Journal of the Association for Information Systems
JAIS

Special Issue

Using Measures of Risk Perception to Predict Information Security Behavior: Insights from Electroencephalography (EEG)

Anthony Vance
Brigham Young University
anthony@vance.name

David Eargle
University of Pittsburgh
dave@daveeargle.com

Bonnie Brinton Anderson
Brigham Young University
bonnie_anderson@byu.edu

C. Brock Kirwan
Brigham Young University
kirwan@byu.edu

Abstract

Users' perceptions of risks have important implications for information security because individual users' actions can compromise entire systems. Therefore, there is a critical need to understand how users perceive and respond to information security risks. Previous research on perceptions of information security risk has chiefly relied on self-reported measures. Although these studies are valuable, risk perceptions are often associated with feelings—such as fear or doubt—that are difficult to measure accurately using survey instruments. Additionally, it is unclear how these self-reported measures map to actual security behavior. This paper contributes to this topic by demonstrating that risk-taking behavior is effectively predicted using electroencephalography (EEG) via event-related potentials (ERPs). Using the Iowa Gambling Task, a widely used technique shown to be correlated with real-world risky behaviors, we show that the differences in neural responses to positive and negative feedback strongly predict users' information security behavior in a separate laboratory-based computing task. In addition, we compare the predictive validity of EEG measures to that of self-reported measures of information security risk perceptions. Our experiments show that self-reported measures are ineffective in predicting security behaviors under a condition in which information security is not salient. However, we show that, when security concerns become salient, self-reported measures do predict security behavior. Interestingly, EEG measures significantly predict behavior in both salient and non-salient conditions, which indicates that EEG measures are a robust predictor of security behavior.

Keywords: *Risk Perception, Information Security Behavior, NeuroIS, Self-reported Measures, EEG, Iowa Gambling Task (IGT), Laboratory Experiment, Security Warning Disregard.*

* Fred Davis was the accepting senior editor. This article was submitted on 30th April 2013 and went through three revisions.

Volume 15, Special Issue, pp. 679-722, October 2014

Using Measures of Risk Perception to Predict Information Security Behavior: Insights from Electroencephalography (EEG)

1. Introduction

Scholars are increasingly recognizing that individual users play a crucial role in the security of information systems (Furnell & Clarke, 2012; Willison & Warkentin, 2013). This is because users often represent the weakest link in the security of a system—if a user can be coaxed into doing something insecure, the security of an entire system can be compromised (Anderson, 2008). The status of users as the weakest link in the security chain is fully recognized by hackers and cybercriminals, who routinely use social engineering tactics to trick users into installing malicious software (malware) or otherwise obviate technical security controls (Abraham & Chengalur-Smith, 2010; Mandiant, 2013). Given this reality, we need to understand how users perceive and respond to information security risks.

Behavioral research on information systems security to date has ma inly used self-reported measures to gauge users' perceptions of information security risks (e.g., Anderson & Agarwal, 2010; Guo, Yuan, Archer, & Connelly, 2011; Johnston & Warkentin, 2010; Malhotra, Kim, & Agarwal, 2004). While studies using self-reported measures have contributed significantly to our theoretical understanding of security and behavior, such measures are prone to certain biases that can undermine the validity of scientific findings (Dimoka et al., 2012). In particular, many emotions, such as fear, uncertainty, and distrust (all intrinsic to risk perceptions), are at least partially experienced unconsciously, which makes them difficult to capture accurately (Dimoka, Pavlou, & Davis, 2011; Winkielman & Berridge, 2004).

Moreover, users' perceptions of risks have predominantly been associated with intentions to behave rather than behavior itself (Crossler et al., 2013). This is problematic in the context of information security because respondents have been shown to profess security concerns and later fail to take action to protect themselves online, even when the costs to do so are minimal (Acquisti & Grossklags, 2004). Due to these concerns, researchers have called for the measurement of security-related cognition and behaviors using alternative means, such as NeuroIS methods (Crossler et al., 2013). Therefore, a gap exists in our understanding of:

1) how to measure information security risk perceptions most accurately, and

2) how these measures map to security behavior.

This paper contributes to IS research by demonstrating that risk perceptions are effectively measured using electroencephalography (EEG) via event-related potentials (ERPs), which measure neural events triggered by specific stimuli or actions. More specifically, we measured ERPs in response to gains and losses in the Iowa Gambling Task (IGT), a widely used technique in the fields of Psychology and Neuroscience that has been shown to be strongly correlated with real-world risky behaviors (Buelow & Suhr, 2009). Further, we show that the differences in these measures of the neural responses to gains and losses strongly predict users' information security behavior in a separate laboratory-based computing task.

