning message.

Implementing the recommendations above should increase the likelihood that the wording of a given warning message aligns with how nonexperts think about the attack's consequence. That should increase their trust in the warning message, which should increase the likelihood that they will comply with it (Bartsch *et al.*, 2013; Bartsch and Volkamer, 2013; Ibrahim *et al.*, 2010).

4.3 Replications of prior results

Inspection of Figure 7 suggests participants thought about risk in general terms, e.g. "more serious" and "will not cause harm." This outcome replicates Bartsch and Volkamer's (2013) results. Specifically, those authors reported that nonexperts discussed risk in nondescript ways, e.g. "comparatively bad" (p. 8). The present results echo that finding.

Inspection of Figure 8 suggests that participants only used the formal name of one specific type of cyberattack, phishing, when evoking the "type of attack" theme. Accordingly, the present study replicated that nonexperts do not know the technical terms cybersecurity professionals use when referring to different types of cyberattacks (Kauer *et al.*, 2012; Modic and Anderson, 2014), which reinforces the general recommendation to avoid technical language when crafting warning messages (Bauer *et al.*, 2013).

5. Future directions

Sections 5.1–5.3 describe several promising avenues for future research.

5.1 Confirming the present results with a more diverse sample

How people think about cyber-related topics differs based on age and education level (Wash and Rader, 2015). For example, adults under the age of 50 years, but not older adults, thought viruses could be caught merely by browsing the internet. Further, people who completed high school, but not people who had postgraduate education, thought cyberattacks target individual computer users.

Participants in the present study were a relatively homogenous sample of mostly young adult, college students. Therefore, it was not possible to examine whether the current

study's results reflect all users or only young adult, college students. Accordingly, future research should replicate the present study with a more diverse sample.

5.2 Developing a more nuanced understanding of perceived severity

As noted in Section 4.1, participants thought some cyberattack consequences were more severe than others. This is noteworthy because nonexperts consider consequence severity when determining whether to act against an attack (Bartsch *et al.*, 2013; Dodel and Mesch, 2017; Ng *et al.*, 2009). For example, when a person considers a consequence to be severe, they are more likely to attempt a behavior they are told will prevent that consequence, even if they do not think that behavior will prevent the consequence, than when they perceive a consequence to not be severe (Ng *et al.*, 2009). Further, reported knowledge of attack severity correlates positively with the reported amount of behaviors to prevent malware (Dodel and Mesch, 2017).

However, the present results do not specify exactly how severe participants perceived specific consequences to be. Foster *et al.* (2021) addressed that issue for phishing attack consequences. Future research should examine the perceived severity of the consequences of other attack types, which will provide an important piece to our understanding of how users think about cyberattack consequences.

5.3 Understanding how users' values affect their cybersecurity

Your values influence your behavior (Maio, 2016). For example, some nonexperts do not value their privacy, so they choose to not protect it (Kang *et al.*, 2015; Renaud *et al.*, 2014). For example, users reported that they do not care whether others read their personal email, so they did not take steps to prevent others from doing so (Renaud *et al.*, 2014).

Such results present a challenge for warning message design. Arguably, the design of most warning messages assumes users value their privacy and will do what is necessary to protect it if they understand their privacy is threatened and how to protect it. For example, a warning message might tell users that attackers could access the user's personal information and suggest behaviors to safeguard their information. However, that warning message may not convince a user who does not value their privacy to do what is necessary to protect it (Jones *et al.*, 2021). Future research should investigate how to encourage users to protect their cybersecurity when users' unsafe behavior reflects their values more so than their understanding of the situation.

6. Summary

The current study determined how nonexperts think about cyberattack consequences. Participants grouped consequences based on perceived similarity and gave each group a label that represented why they perceived the grouped consequences to be similar. The results suggested participants mainly thought about consequences in terms of what the attacker is doing ("action") and what will be affected ("object of attack"). Further, the results suggest participants thought about certain aspects of the consequences, "action." "object of attack" and "outcome for object of attack," in concrete terms and the other aspects of the consequences in general terms. Therefore, the present results suggest warning message wording should be anchored on what the attacker is doing ("action") and what will be affected ("object of attack") and reflect whether nonexperts think about a given aspect of the consequence in concrete or general terms. Doing so should increase the likelihood that the wording of a given warning message aligns with how nonexperts think about the attack's consequence. That should increase their trust in the warning message, which should increase the likelihood that they will comply with it (Bartsch *et al.*, 2013; Bartsch and Volkamer, 2013; Ibrahim *et al.*, 2010).

