Warning! A Comprehensive Model of the Effects of Digital Information Security Warning Messages

Research-in-progress

Mario Silic
University of St Gallen/ZSEM
mario.silic@unisg.ch

Jordan Barlow
California State University
jobarlow@ fullerton.edu

Dustin Ormond
Creighton University
DustinOrmond@creighton.edu

ABSTRACT

Despite existing countermeasures to combat malicious actions, users are the last line of defense to protect personal and organizational assets. Given that users often ignore warning messages that motivate compliant behavior, the issue of protecting personal and organizational assets is exacerbated. Messages that are largely ignored cannot have any impact on attitudes, motivation, or behavior. Therefore, crafting messages that increase attention and comprehension regarding specific threats and ways to cope with these threats is vital. This research combines the communication-human information processing (C-HIP) model with protection motivation theory (PMT) to assess how warning message content affects adherence especially when users pay attention to the content of the warning message. In essence, this study considers a holistic view of examining the channel (warning message), attention, comprehension and their influence on attitudes and beliefs, motivation, and behavior. Additionally, we propose including alternative courses of action in digital warning messages to increase secure attitudes, beliefs, and behavior. We test this holistic model through a series of field and lab experiments to evaluate message comprehension, attitudes, and beliefs and capture actual attention and secure behavior.

INTRODUCTION

Cybercrime is increasing, targeting individuals, organizations, and governments at a rapid rate. The estimated cost of cybercrime for the global economy is around $445 billion each year, where 800 million people in 2013 were affected by cyber espionage and loss of private information (McAffee 2014). Despite many existing countermeasures aiming at protecting users' integrity (e.g., antivirus software, firewalls, operating system mechanisms such as password protection when installing new software, etc.), in practice users represent the last line of defense against malicious actions. Such actions can be either directed against themselves (e.g., malware destroying user's hard drive) or against organizational assets where the user is used as the backdoor by cybercriminals.

Information security research has examined several different methods and techniques for persuading users to behave securely in organizations and other settings (e.g., deterrence techniques, anti-neutralization techniques, SETA training programs, etc.). However, these techniques have not been evaluated when examining pop-up warning messages. Warnings represent communication designed to prevent users from hurting themselves or others (Wogalter 2006b) and as such, physical warnings have been shown to be very effective in preventing hazards or criminal incidents (Coleman 2007; Goldstein, Cialdini, & Griskevicius 2008; Schultz & Tabanico 2009). Less is known about digital or computer warning messages.

Digital warnings are unique from other security measures in that they are usually not the first line of defense for users. According to the "hazard control" hierarchy (Wogalter 2006b), the first step to control or remove risk is an attempt to eliminate or minimize the hazard as much as possible. The second step strives to minimize the interaction between the user and the hazard. Finally, the third step provides warning messages to the user which may reduce risk by enabling better decision-making. In other words, warning messages are unique in that they are provided only when other, potentially more powerful, security measures are not able to keep risk from the user. An example of non-digital warnings is that of the tobacco industry. Users are constantly informed about the health risks of smoking and its consequences. However, warnings are quite often ignored by users and may even produce the "boomerang" effect—that is, warnings have the potential to increase harmful behavior by drawing attention to such behavior (Bushman 2006).

A similar phenomenon seems to be happening in the digital world. For example, Egilman and Bohme (2006) argue that people do not read digital warnings, as they are habituated to them. Other studies found that users ignore web browser SSL warnings and simply skip them (e.g. Akhawe & Felt 2013; Sunshine et al. 2009a). Research on computer warning messages indicates that HCI elements are integral parts of these messages (Bravo-Lillo et al. 2011a); however, it does not assess the psychological effects of these messages—wording of warning messages appears to be based on trial and error rather than persuasion or communication theories (Modic & Anderson 2014). However, it is important to test theory-based communication in the unique warning messages context that is less direct than the previously-tested theory-based security trainings and other security information. The few research studies that have addressed computer security warning content have neglected that users commonly ignore warnings in the first place (e.g. Akhawe & Felt 2013; Sunshine et al. 2009a). For example, Egelman, Cranor, and Hong (2008) manipulated the content of the malware warning to understand the effects on the user's behavior but did not take into account the initial warning ignorance where users do not read warning at all.

On the other hand, recent research on warnings (Anderson et al. 2014a, 2014b) focuses on why computer warning messages are largely ignored and ways (e.g. polymorphic warnings) to have people pay more attention to them (Anderson et al. 2014a, 2014b). However, even if people actually read the warning messages, they may reject them based on their content. Therefore, research should assess how warning message content, based on theory, may affect adherence especially when users pay attention to the content of the warning message. Essentially, an empirical research study is needed to understand both the attention and content aspects of computer warning messages and their effects on users.

In this study, we measure the total time people spend (attention) reading a variety of warning messages (content). Then we examine the effect of the content for only those who actually paid attention to the warning message. As a foundation for this research, we evaluate the Communication-Human

Information Processing (C-HIP) Model to understand and test the process and interactions of attention, comprehension, attitudes, beliefs, and motivation on ultimate user behavior when they encounter computer security warning messages.

Further, based on the C-HIP model and other related theories of communication and persuasion, such as the Health Beliefs Model, this study proposes a new content element of computer security warning messages (i.e., suggesting alternative secure courses of action) that users who pay attention to warning message content may be more persuaded to behave securely. Thus, this leads to our research questions:

RQ1. What aspects of warning messages are most powerful in keeping individuals from performing potentially insecure IT behavior, particularly considering the attention and comprehension of the user toward the message?

RQ2. When warning messages include content directing users to alternative courses of action, is the likelihood to heed the warning increased?

THEORETICAL BACKGROUND

Research on warnings in the physical world has been categorized into the Communication-Human Information Processing (C-HIP) Model (Conzola & Wogalter 2001; Wogalter 2006a), shown below in Figure 1.

This framework shows that in order to communicate a message (such as a warning), you have to consider the source, the channel, and multiple aspects of the receiver. These aspects start with gaining and retaining attention and then proceed to comprehension, attitudes, beliefs, motivation, and ultimately behavior. *Source* refers to the person or entity delivering the message. In the case of digital information security warnings, the source could be anti-virus software, an organization’s IT department, or others. However, the source is often hidden from the user—they only see warning messages as appearing on the screen “out of nowhere.” Although the source is an important attribute in successfully communicating a message through a warning channel, for the sake of simplicity and brevity we do not focus on source characteristics in this study. Future research should address this issue.

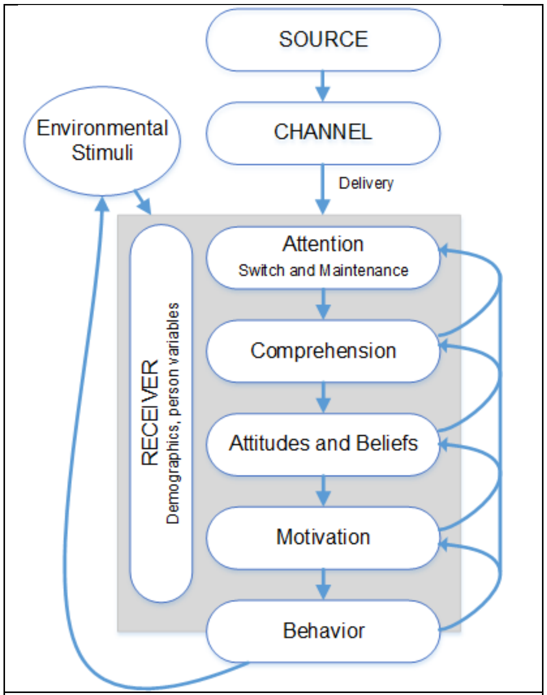


Figure 1. C-HIP model. Adapted from Conzola and Wogalter (2001) and Wogalter (2006a)

Channel is the method of delivering the communication. In the case of digital warning messages, the warning message itself is the channel. Previous research has indicated that the source of the warning message communicates the presence of a hazard through some media channel to a recipient (Chen et al. 2014). For instance, Bravo-Lillo et al. (2013) designed “attractors” (i.e. user interface modifications) to draw attention to the most important and pertinent information to aid in decision-making. Another study tracked eye movements and found that users paid no attention to SSL icons (Grier, Tang, & King 2008). However, these studies addressed the effect of digital warning channel attributes either on (1) behavior, without investigating the receiver attributes that mediate or moderate this relationship, or (2) attention, without considering that attention is only one aspect of the receiver affecting their ultimate behavior.