In addition, we compare the predictive power of EEG measures to that of self-reported measures of information security risk perceptions. Our experiments show that these self-reported measures of risk perception are ineffective in predicting security behaviors under a condition in which information security is not salient. However, we show that, when security concerns become salient (through a simulated malware incident on participants' personal computers), these same self-reported measures of security risk perception do predict security behavior. Interestingly, EEG measures significantly predict behavior in both salient and non-salient conditions, which indicates that EEG measures are a robust predictor of security behavior.

This paper proceeds as follows. In Section 2, we review the literature to show how perceived information security risk has been previously captured using self-reported measures and neural measures. In Section 3, we develop our hypotheses for the predictive validity of self-reported and EEG measures of risk and under what conditions they are most effective. In Section 4, we describe our methodology involving a series of surveys and experimental tasks. In Section 5, we present our

analysis and the results of our hypothesis testing. Finally, in Section 6, we discuss the implications of our findings, their limitations, and directions for future research on the use of NeuroIS methods to measure the construct of information security risk perceptions.

2. Literature Review

2.1. Background

In this section, we set a foundation for our hypotheses and experimental task by reviewing how information security risk perceptions have previously been studied in the IS field. We also discuss methodological issues for capturing risk perceptions using self-reported and NeuroIS methods.

2.2. Information Security Risk Perceptions

Risk perception is an interesting area of study because it is a complex combination of social, cultural, economic, psychological, financial, and political factors (e.g., Brooker, 1984; Dholakia, 2001; Grewal, Gotlieb, & Marmorstein, 1994; Kaplan, Szybillo, & Jacoby, 1974; Slovic, 1987). IS researchers have examined risk perceptions in the domains of information security and privacy (e.g., Anderson & Agarwal, 2010; Guo et al., 2011; Johnston & Warkentin, 2010; Malhotra et al., 2004). A primary theoretical perspective used is protection motivation theory (PMT) and related health-belief models (Rogers, 1975). PMT explains how people become motivated to cope with a threat, with two principal drivers being perceived severity and perceived susceptibility. Perceived susceptibility refers to the likelihood of becoming exposed to a threat, whereas perceived severity is the impact of potential consequences posed by the threat (Prentice-Dunn & Rogers, 1986). Together, these two constructs essentially measure perceived risk.

Researchers have used PMT to explain the adoption of anti-spyware software (Johnston & Warkentin, 2010), information security policy (ISP) compliance (Herath & Rao, 2009; Vance, Siponen, & Pahnila, 2012), and security behaviors of employees (Workman, Bommer, & Straub, 2008) and home users (Anderson & Agarwal, 2010). Liang and Xue (2010) used the technology threat avoidance theory (TTAT), which draws on PMT as its theoretical base, to explain how threat severity and susceptibility contribute to the avoidance of spyware threats.

Privacy researchers have also used the construct of perceived risk (Hong & Thong, 2013; Jarvenpaa, Tractinsky, & Saarinen, 1999; Xu, Luo, Carroll, & Rosson, 2011). These studies use perceived risk to explain Internet users' willingness to share information about themselves online. Although these studies do not measure threat severity and threat susceptibility separately, they are both implicit in the measurement items (see Dinev & Hart, 2006; Malhotra et al., 2004 for commonly used measures of privacy perceived risk). In such cases, perceived risk is measured in terms of the likelihood of a negative privacy outcome (such as a company selling one's personal information).

2.3. Self-reported Measures

All of the studies mentioned above measured perceptions of risk with self-reported measures. The advantages of such measures are that they are fairly easy to develop, distribute, collect, and analyze. A straightforward means of measuring someone's perceptions is simply to ask that person. However, self-reported measures are subject to a range of well-known biases and demand effects (Dimoka et al., 2011), including the social desirability bias, subjectivity bias, common methods bias, and demand bias. "Social desirability bias" is the tendency of individuals to portray themselves and their behavior in ways that are more socially acceptable. It includes exaggerated positive self-reports and diminished or non-disclosure of negative self-reports (Paulhus, 1991). "Subjectivity bias" refers to the difficulty of capturing reality by soliciting individuals' subjective perceptions. Individual differences between respondents can distort measures of objective reality (Theorell & Hasselhorn, 2005). "Common methods bias" describes variance that is attributable to artifacts of the survey instrumentation rather than to actual variance between different constructs (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). "Demand bias" relates to the effects of the roles that participants may perceive as part of the tacit social contract formed between participants and the experimenters in

undertaking a study. Demand-induced behaviors include attempting to discern and intentionally confirm or disconfirm the experimenter's hypotheses (Orne & Whitehouse, 2000).

Additionally, self-reported measures require conscious decision making. Several studies in Psychology (e.g., Greenwald & Banaji, 1995; Skowronski & Lawrence, 2001; Spangler, 1992) have shown that self-reported measures may correlate poorly with more implicit measures where participants may or may not be unaware in control of the impact of their attitude and cognition (Fazio & Olson, 2003). These types of non-conscious reactions are inherently impossible to self-report.