References

Acar, Y., Fahl, S. and Mazurek, M.L. (2016), "You are not your developer, either: a research agenda for usable security and privacy research beyond end users", *2016 IEEE Cybersecurity Development (SecDev)*, IEEE, pp. 3-8.

Agrawal, N., Zhu, F. and Carpenter, S. (2020), "Do you see the warning? Cybersecurity warnings via nonconscious processing", *2020 ACM Southeast Conference (ACMSE 2020)*, Association for Computing Machinery, *Tampa, FL*, pp. 260-263.

Akhawe, D. and Felt, A.P. (2013), "Alice in Warningland: a large-scale field study of browser security warning effectiveness", *22nd USENIX Security Symposium*, *USENIX Association*, *Washington,* DC, pp. 257-272.

Amran, A., Zaaba, Z.F., Singh, M.M. and Marashdih, A.W. (2017), "Usable security: revealing end-users comprehensions on security warnings", *Procedia Computer Science*, Vol. 124, pp. 624-631.

Asgharpour, F., Liu, D. and Camp, L.J. (2007), "Mental models of computer security risks", in Dietrich, S. and Dhamija, R. (Eds), *Financial Cryptography and Data Security*, Springer, Berlin, Heidelberg, FC 2007 and USEC 2007, Trinidad and Tobago, pp. 367-377.

Barnett, J. (2004), "The multiple sorting procedure (MSP)", in Breakwell, G.M. (Ed.), *Doing Social Psychology Research*, Wiley Online Library, Oxford, pp. 289-304.

Bartsch, S. and Volkamer, M. (2013), "Effectively communicate risks for diverse users: a mental-models approach for individualized security interventions", in Horbach, M. (Ed.), *INFORMATIK 2013*, GER, *Bonn*, pp. 1971-1984.

Bartsch, S., Volkamer, M., Theuerling, H. and Karayumak, F. (2013), "Contextualized web warnings, and how they cause distrust", *International Conference on Trust and Trustworthy Computing*, Springer, Berlin, Heidelberg, pp. 205-222.

Bauer, L., Bravo-Lillo, C., Cranor, L.F. and Fragkaki, E. (2013), "Warning design guidelines", *CMU-CyLab*, Vol. 13, pp. 1-27.

Bravo-Lillo, C., Cranor, L.F., Downs, J.S. and Komanduri, S. (2011), "Bridging the gap in computer security warnings: a mental model approach", *IEEE Security and Privacy Magazine*, Vol. 9 No. 2, pp. 18-26, doi: 10.1109/MSP.2010.198.

Clare, D.D. and Levine, T.R. (2019), "Documenting the truth- default: the low frequency of spontaneous unprompted veracity assessments in deception detection", *Human Communication Research*, Vol. 45 No. 3, pp. 286-308, doi: 10.1093/hcr/hqz001.

Dodel, M. and Mesch, G. (2017), "Cyber-victimization preventive behavior: a health belief model approach", *Computers in Human Behavior*, Vol. 68, pp. 359-367.

Dourish, P., De La Flor, J.D. and Joseph, M. (2003), "Security as a practical problem: some preliminary observations of everyday mental models", *Proceedings of CHI 2003 workshop on HCI and security systems*, ACM.

Dupuis, M., Jennings, A. and Renaud, K. (2021), "Scaring people is not enough: an examination of fear appeals within the context of promoting good password hygiene", *Proceedings of the 22st Annual Conference on Information Technology Education*, pp. 35-40.

Egelman, S., Cranor, L.F. and Hong, J. (2008), "You've been warned: an empirical study of the effectiveness of web browser phishing warnings", *CHI 2008: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, *Florence*, pp. 1065-1074.

Fincher, S. and Tenenberg, J. (2005), "Making sense of card sorting data", *Expert Systems*, Vol. 22 No. 3, pp. 89-93.

Foster, E.K., Jones, K.S., Armstrong, M.E. and Namin, A.S. (2021), "User perceptions of phishing consequence severity and likelihood, and implications for warning message design", *International Conference on Applied Human Factors and Ergonomics*, Springer, *Cham*, pp. 265-273.