This study focuses on the various stages of C-HIP involving the *receiver* (i.e., the person toward whom the warning message is addressed) and the effects of channel content on the receiver. As shown in the gray box in the figure above, there are several steps involved in communicating a message, and several different factors that could affect the ultimate behavior of the receiver.

The first step is *attention*. If receivers are not paying attention to the message, it cannot have any further impact on their behavior. Attention can often be gained through simple visual aspects (e.g., size, colors, graphics) (Laughery & Wogalter 2006). Another aspect that can have high impact on the user's attention is the environment itself, which can be cluttered and noisy. Thus, to attract attention, a warning has to be conspicuous or salient relative to its context (Sanders & McCormick 1987). According to Wogalter and Laughery (1996), a user's attention will be driven by (1) spatial and temporal factors such as novelty, size, illumination, and contrast, (2) signal words such as "DANGER", (3) signal icons such as an exclamation point, (4) color such as red which signals danger in many cultures, and (5) pictures such as a pictorial sign displaying smoking consequences. One study on web browser warnings, such as those that appear when users visit suspected phishing websites, showed that altering text and color led to a significant increase of user's attention (Egelman & Schechter 2013). Because the effects of channel aesthetics on attention are complex and have been studied extensively in the literature, we do not focus on this aspect for the purposes of empirical testing in our model.

The next step is *comprehension*. Comprehension evaluates an individual's level of understanding of the message itself and the consequences associated with disregarding the message. Given the variety of knowledge among users, warning messages should be crafted to target the least-skilled user to ensure all recipients understand the messages (Wogalter & Laughery 1996). When examining messages related to information security, technical jargon may increase the difficulty of comprehension for some users and should be avoided where possible (Bravo-Lillo et al. 2011a; Bravo-Lillo et al. 2011b).

The last three stages in the receiver portion of C-HIP are *attitudes and beliefs*, *motivation*, and *behavior*. Once a user pays attention to and comprehends a warning, their attitudes and beliefs can be changed, which is essential to affect their motivation to behave in a certain way. C-HIP postulates that in order for the receiver to change their behavior based on the communication from the source, they must be motivated by attitudes and beliefs. This part of the C-HIP process corresponds to several research models regarding the effects of attitudes and beliefs on ultimate behavior (e.g., Theory of Reasoned Action, Theory of Planned Behavior, Protection Motivation Theory). Motivation, or intentions, explains the connection between attitudes/beliefs and behavior in these models.

To examine the effects of attitudes and beliefs on the success of warning messages, we utilize Protection Motivation Theory (PMT), which is based on the Health Beliefs Model (HBM). Although some information security research has been conducted using HBM (e.g. LaRose et al. 2005a, 2005b; Ng, Kankanhalli, & Xu 2009; Ng & Xu 2007; Woon, Tan, & Low 2005), recent literature has primarily examined PMT in the information security context. Several studies have evaluated PMT together with the theory of planned behavior (Anderson & Agarwal 2010; Ifinedo 2012), deterrence (Herath & Rao 2009), habit (Vance, Siponen, & Pahnila 2012), bring your own device (Loraas et al. 2014), fear appeals (Johnston & Warkentin 2010; Johnston, Warkentin, & Siponen 2015), spyware/malware (Gurung, Luo, & Liao 2008; Lee & Larsen 2009), and user interface (Vance, Lowry, & Egget forthcoming), among others, to determine its impact on information security behaviors. Despite this extensive amount of research, these theories have not been fully applied to the effects of digital warning messages.

According to the Health Beliefs Model, to change users' behavior, three conditions have to be met: (1) the individual must be personally susceptible to the health problem; (2) the individual should understand that risk can lead to serious harm; and (3) the individual must understand what actions can be taken to avoid harm and the costs or benefits of those actions (Janz & Becker 1984; Rogers 1975; Witte 1992, 1994). PMT examines the same three conditions, respectively named perceived threat susceptibility, perceived threat severity, and perceived response efficacy (Rogers 1975). In addition, PMT introduces a fourth condition, perceived self-efficacy, which is the perception that one can successfully enact a recommended response (Maddux & Rogers 1983). Finally, each condition can be affected by the receiver's personal characteristics (e.g., demographics) and by environmental factors. Such factors should always be considered in models of warning effectiveness (Wogalter, 2006).

While some studies on digital warning messages have addressed some individual aspects of the C-HIP model, none have examined it in one holistic model. Little research exists that incorporates the C-

HIP model and, where research does exist, it mostly explains human processing without directly applying, testing, or adapting the model to the warnings context (e.g. Bravo-Lillo et al. 2011a; Chen et al. 2014; Schmuntzsch, Sturm, & Roetting 2014; Wogalter & Conzola 2002). For example, one study (Egelman et al. 2008) cites the C-HIP model as relevant to computer warning messages, but does not empirically investigate the model regarding the effectiveness of web browser phishing warnings.

Most research on computer-based IT security warning messages has focused on the channel (e.g. Bravo-Lillo et al. 2013; Egelman et al. 2008; Egelman & Schechter 2013; Sunshine et al. 2009b) or the attention aspects of warnings (e.g. Anderson et al. 2014a; Anderson et al. 2014b). Few research studies have considered a holistic view of examining attention together with the warning message content and its influence on attitudes and beliefs. It is important to understand how the content of the message can affect the attitudes, beliefs, and motivation of users, but also consider these effects when measuring the user attention and comprehension of the warning message. This study seeks to fill this gap by manipulating the content of a warning message, using theory-based methods, and measuring its effects on user behavior where the user pays attention to the warning.

Additionally, this study examines the impact of providing an alternative course of action to lower compliance costs. In other words, even if users have attitudes or beliefs that certain behaviors could be insecure, they may not have the motivation to behave securely if their attitudes and beliefs about the threat do not reflect a knowledge of alternative courses of action (Wogalter 2006a).

Hypothesis development

To gain a more holistic understanding of the effect of digital warning messages, we propose a variance model that combines HBM and PMT with the C-HIP model to examine attention, comprehension, attitudes, beliefs, motivation, and behavior. This model is depicted below in Figure 2. We present the associated hypotheses starting with the dependent variable and then working back from there.

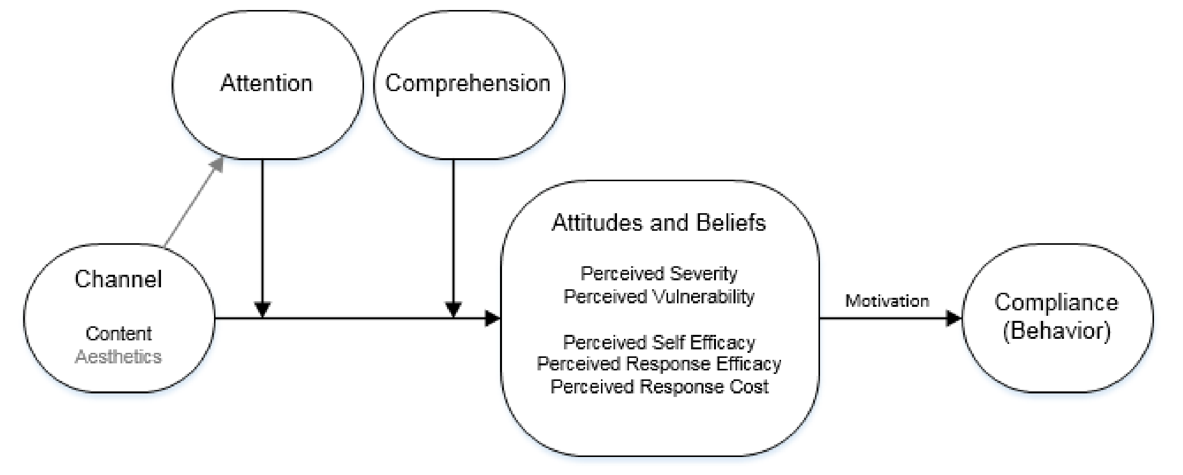


Figure 2. C-HIP variance model and study hypotheses.

Attitudes and Beliefs, Motivation, and Behavior

C-HIP postulates that in order for the receiver to change their behavior based on the communication from the source, they must be motivated by attitudes and beliefs. This part of the C-HIP process corresponds to several research models regarding the effects of attitudes and beliefs on ultimate behavior (e.g., Theory of Reasoned Action, Theory of Planned Behavior, Protection Motivation Theory, etc.). Motivation, or intentions, explains the connection between attitudes/beliefs and behavior in these models. Based on this large body of literature, we theorize that, after receiving the warning message, a user's attitudes and beliefs will motivate the user's ultimate compliance or violation behavior.

Specifically, to test the effects of attitudes and beliefs on user behavior, we use Protection Motivation Theory. This theory states that persuasive communication in the form of fear appeals increases motivation to protect oneself based on one's threat appraisal and coping appraisal (Maddux & Rogers 1983; Rogers 1975). The threat appraisal includes perceptions of threat severity and threat susceptibility. The coping appraisal includes perceptions of response efficacy, self-efficacy, and response cost.

Witte (1992, 1994) states that threats exist regardless of whether or not they are perceived by individuals. PMT and its extensions have identified that fear appeals should be constructed to not only help individuals identify existing threats but also understand the seriousness of these threats and their probability of occurrence (Rogers 1975; Witte 1992, 1994). As individuals become more conscious of a threat, they will establish beliefs about its seriousness and its probability of occurrence (Johnston et al. 2015).

Perceived threat severity refers to the user's conviction regarding the seriousness of the problem and its associated consequences (Ng et al. 2009; Ng & Xu 2007; Rogers 1975). In information security, the risk of any given threat is not only limited to damaging organizational assets (e.g. information, systems, operations), but could include other adverse outcomes such as negative impacts on job performance. Considering the impact these risks may have on organizational or individual performance, users' perception of threat severity will vary; therefore, increasing understanding regarding the consequences of insecure behavior may motivate compliance behavior. For example, successful warning messages that increase perceptions regarding a given threat may affect individual attitudes to engage in harmful behavior.

Perceived threat susceptibility is the user's conviction regarding the probability that a threat will occur (Rogers 1975). The level of a threat's severity is known to impact vulnerability (Milne, Sheeran, & Orbell 2000). Rogers (1975) further states that as compared to low-fear appeals, high-fear appeals increase perceptions of threat susceptibility and facilitate attitude change. In this same study, a unique motivator to help smokers become aware of their personal vulnerability was to role-play with them. Similarly, warning messages could resemble role-playing by heightening individual awareness of personal vulnerabilities. This awareness would facilitate attitude change against engaging in harmful behavior. Therefore, warning messages should not only address the severity of threats but also its probability of occurrence. Base on the above rationale, we hypothesize that:

H1a. Individual perceptions of threat severity, after experiencing a warning message, are positively associated with compliance behavior.

H1b. Individual perceptions of threat vulnerability, after experiencing a warning message, are positively associated with compliance behavior.

As part of the coping appraisal, effective fear appeals should cause individuals to form cognitions about efficacy, which include perceptions about the recommended response (i.e., response efficacy) and perceptions of efficacy of the individual performing the response, i.e., self-efficacy (Witte 1994). Response efficacy is “the degree to which an individual believes the response to be effective in alleviating a threat” (Johnston & Warkentin 2010; Rogers 1975). Self-efficacy pertains to the level of confidence in “one’s ability to undertake the recommended preventive behavior” (Herath & Rao 2009; Maddux & Rogers 1983).

Despite the actual response efficacy against a specific threat, users still cognitively evaluate the efficacy of the response and develop their own perceptions. Ultimately, this evaluation will determine whether the user follows the recommendation (Maddux & Rogers 1983). If the response is considered moderately or highly effective then users are likely to enact a recommended response (Maddux & Rogers 1983). Given that users may not know the recommended response to a particular threat, warning messages serve as an appropriate medium for users to identify the most effective response to overcome the threat.

Even if users believe that a response is effective, they also consider whether or not they are capable of enacting the recommended response (Maddux & Rogers 1983; Witte 1992). Events that spark high levels of emotional arousal (e.g., perceived security threats) are known to negatively impact self-efficacy (e.g., lower perceptions of one’s ability to use a computer) (Marakas, Yi, & Johnson 1998). Therefore, as users perceive threatening events (e.g., viruses, spyware) as severe and probable, they doubt their ability to function adequately within the heightened threat conditions (Johnston & Warkentin 2010). Therefore, when users view warning messages which convey how to complete a recommended response, doubts may diminish about one’s own ability and attitude may increase about enacting the recommended response. Thus, we posit that:

H1c. Individual perceptions of self-efficacy, after experiencing a warning message, are positively associated with compliance behavior.

H1d. Individual perceptions of response efficacy, after experiencing a warning message, are positively associated with compliance behavior.

A third factor of the coping appraisal is response costs. Response costs are those that are incurred when coping with the threat (Lee & Larsen 2009). Contrary to the other aspects of the coping appraisal, response costs preclude individuals from enacting a particular response. Given an extensive amount of time, effort, money, and other requirements to adopt the recommend response, individuals may decide to forego this effort (Milne et al. 2000). When evaluating response costs within the realm of information security, security practices are often considered a hindrance leading employees to neglect security practices (Herath & Rao 2009). Therefore, as the cost to heed warning messages increases, compliance with these warning messages will decrease. We then hypothesize that:

H1e. Individual perceptions of response cost, after experiencing a warning message, are negatively associated with compliance behavior.

Channel Effects on Attitudes and Beliefs

As individuals encounter warning messages, they will formulate attitudes and beliefs about the threat appraisal and coping appraisal, in other words, they will develop cognitions about the repercussions of a given threat and the likelihood of its occurrence together with how to combat this threat. For example, a successful warning message may identify one or more responses to a given threat with steps to enact the recommended response(s). As a result, perceptions related to response efficacy and self-efficacy will increase. Although PMT and the Health Beliefs Model explain some effects that impact attitudes and beliefs, this study expounds on current research by identifying whether warning message content will be strong enough to actually changes attitudes and beliefs.

Even when users change their attitudes or beliefs, the C-HIP model shows that these attitudes or beliefs must lead to motivation in order to ultimately change behavior. In this step, it is essential to understand the cost of compliance (e.g. effort, time, money). Often, even if users are aware of the benefits or consequences of compliance/noncompliance, they may not be motivated to comply because they are not aware of alternative safe courses of action (Wogalter 2006a).

One challenge with existing warning designs is that they usually do not offer any alternative solutions that may influence user's behavior. Most of the web based computer warnings offer two choices to the user (e.g., "proceed" or "cancel"), and, as such, present a limited binary decision making process. Only recently, web browsers (e.g., Google Chrome) started to incorporate alternative options to the user such as a "help me understand" link where the user is presented with a third option (Sophos 2015). In their recent study (Felt et al. 2015) introduced opinionated design, defined by "the use of visual design cues to promote a recommended course of action", which resulted in substantially increased adherence rates (nearly 30% more total users decided to change their course of action), positively impacting user safety and decision making. Therefore, successful warning messages should (1) increase cognizance about the severity of the threat, (2) identify the likelihood that threat will occur, (3) pinpoint effective responses to a given threat, (4) stimulate one's perception about their ability to complete the recommended response, and (5) diminish costs associated with the recommended response. One unique way to diminish perceived costs of compliance that we propose is that warning messages should include alternative courses of action to the user. That is, if users heed a digital warning telling them not to use a certain software, it could be more effective if the warning message also suggested a way to reduce the costs of compliance by presenting an alternative way to complete the work. By presenting alternative courses of action, users are better able to cope with the threat presented and will better understand the efficacy of compliance and perceive reduced costs of compliance. In summary, we hypothesize:

H2a. Successful warning messages that identify the threat severity are positively associated with attitudes and beliefs about compliance.