Hardee, J.B., West, R. and Mayhorn, C.B. (2006), "To download or not to download: an examination of computer security decision making", *Interactions*, Vol. 13 No. 3, pp. 32-37.

Herley, C. (2009), "So long, and no thanks for the externalities: the rational rejection of security advice by users", *Proceedings of the 2009 Workshop on New Security Paradigms Workshop (NSPW '09)*, ACM, pp. 133-144.

Ibrahim, T., Furnell, S.M., Papadaki, M. and Clarke, N.L. (2010), "Assessing the usability of end-user security software", *International Conference on Trust, Privacy and Security in Digital Business*, Springer, pp. 177-189.

Ion, I., Reeder, R., Consolvo, S. (2015), "'. . . no one can hack my mind': comparing expert and non-expert security practices", *Symposium on Usable Privacy and Security (SOUPS) 2015*, USENIX Association, Ottawa, pp. 327-346.

Jeong, R. and Chiasson, S. (2020), "'Lime', 'open lock', and 'blocked' children's perception of colors, symbols, and words in cybersecurity warnings", *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ACM, *Honolulu, HI*, pp. 1-13.

Jones, K.S., Lodinger, N.R., Widlus, B.P., Namin, A.S. and Hewett, R. (2021), "Do warning message design recommendations address why non-experts do not protect themselves from cybersecurity threats? A review", *International Journal of Human–Computer Interaction*, Vol. 37 No. 18, pp. 1709-1719.

Kakali, C. and Papatheodorou, C. (2010), "Exploitation of folksonomies in subject analysis", *Library and Information Science Research*, Vol. 32 No. 3, pp. 192-202.

Kang, R., Dabbish, L., Fruchter, N. and Kiesler, S. (2015), "'My data just goes everywhere': user mental models of the internet and implications for privacy and security", *Symposium on Usable Privacy and Security (SOUPS)*, *USENIX Association*, pp. 39-52.

Kauer, M., Günther, S., Storck, D. and Volkamer, M. (2013), "A comparison of American and German folk models of home computer security", *International Conference on Human Aspects of Information Security, Privacy, and Trust*, Springer, *Berlin, Heidelberg*, pp. 100-109.

Kauer, M., Pfeiffer, T., Volkamer, M., Theuerling, H. and Bruder, R. (2012), "It is not about the design – it is about the content! Making warnings more efficient by communicating risks appropriately", in Suri, N. and Waidner, M. (Eds), *SICHERHEIT 2012 – Sicherheit, Schutz Und Zuverlassigkeit*, GER, Bonn, pp. 187-198.

Kohnfelder, L. and Praerit, G. (1999), "The threats to our products", Microsoft Corporation.

Levine, T.R. (2014), "Truth-default theory (TDT): a theory of human deception and deception detection", *Journal of Language and Social Psychology*, Vol. 33 No. 4, pp. 378-392.

Liang, K.Y. and Zeger, S.L. (1986), "Longitudinal data analysis using generalized linear models", *Biometrika*, Vol. 73 No. 1, pp. 13-22.

Lombard, M., Snyder-Duch, J. and Bracken, C.C. (2002), "Content analysis in mass communication: assessment and reporting of intercoder reliability", *Human Communication Research*, Vol. 28 No. 4, pp. 587-604.

Maio, G.R. (2016), *The Psychology of Human Values*, Psychology Press, New York, NY.

Modic, D. and Anderson, R. (2014), "Reading this may harm your computer: the psychology of malware warnings", *Computers in Human Behavior*, Vol. 41, pp. 71-79.

Ng, B.Y., Kankanhalli, A. and Xu, Y.C. (2009), "Studying users' computer security behavior: a health belief perspective", *Decision Support Systems*, Vol. 46 No. 4, pp. 815-825.

Optimal Workshop Ltd. (2020), "OptimalSort", available at: www.optimalworkshop.com/optimalsort/

Prettyman, S.S., Furman, S., Theofanos, M. and Stanton, B. (2015), "Privacy and security in the brave new world: the use of multiple mental models", in Tryfonas, T. and Askoxylakis, I. (Eds), *International Conference on Human Aspects of Information Security, Privacy, and Trust*, Springer, *Cham*, pp. 260-270.