H2b. Successful warning messages that identify the threat susceptibility are positively associated with attitudes and beliefs about compliance.

H2c. Successful warning messages that identify the effective response(s) to the threat are positively associated with attitudes and beliefs about compliance.

H2d. Successful warning messages that stimulate self-efficacy are positively associated with attitudes and beliefs about compliance.

H2e. Successful warning messages that reduce response costs through providing alternative courses of action are positively associated with attitudes and beliefs about compliance.

Attention

The main objective for many designers of warning messages is to capture users' attention and convey information about the possible hazard (Bravo-Lillo et al. 2013). The warning design is composed of several different aspects which may impact its attractiveness so it is more visible, and consequently, has higher impact on the user. Consequently, in this communication delivery process, if a user's attention is switched to the warning message, we can expect to see increased compliance and better decision making. However, this is not a direct effect. In fact, in our model we do not theorize that paying attention to a warning message will have a direct effect on compliance. Rather, following the C-HIP model, we propose that attention is necessary in order for warning messages to affect attitudes and beliefs. Without this attention, the warning message will have no effect. But simply getting a user to pay attention to a warning is not enough in itself to affect ultimate behavior. Thus, we hypothesize the following interaction effect:

H3. Warning messages positively influence attitudes and beliefs only when users pay careful attention to the warning.

Comprehension

Once a warning message captures a user's attention, the next step is comprehension. One common mistake of warning designers is to assume that an average user will understand the hazard and its consequences and risks (Wogalter & Laughery 1996). Given that users range from novices to experts, they act and behave differently based on their level of warning comprehension. For example, due to technical complexity, novice users may not fully understand what is an SSL warning (a web browser warning that appears when a potentially unverified secure connection is about to be established with the remote server). When general users were presented with terms such as "startup disk, encryption, virus, attachment, macro, and certificate," they indicated that they heard of them but had difficulties to make sense of them (Bravo-Lillo et al. 2011a; Bravo-Lillo et al. 2011b). Therefore, the content of warning messages should target the least-skilled users because messages that include complex technical terms are not likely to be comprehended (Wogalter & Laughery 1996). For example, Sunshine et al. (2009b) found that users who understood the web browser SSL warnings behaved very differently compared to those who did not understand them. Thus, we expect that users, regardless of their technical expertise, who fully comprehend warning messages will behave differently. Thus, following the C-HIP model, we hypothesize a second moderating effect:

> *H4. Warning messages positively influence user attitudes and beliefs only when users fully comprehend the meaning of the warning.*

METHODS

General research design

To test the hypotheses, we design a series of field and lab experiments where users either download an application from the Web or use a Web browser-based application. The applications used in the studies were created by one of the authors using VB.net. After running the application, users experience a variation of a warning message created for this study. The warning message briefly describes the consequences of using the 'potentially insecure' software and asks the user if they would like to "Continue" to use the software or "Exit." After the user presses continue, the message disappears and they can use the software. If the user presses exit, the application is closed.

In each study, three measures are collected when the user opens the application for the first time. First, the user's *decision* is recorded as 0 if he or she clicks on "Exit" or 1 if he or she clicks on "Continue." Second, we record the *duration* (in ms) from the start of the application until the decision. Finally, we record a unique PC identifier (IP address and/or MAC address) of the user so that we can measure the behavior of the users only the first time they encounter the warning, and potentially measure any interesting behavior that occurs if/when users subsequently open the application another time.

We vary the possible warning messages that users can see and test their individual effects on the user's decision and time taken to make the decision.

Pilot Study Design

Pilot data was collected from March to May, 2015. This pilot data was used only to test our *procedures*, but not our *theory* (which continues to be refined). The pilot study was also done to analyze general differences between behavior when users see a warning message vs. when they do not. In addition to this pilot data, we also plan to conduct an expert panel review and further validity testing of our instrument before completing a full-scale data collection.

In the pilot data, we loosely used the Health Beliefs Model and PMT to guide the design of various types of warning messages. While many organizational programs only focus on negative consequences of engaging in insecure behavior (e.g., deterrence theory, protection motivation theory), they neglect the benefits and costs of NOT performing the insecure behavior. Based on the Health Beliefs Model, users of this study are presented with warning messages that focus on the consequences (both positive and negative) of doing the "right" thing (e.g. benefits and drawbacks of discontinuing unauthorized software use) rather than the consequences of doing the "wrong" thing (e.g. sanctions for continuing to use the unauthorized software). As this is a pilot study, the full-scale study will contain different treatments and wording.

Participants

The participants in the pilot data collection are general Internet users who download the PDF application after finding it on SourceForge. Participants are not directly recruited for the study, nor are they aware of its purpose, which improves the validity of results.

Procedures

In the pilot study, users downloaded an application from the Web designed to manipulate PDF documents (e.g., split, merge, etc.). The application was created by one of the authors using VB.net and placed on SourceForge[1]. When the user opens the application, either a warning message or a control message popped up. The display of the warning or the control message is controlled by the random

[1] http://sourceforge.net/projects/pdfsplitextractandmerge/?source=navbar

function within the application; however, the function was set to favor warning messages over control messages because there were many variations of the warning message but only one type of control message. The particular warning message that any given user sees is also controlled by the random function within the software.

The warning message briefly describes the consequences of using unauthorized software and asks the user if they would like to “Continue” to use the software or “Exit.” The non-warning message (control group) simply thanks the user for using the application and proposes two options (“Continue” or “Exit”), which lead to the same outcomes. The control treatment was used in this pilot study simply to compare the effect of using any warning message against the behavior of users when they do not see a warning message. The control treatment is only for the pilot study to analyze general differences in behavior between seeing a warning message or not. Such a control treatment likely will not be needed in future data collection (instead, we will simply compare different versions of warning messages).

Warning Variations

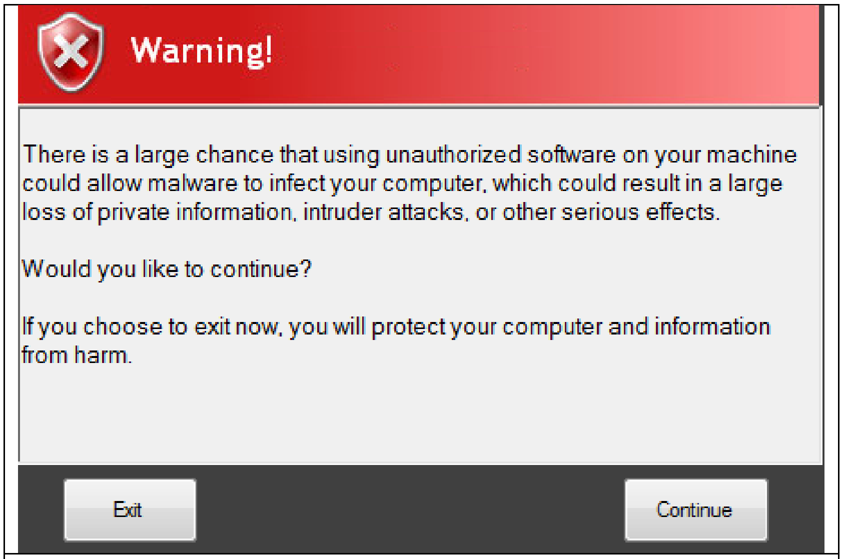
In this study, there are 16 (2 x 2 x 2 x 2) variations of the warning message: high/low severity x high/low certainty x benefit listed/not listed x barrier listed/not listed. The warning message starts by describing the threat from using unauthorized software, describing the relative certainty and severity of the threat occurring.