Princeton University (2010), "About WordNet", available at: http://wordnetweb.princeton.edu/perl/webwn

Renaud, K., Volkamer, M. and Renkema-Padmos, A. (2014), "Why doesn't Jane protect her privacy?", in De Cristofaro, E. and Murdoch, S.J. (Eds), *Privacy Enhancing Technologies (PETS). Lecture Notes in Computer Science*, Springer, p. 8555.

Renaud, K., Searle, R. and Dupuis, M. (2021), "Shame in cyber security: effective behavior modification tool or counterproductive foil?", *New Security Paradigms Workshop 2021*.

Sasse, A. (2015), "Scaring and bullying people into security won't work", *IEEE Security and Privacy*, Vol. 13 No. 3, pp. 80-83.

Sasse, M.A., Bronstoff, S. and Weirich, D. (2001), "Transforming the 'weakest link': a human-computer interaction approach for usable and effective security", *BT Technology Journal*, Vol. 19 No. 3, pp. 122-131, doi: 10.1023/A:1011902718709.

Sombatruang, N., Omiya, T., Miyamoto, D., Sasse, M.A., Kadobayashi, Y. and Baddeley, M. (2020), "Attributes affecting user decision to adopt a virtual private network (VPN) app", *International Conference on Information and Communications Security*, Springer, *Cham*, pp. 223-242.

Stemler, S.E. (2004), "A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability", *Practical Assessment, Research and Evaluation*, Vol. 9 No. 4, pp. 1-19.

Theofanos, M., Stanton, B., Furman, S., Prettyman, S.S. and Garfinkel, S. (2017), "Be prepared: how US government experts think about cybersecurity", *Network and Distributed System Security Symposium (NDSS)*, Information Society, *San Diego, CA*, pp. 1-11, doi: 10.14722/usec.2017.23006.

Ur, B., Bees, J., Segreti, S.M., Bauer, L., Christin, N. and Cranor, L.F. (2016), "Do users' perceptions of password security match reality?", *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ACM, pp. 3748-3760.

Vaniea, K.A., Rader, E. and Wash, R. (2014), "Betrayed by updates: how negative experiences affect future security", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, pp. 2671-2674, doi: 10.1145/2556288.2557275.

Volkamer, M. and Renaud, K. (2013), "Mental models – general introduction and review of their application to human-centred security", *Number Theory and Cryptography*, Springer, Berlin, Heidelberg, pp. 255-280.

Wash, R. (2010), "Folk models of home computer security", *Proceedings of the Sixth Symposium on Usable Privacy and Security (SOUPS) 2010*, ACM, *Redmond, WA*, pp. 1-16.

Wash, R. and Rader, E.J. (2015), "Too much knowledge? Security beliefs and protective behaviors among United States internet users", *Symposium on Usable Privacy and Security (SOUPS) 2015*, *USENIX Association*, *Ottawa*, pp. 309-325.

Weirich, D. and Sasse, M.A. (2001), "Pretty good persuasion: a first step towards effective password security in the real world", *Proceedings of the 2001 workshop on new security paradigms (NSPW '01)*, ACM, pp. 137-143.

West, R., Mayhorn, C., Hardee, J. and Mendel, J. (2009), "The weakest link: a psychological perspective on why users make poor security decisions", *Social and Human Elements of Information Security: Emerging Trends and Countermeasures*, IGI Global, Hershey, PA, pp. 43-60.

Wu, J. and Zappalla, D. (2018), "When is a tree really a truck? Exploring mental models of encryption", *Proceedings of the fourteenth symposium on usable privacy and security (SOUPS)*, *USENIX Association*, pp. 395-409.

Zaaba, Z.F., Furnell, S.M. and Dowland, P.S. (2016), "Literature studies on security warnings development", *International Journal on Perceptive and Cognitive Computing*, Vol. 2 No. 1, pp. 8-18.

Further reading

Krol, K., Moroz, M. and Sasse, M.A. (2012), "Don't work. Can't work? Why it's time to rethink security warnings", *7th International Conference of Risks and Security of Internet Systems (CRISIS)*, IEEE, *Cork*, pp. 1-8.

Norman, D.A. (2014), "Some observations on mental models", *Mental Models*, Psychology Press, New York, NY, pp. 15-22.

Rumelhart, D.E. and Norman, D.A. (1983), "Representation in memory", Working paper [ONR 8302], University of California, San Diego, La Jolla, CA, June.

Corresponding author

Keith S. Jones can be contacted at: keith.s.jones@ttu.edu

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**