- High certainty: There is a large chance that using unauthorized software on your machine could allow malware to infect your computer…
- Low certainty: There is a small chance that using unauthorized software on your machine could allow malware to infect your computer…
- High severity: …which could result in a large loss of private information, intruder attacks, or other serious effects.
- Low severity: …which may be a nuisance to address.

Along with asking the user if they would like to continue using the software, the warning message also includes some combination of listing (or not listing) the benefit and/or barrier of exiting the unauthorized software in order to comply with good practice.

- Benefit: If you choose to exit now, you will protect your computer and information from harm.
- Barrier: If you choose to exit now, you will need to find a different, authorized software to perform your work.

An example warning message (high certainty / high severity / benefit listed / barrier not listed) is shown in Figure 3 below.


Warning!

There is a large chance that using unauthorized software on your machine could allow malware to infect your computer, which could result in a large loss of private information, intruder attacks, or other serious effects.

Would you like to continue?

If you choose to exit now, you will protect your computer and information from harm.

Figure 3. Example warning message from pilot study

Pilot Study Results

Data Validation

Because the pilot study was collected as we developed our procedures, there is a range of quality in the data collected. When data collection began, we did not account for users attempting to open the software multiple times. Thus, part way through the collection we started collecting the IP addresses of users. Although the data suggests that most IP addresses are from the same user (since the same IP address accessed the software within minutes), IP addresses are limited in that multiple different users could be identified with the same IP address (in the case of proxy addresses). Thus, we report the results of (1) the full data set with all responses (including those with no IP address recorded and multiple attempts by the same user; n=365, including 253 receiving a warning message) and (2) the results of only the first attempt of a recorded IP address (n=186, including 90 receiving a warning message).

Analysis with Continue/Exit Decision as Dependent Variable

A summary of responses for users who saw a warning message is shown in Tables 1 and 2 below. Most users chose to "continue" using the software after seeing the warning message, although there were some differences between warning types.

Table 1. Summary Data for Full Data Set

Factor	n	# pressed "Continue"	# pressed "Exit"	% pressed "Continue"
Severity-High	125	88	37	70.4%
Severity-Low	128	107	21	83.6%
Certainty-High	123	96	27	78.0%
Certainty-Low	130	99	31	76.2%
Benefit-Present	136	104	32	76.5%
Benefit-Absent	117	91	26	77.8%
Barrier-Present	129	92	37	71.3%
Barrier-Absent	124	103	21	83.1%
Total Warning Treatment	**253**	**195**	**58**	**77.1%**
Non-warning	112	100	12	89.3%

Table 2. Summary Data for First Time Viewing a Warning Message from Unique IP Addresses

Factor	n	# pressed "Continue"	# pressed "Exit"	% pressed "Continue"
Severity-High	45	35	10	77.8%
Severity-Low	45	36	9	80.0%
Certainty-High	41	31	10	75.6%
Certainty-Low	49	40	9	81.6%
Benefit-Present	52	39	13	75.0%
Benefit-Absent	38	32	6	84.2%
Barrier-Present	45	33	12	73.3%
Barrier-Absent	45	38	7	84.4%
Total Warning Treatment	**90**	**71**	**19**	**78.9%**
Non-warning	96	85	11	88.5%

In the full data set, it appears that the "severity" statement and the "barrier" statement are more powerful than the other factors. Contrary to expectations, users were more likely to heed the warning and exit the software when a barrier to compliance was listed. We would have expected that listing a barrier to comply would make users less likely to heed the warning, but the opposite occurred. This is confirmed with the multiple logistic regression analysis summarized in Table 3 below. In this analysis, the "severity" statement had a statistically significant effect (p=0.025), and the "barrier" statement had a strong effect as well (p=0.051).

Table 3. Logistic Regression on Full Data Set

	B	S.E.	Wald	Sig.	Exp(B)
Severity	-0.702	.313	5.019	.025	0.496
Certainty	0.020	.307	0.004	.948	1.020
Benefit	-0.013	.308	0.002	.966	0.987
Barrier	-0.611	.313	3.796	.051	0.543
Constant	1.932	.373	26.863	.000	6.907

In the partial data set, there are not strong effects of any of the factors, although the sample size is quite small. The multiple logistic regression on this data set is summarized in Table 4 below. The treatments (particularly pertaining to the "certainty" variable, which only had one word difference between the high and low conditions) may not have been strong enough in this pilot study. The wording and treatments were not validated and refined for the pilot collection.

Table 4. Logistic Regression on Filtered Data Set

	B	S.E.	Wald	Sig.	Exp(B)
Severity	-0.064	.530	0.015	.904	0.938
Certainty	-0.356	.532	0.448	.503	0.700
Benefit	-0.503	.561	0.805	.370	0.605
Barrier	-0.683	.539	1.607	.205	0.505
Constant	2.207	.642	11.821	.001	9.089

Analysis with Time as Dependent Variable

Next, we examined the amount of time users take to make a decision of whether or not to heed the warning message, and how this is affected by the warning message and its contents. Because we cannot simply compare the average duration, due to the right skewed distribution, we use an event history analysis technique, also known as a survival technique. In these techniques, the time between the display of the warning message (or the control message) until the user's choice (Continue or Exit) is referred to as the *hazard event*, which has a beginning and an end. We timed this event by calculating the difference between the start and the end of the event, which provided us the total duration of the event.

Specifically, the survival technique we used is the Kaplan-Meier Survival estimator (Kaplan & Meier 1958), which enables dealing with differing survival times (i.e., times-to-event), especially when not all the subjects continue in the study (Rich et al. 2010). The Kaplan-Meier survival estimator is largely used in medical research to measure patients and the amount of time they live after receiving a specific treatment. It is also widely used in many other contexts such as determining the length of time people remain unemployed after losing their jobs. In our context, we are using the Kaplan-Meier survival estimator to understand behavior throughout the duration after being exposed to warning or non-warning messages.

In Figure 4 we plot the cumulative survival functions for the control treatment vs. the warning treatment (which includes all different subgroups). The cumulative survival function represents the cumulative survival proportion against time for each treatment. The lengths of the horizontal lines along the X-axis represents the survival duration for that interval. The interval is terminated by the occurrence of the event of interest (Continue =1 or Exit = 0). In other words, the x-axis denotes time (in our case expressed in milliseconds). The y-axis denotes the percentage of subjects who have survived. Also, each drop in the curve represents an event (i.e. one user entry point). In simpler words, the Kaplan-Meier survival estimator enables us to understand if warning messages have any effect on the time taken during the user's decision-making process; hence, if a warning is presented to the user, we would expect the user to spend more time reflecting on the action to take. More precisely, if the warning message has any effect on the user's behavior, then we should see longer durations of the event where the user spends more time heeding the warning message and trying to making sense of it.

We can clearly see that the cumulative survival proportion of the warning group is higher compared to the control group. In other words, users who received the warning treatment have a higher survival rate than those that received the non-warning treatment (i.e., control group).

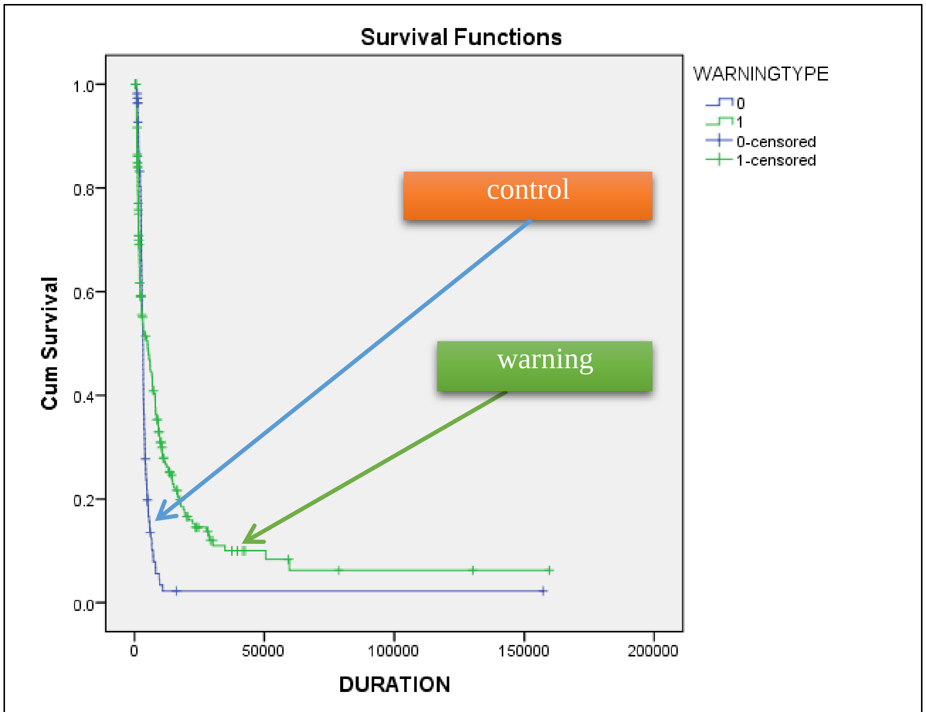


Figure 4. Kaplan-Meier survival function

A log rank test was run to determine if there were differences in the survival distribution for the different types of interventions (warning vs no-warning). The survival distributions for the two interventions were statistically significantly different (see Table 4), χ2(2) = 14.974, p < .0005.

Table 4. Test of equality of survival distributions for the different levels of warning type.

	Chi-Square	p value
Log Rank (Mantel-Cox)	14.974	.000
Breslow (Generalized Wilcoxon)	.825	.364
Tarone-Ware	5.581	.018

Finally, to understand if the effect of a warning on the duration of the event is significant (i.e. whether warning message really impacts the user's behavior) we use Cox proportional-hazard regression (Box-Steffensmeier & Jones 2004).

The results from the estimate Cox model are presented in Table 5. The hazard ratio estimate of the warning type (Exp(B)=1.650, Coefficient 0.131 with p<0.0005) indicates that in the presence of a warning message, users are 1.65 times more likely to exit the software.

Table 5. Cox regression model over warning type

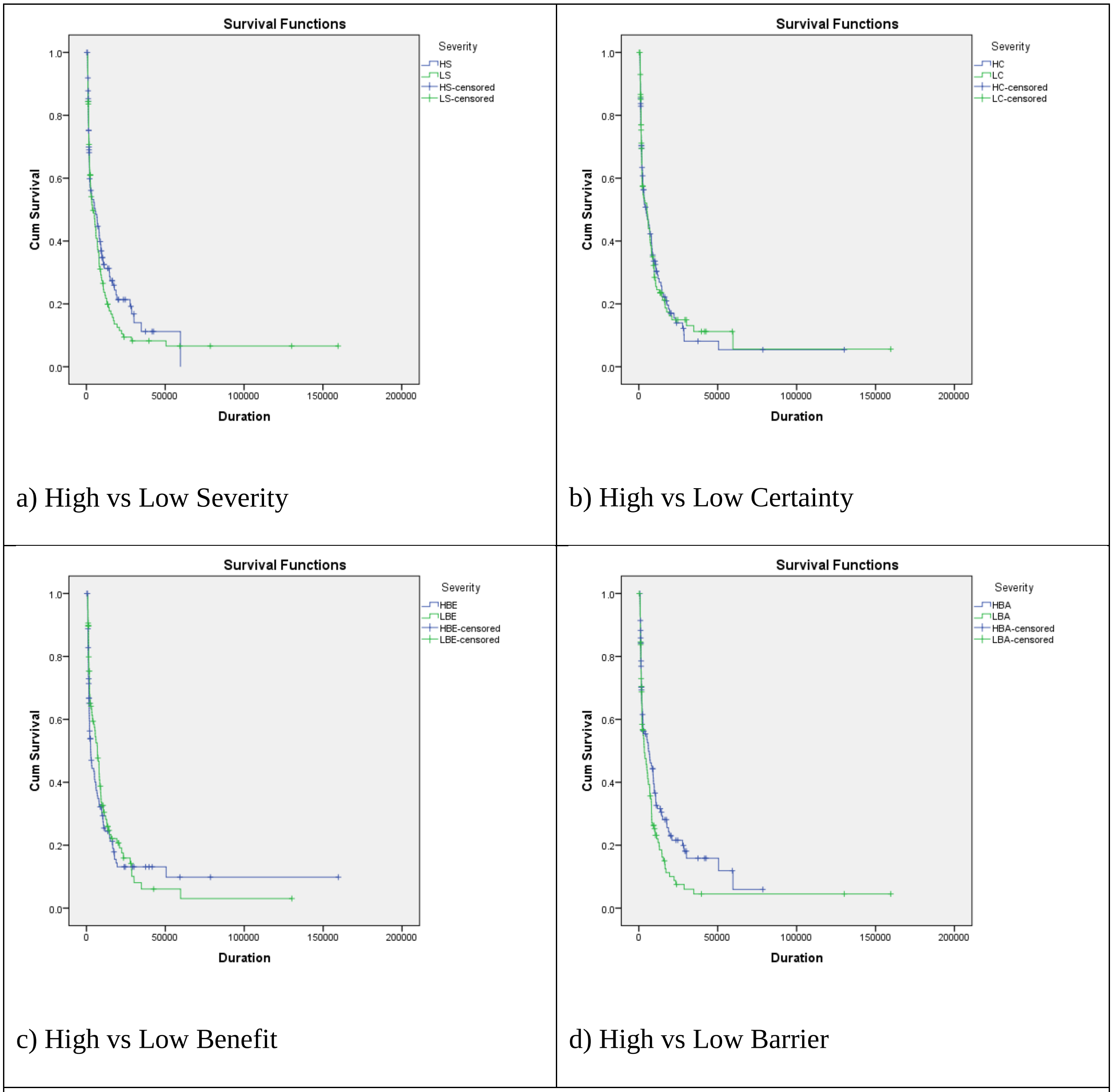
	B	SE	Wald	Sig.	Exp(B)	95.0% CI for Exp(B)	
						Lower	Upper
Warning vs Control	.501	.131	14.703	<.001	1.650	1.277	2.131

In Figure 5, we present the four different warning types and their corresponding survival distribution times. The cumulative survival proportion of the 'High Severity (HS)' has higher survival rate compared to the 'Low Severity (LS)'. The same result can be seen with 'High Barrier (HBA)' vs. 'Low Barrier (LBA)'. However, 'High Certainty (HC)' vs. 'Low Certainty (LC)' and 'High Benefit (HBE)' vs. 'Low Benefit (LBE)' have very similar cumulative survival proportions which indicates that the warning message (high vs. low) does not have any impact on the survival rate. However, the Log rank test (Table 6) showed that only the survival distribution of HBA vs. LBA is statistically significantly different, $\chi2(2) = .002$, $p < .05$. For HS vs. LS the Log rank test showed $\chi2(2) = 1.516$, $p > .05$ which means that we do not have a statistically significant result between high severity compared to low severity. This could mean that users could not really differentiate the high vs low severity warning message in terms of the risk behind them.

Table 6. Overall comparisons for Severity, Certainty, Benefit and Barrier

	HS vs LS		HC vs LC		HBE vs LBE		HBA vs LBA	
	Chi-Square	Sig.	Chi-Square	Sig.	Chi-Square	Sig.	Chi-Square	Sig.
Log Rank (Mantel-	1.516	.218	.001	.977	1.214	.270	5.002	.025

Cox)								
Breslow (Generalized Wilcoxon)	.379	.538	.005	.941	4.479	.034	2.177	.140
Tarone-Ware	.914	.339	.013	.908	3.323	.073	3.458	.063



a) High vs Low Severity

b) High vs Low Certainty

c) High vs Low Benefit

d) High vs Low Barrier

Figure 5. Kaplan-Meier survival function for Severity, Certainty, Benefit, and Barrier

Overall, we can see that in the presence of a warning message, the duration of the decision-making process is affected. Further, when comparing the four main factors manipulated in the warning messages (severity, certainty, benefit of compliance, and barrier to compliance) and their effectiveness on users' decision making, we see that only the manipulation of the presentation of severity in the warning message has a strong impact compared to three other warning types.[2]

Full-Scale Data Collection (Proposed)

Based on feedback from information security experts in academia and industry on our pilot results, we will implement a full scale data collection that, unlike the pilot study, will also include survey measures to more fully assess the complete theoretical model.

Participants

We have three potential different samples for the full scale study: (1) students recruited from multiple universities with an extra credit or course credit incentive to participate, (2) Internet users recruited using Mechanical Turk, and/or (3) employees recruited from a large international company. We have gained access to all three potential subject pools, and we hope to collect from all of them. Collecting from multiple sources will strengthen our study in multiple ways. First, the sample size in some of the populations may be small (e.g., in the industry company, we may only be able to access 70-90 responses) so multiple sources will increase the sample size. Second, testing the hypotheses in multiple contexts with multiple types of participants will increase the generalizability of the results. Each of the potential subject pools has particular strengths. We will report the results of the three field/lab studies separately.

Procedures

Because the participants are aware they are taking part in a research study, different procedures must be used in order to disguise the hypotheses and solicit honest behavior.

[2] For the sake of brevity, only the analysis on duration for the full data set is reported. In the full-scale data collection, analysis will be done that validates only unique users of the software.

Participants are first instructed that the study will require them to complete some work that involves either (1) editing PDF documents (similar to the application used in the pilot study) or (2) completing an online task on Mechanical Turk. In the first option, the editing will include merging, splitting, and/or extracting pages from PDFs. Participants are then instructed that in order to complete these requirements, they will need to download special software in order to complete the task. Participants are then given a list of links to software applications that could be downloaded in order to complete the work. Unbeknownst to the participants, the list of links actually all lead to the same software, and they are only able to click on one link. Once the participants download the software and open it, one of several warning messages will appear.

Regarding the online task option, participants from Mechanical Turk will use a Web application to rate their preference for a particular product photo. This alternative web application is necessary for collecting data in Mechanical Turk because these users will not be able to download the PDF software. However, the procedures will be similar. The participants will be instructed that the study will require them to rate product photos. When the user clicks “Start” to open the Web application, a variation of a warning message will appear, warning the user about potential dangers of using the Web application.

Using both the downloaded PDF application and the online Web application also strengthens our proposed study in multiple ways. First, using different variations of software allows us to access multiple types of subject pools. Second, our results become more generalizable (e.g., they will not be limited to warning messages on software alone or warning messages in Web applications alone).

After both scenarios, participants are directed to a survey that informs them the warning message was part of the study and asks several questions about variables of interest in the study. There could also be a pre-survey to ask some of the survey questions before users are directed to the task that results in a warning message.

Treatments

When the user opens the application or accesses the website to complete the task, a warning message will pop up. Several variations of the warning messages are used as the treatment in this study.

The particular warning message that any given user sees is random determined by the random function within the software.

Validation of instruments

The procedures and treatments of this study will be refined based on (1) discussion at the IFIP IT Security workshop and (2) results of an expert panel to be completed after the workshop.

DISCUSSION

The present paper examines how theory-based communication, through the use of warning messages, influences information security policy compliance. Specifically, based on the C-HIP model and supported by Health Belief Model and Protection Motivation Theory, we suggest a new content element of computer security warning messages (i.e., suggesting alternative secure courses of action) that users who pay attention to warning message content may be more persuaded to behave securely. Through the preliminary pilot study, we observe some interesting findings.

The pilot study results suggest that both the "severity" and "barrier" statements are more powerful than the "certainty" and "benefit" statements. This is an interesting finding which indicates that users are more likely to heed the warning message and exit the software when a barrier to compliance statement is present. We expected the "benefit" statement to have higher importance, than indicated by our preliminary results, granted that users who feel their data is susceptible to compromise would engage in behaviors to properly safeguard their data. We believe this issue may be due to the wording that was included in the warning message. In subsequent studies, the warning message will be refined and undergo expert panel reviews to better determine the impact of barrier statements on information security policy compliance.

Additionally, the presence of a warning message does not lead to an immediate incident termination. Over 77% of users clicked on 'Continue' and used the application despite the warning message. This is consistent with previous studies showing that users often ignore warning messages (Maimon et al. 2014). However, we can clearly see that in the presence of a warning message, the duration of the decision-making process is affected. As expected, users put more time in trying to make sense of the warning message and better understand how warning content may impact them. Interpretation of this could be that users spend more time in thinking, reading and reflecting about the content of the warning message and the consequences their acts may have if they click on ‘continue’ vs the ‘exit’ action. That is, even though a majority chose to disregard the potential consequences presented by the warning message, it may not be only because of a lack of paying attention. While some users were inattentive to the message, many stopped to carefully consider the warning before ultimately deciding not to heed its contents. This suggests that warning message may activate an automatic cautious behavior (Bargh, Chen, & Burrows 1996). This further highlights the need for studies such as ours that examine both attention as well as attitudes and beliefs, because users could ignore the warning message for either of these reasons.

Further refinement and additional lab and field experiments will be conducted to improve warning message content and reduce issues such as technical jargon, which is quite problematic for some users (Modic & Anderson 2014). Specifically, theory-driven communication will be presented to users that identifies risks to which they may be exposed. Given the vast difference in the mental approach between novice and advanced users (Bravo-Lillo et al. 2011a), such theory-driven communication will be useful in distinguishing between all users. Moreover, a challenge users frequently experience is the binary decision of continuing or exiting based on a warning message rather than alternative actions. Only recently, Felt et al. (2015) proposed opinionated design, which was implemented in Google Chrome. Though this is an additional step where the user is invited to think again before deciding, it is still limited due to the lack of a relevant alternative that would be useful and increase compliance.

Overall, further studies will increase understanding of how user decisions are impacted by the different stages of the C-HIP model. This model identifies each layer to increase information security policy compliance.

REFERENCES

Akhawe, D., & Felt, A. P. (2013). Alice in Warningland: A large-scale field study of browser security warning effectiveness. *Proceedings of the 22nd USENIX Conference on Security*: USENIX Association.

Anderson, B. B., Vance, A., Kirwan, B., Eargle, D., & Howard, S. (2014a). Users aren't (necessarily) lazy: Using neuroIS to explain habituation to security warnings. In *2014 International Conference on Information Systems (ICIS)*.

Anderson, B. B., Vance, A., Kirwan, B., Eargle, D., & Howard, S. (2014b). *Why users habituate to security warnings: Insights from fMRI*. Paper presented at 2014 IFIP 8.11 Dewald Roode Security Workshop.

Anderson, C. L., & Agarwal, R. (2010). Practicing safe computing: A multimethod empirical examination of home computer user security behavioral intentions. *MIS Quarterly, 34*(3), 613-643.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of personality and social psychology, 71*(2), 230.

Box-Steffensmeier, J. M., & Jones, B. S. (2004). *Event history modeling: A guide for social scientists*: Cambridge University Press.

Bravo-Lillo, C., Cranor, L. F., Downs, J., Komanduri, S., Reeder, R. W., Schechter, S., & Sleeper, M. (2013). Your attention please: Designing security-decision UIs to make genuine risks harder to ignore. In *SOUPS '13 Proceedings of the Ninth Symposium on Usable Privacy and Security*.

Bravo-Lillo, C., Cranor, L. F., Downs, J. S., & Komanduri, S. (2011a). Bridging the gap in computer security warnings. *IEEE Security and Privacy*, *9*(March/April), 18-26.

Bravo-Lillo, C., Cranor, L. F., Downs, J. S., & Komanduri, S. (2011b). Bridging the Gap in Computer Security Warnings: A Mental Model Approach. *Ieee Security & Privacy, 9*(2), 18-26.

Bushman, B. J. (2006). Effects of warning and information labels on attraction to television violence in viewers of different ages. *Journal of Applied Social Psychology, 36*(9), 2073-2078.

Chen, T.-C., Stepan, T., Dick, S., & Miller, J. (2014). An anti-phishing system employing diffused information. *ACM Transactions on Information and System Security (TISSEC), 16*(4), 16.

Coleman, S. (2007). The Minnesota income tax compliance experiment: replication of the social norms experiment. *Available at SSRN 1393292*.

Conzola, V. C., & Wogalter, M. S. (2001). A Communication-Human Information Processing (C-HIP) approach to warning effectiveness in the workplace. *Journal of Risk Research, 4*(4), 309-322.

Egelman, S., Cranor, L. F., & Hong, J. (2008). You've been warned: An empirical study of the effectiveness of web browser phishing warnings. *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems*.

Egelman, S., & Schechter, S. (2013). The importance of being earnest [in security warnings]. In A.-R. Sadeghi (Ed.), *Financial Cryptography and Data Security* (pp. 52-59). Okinawa, Japan.

Egilman, D., & Bohme, S. (2006). A brief history of warnings. *Handbook of Warnings. Lawrence Erlbaum Associates, Mahwah, NJ,* 35-48.

Felt, A. P., Ainslie, A., Reeder, R. W., Consolvo, S., Thyagaraja, S., Bettes, A., . . . Grimes, J. (2015). *Improving SSL Warnings: Comprehension and Adherence*. Paper presented at Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Republic of Korea.

Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of consumer Research, 35*(3), 472-482.

Grier, C., Tang, S., & King, S. T. (2008). Secure web browsing with the OP web browser. In *Security and Privacy, 2008. SP 2008. IEEE Symposium on*.

Gurung, A., Luo, X., & Liao, Q. (2008). Consumer motivations in taking action against spyware: An empirical investigation. *Information Management & Computer Security, 17*(3), 276-289.

Herath, T., & Rao, H. R. (2009). Protection motivation and deterrence: A framework for security policy compliance in organisations. *European Journal of Information Systems, 18*(2), 106-125.

Ifinedo, P. (2012). Understanding information systems security policy compliance: An integration of the theory of planned behavior and the protection motivation theory. *Computers & Security, 31*(1), 83-95.

Janz, N. K., & Becker, M. H. (1984). The health belief model: A decade later. *Health Education & Behavior, 11*(1), 1-47.

Johnston, A. C., & Warkentin, M. (2010). Fear Appeals and Information Security Behaviors: An Empirical Study. *Mis Quarterly, 34*(3), 549-566.

Johnston, A. C., Warkentin, M., & Siponen, M. (2015). An enhanced fear appeal rhetorical framework: Leveraging threats to the human asset through sanctioning rhetoric. *MIS Quarterly, 39*(1), 113-134.

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association, 53*(282), 457-481.

LaRose, R., Rifon, N., Liu, S., & Lee, D. (2005a). Online safety strategies: a content analysis and theoretical assessment. In *The 55th Annual Conference of the International Communication Association, New York City*.

LaRose, R., Rifon, N., Liu, S., & Lee, D. (2005b). Understanding online safety behavior: A multivariate model. In *The 55th Annual Conference of the International Communication Association, New York City*.

Laughery, K. R., & Wogalter, M. S. (2006). Designing effective warnings. *Reviews of human factors and ergonomics, 2*(1), 241-271.

Lee, Y., & Larsen, K. R. (2009). Threat or coping appraisal: Determinants of SMB executives' decision to adopt anti-malware software. *European Journal of Information Systems, 18*(2), 177-187.

Loraas, T. M., Crossler, R. E., Long, J. H., & Trinkle, B. S. (2014). Understanding compliance with BYOD (bring your own device) policies utilizing protection motivation theory: Bridging the intention-behavior gap. *Journal of Information Systems, 28*(1), 209-226.

Maddux, J. E., & Rogers, R. W. (1983). Protection motivation and self-efficacy: A revised theory of fear appeals and attitude change. *Journal of Experimental Social Psychology, 19*, 469-479.

Maimon, D., Alper, M., Sobesto, B., & Cukier, M. (2014). Restrictive deterrent effects of a warning banner in an attacked computer system. *Criminology, 52,* 33-59.

Marakas, G. M., Yi, M. Y., & Johnson, R. D. (1998). The multilevel and multifaceted character of computer self-efficacy: Toward clarification of the construct and an integrative framework for research. *Information Systems Research, 9*(2), 126-163.

McAffee. (2014). Net Losses: Estimating the Global Cost of Cybercrime.

Milne, S., Sheeran, P., & Orbell, S. (2000). Prediction and intervention in health-related behavior: a meta-analytic review of protection motivation theory. *Journal of Applied Social Psychology, 30*(1), 106-143.

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

Ng, B.-Y., Kankanhalli, A., & Xu, Y. C. (2009). Studying users' computer security behavior: A health belief perspective. *Decision Support Systems, 46*(4), 815-825.

Ng, B.-Y., & Xu, Y. (2007). Studying users' computer security behavior using the Health Belief Model. *PACIS 2007 Proceedings*, 45.

Rich, J. T., Neely, J. G., Paniello, R. C., Voelker, C. C., Nussenbaum, B., & Wang, E. W. (2010). A practical guide to understanding Kaplan-Meier curves. *Otolaryngology-Head and Neck Surgery, 143*(3), 331-336.

Rogers, R. W. (1975). A protection motivation theory of fear appeals and attitude change. *The Journal of Psychology, 91*, 93-114.

Sanders, M. S., & McCormick, E. J. (1987). *Human factors in engineering and design*: McGRAW-HILL book company.

Schmuntzsch, U., Sturm, C., & Roetting, M. (2014). The warning glove–Development and evaluation of a multimodal action-specific warning prototype. *Applied ergonomics, 45*(5), 1297-1305.

Schultz, P., & Tabanico, J. J. (2009). Criminal beware: a social norms perspective on posting public warning signs*. *Criminology, 47*(4), 1201-1222.

Sophos. (2015). Google redesigns security warnings after 70% of Chrome users ignore them Retrieved May 2015, from https://nakedsecurity.sophos.com/2015/02/03/google-redesigns-security-warnings-after-70-of-chrome-users-ignore-them/

Sunshine, J., Egelman, S., Almuhimedi, H., Atri, N., & Cranor, L. F. (2009a). Crying Wolf: An Empirical Study of SSL Warning Effectiveness. In *USENIX Security Symposium*.

Sunshine, J., Egelman, S., Almuhimedi, H., Atri, N., & Cranor, L. F. (2009b). Crying wolf: An empirical study of SSL warning effectiveness. In *18th USENIX Security Symposium*.

Vance, A., Lowry, P. B., & Egget, D. (forthcoming). Increasing accountability through user-interface design artifacts: A new approach to addressing the problem of access-policy violations. *MIS Quarterly*.

Vance, A., Siponen, M., & Pahnila, S. (2012). Motivating IS security compliance: Insights from habit and protection motivation theory. *Information & Management, 49*(3-4), 190-198.

Witte, K. (1992). Putting fear back into fear appeals: The extended parallel process model. *Communication Monographs, 59*(4), 329-349.

Witte, K. (1994). Fear control and danger control: A test of the extended parallel process model (EPPM). *Communication Monographs, 61*, 113-134.

Wogalter, M. S. (2006a). Communication-Human Information Processing (C-HIP) Model. In M. S. Wogalter (Ed.), *Handbook of Warnings* (pp. 51-61). Mahwah, NJ: Lawrence Erlbaum Associates.

Wogalter, M. S. (2006b). Purposes and scope of warnings. *Handbook of Warnings. Lawrence Erlbaum Associates, Mahwah, NJ*, 3-9.

Wogalter, M. S., & Conzola, V. C. (2002). Using technology to facilitate the design and delivery of warnings. *International Journal of Systems Science, 33*(6), 461-466.

Wogalter, M. S., & Laughery, K. R. (1996). Warning! Sign and label effectiveness. *Current Directions in Psychological Science*, 33-37.

Woon, I., Tan, G.-W., & Low, R. (2005). A protection motivation theory approach to home wireless security. *ICIS 2005 Proceedings*, 31.