Further, in almost all of the previous studies that measured perceived risk, the authors measured intentions, rather than actual behavior, as the dependent variable (Workman et al., 2008 is an exception). This is problematic because studies of IS security and privacy have shown that people behave inconsistently with their self-reported concern for their privacy and security (Acquisti & Grossklags, 2004; Belanger, Hiller, & Smith, 2002; Norberg, Horne, & Horne, 2007). For these reasons, our understanding of perceived risks and their effect on security behavior may be incomplete.

2.4. Neurophysiological Measures

A promising approach to investigate the effectiveness of security warnings is Cognitive Neuroscience applied to Information Systems (NeuroIS). In particular, the neural bases for human cognitive processes can offer new insights into the complex interaction between information processing and decision making. D'Arcy and Herath (2011, p. 694) present a call for research that applies NeuroIS to human-computer interaction, which includes "inferring temporal ordering among brain areas" as an important area of inquiry. They explain that, to understand the design of IT artifacts better, it is desirable to study the timing of brain activations while completing decision tasks. It has been suggested that NeuroIS techniques are a particularly promising means of measuring information security-related behaviors and attitudes (Crossler et al., 2013).

2.5. EEG Measures

The neurophysiological measure we used in this study is the P300 component of an event-related potential (ERP) measured with electroencephalography (EEG). The P300 is a positive-going component that peaks between 250 and 500 milliseconds after stimulus onset and has been observed in tasks that require stimulus discrimination (Polich, 2007). Passive stimulus processing generally produces smaller P300 amplitudes than active tasks; when task conditions are undemanding, the P300 amplitude is smaller. It has been proposed that the P300 reflects processes related to updating the mental representations of the task structure (Donchin, 1981; Donchin & Coles, 1988). According to the "context-updating theory", incoming stimuli are compared against previous stimuli held in working memory. If the new stimulus matches previous stimuli, no updating is required and no P300 is generated. If, however, the new stimulus produces a mismatch with the stimuli held in working memory, the context for that stimulus is updated and a P300 is generated. It is believed that because infrequent, low-probability stimuli can be biologically important, it is adaptive to inhibit unrelated activity to promote processing efficiency, and thereby yield large P300 amplitudes (Polich, 2007).

EEG has been used in the Psychology and Neuroscience literatures to study risk-taking tendencies in individuals. Many EEG studies on risk-taking have had participants perform a gambling task while measuring the EEG either before or during the task. Some studies have related self-reported survey responses on risk-taking propensities to different EEG measurements, including ERP amplitudes and latencies, the power in different frequency bands of the EEG (Massar, Rossi, Schutter, & Kenemans, 2012), or resting-state EEG measurements (Massar, Kenemans, & Schutter, 2013). A group of studies have focused on ERPs for predictions of risk-taking behaviors during gambling or other card tasks. These studies looked at various ERP components under conditions when participants experience a negative (loss) event or a positive (reward) event.

3. Theory and Hypotheses

In this section, we lay out our theory and associated hypotheses. Before doing so, however, we define the specific type of security behavior we examine in this paper—security warning disregard.

3.1. Security Warning Disregard

A common defense against security threats is a security warning presented to the user by the operating system and software, such as email clients and Web browsers (Yee, 2004). While technically effective, these security warnings are undermined by users who either willfully disregard them or fail to recognize the importance of the threat (Schneier, 2004). This commonly observed behavior is contained in a well-known epigram in information security: “given a choice between dancing pigs and security, users will pick dancing pigs every time” (McGraw & Felten, 1999, p. 29). This means that, given a Web link promising to show some amusing entertainment on one hand and a security warning for the same link on the other, users will often ignore the security warning and access the Web link anyway.

We formally define this behavior as “security warning disregard”, which is behaving against the recommended course of action of a security warning. We chose this particular form of security behavior as the dependent variable for the study because, by heeding or disregarding a security warning, users explicitly accept or reject taking on added risk to their information security. Consequently, this behavior provides an observable indication of a user’s information security risk tolerance, which can then be directly compared to measures of perceived risk.

3.2. The Effect of Security Incidents on Behavior and Risk Perception

For fear-arousing stimuli intended to cause individuals to perceive a threat and take a certain action to avert that threat to be effective, both the threat severity and the threat susceptibility should be conveyed (Rogers, 1975). The concept of a threat in a fear-arousing stimulus is theoretically analogous to the concept of risk—a threat is an event with a potentially adverse consequence (Witte, 1992), and a risk describes an event with a potentially negative consequence.

However, computer users are susceptible to becoming desensitized to fear-arousing stimuli and warning messages in general. Studies on warning communications and information system security have empirically validated that users can become habituated to seeing warning messages. For example, Egelman, Cranor, and Hong (2008) demonstrate habituation to Web browser phishing warning messages that appear visually similar after repeated exposure to the warning. Similarly, Amer and Maris (2007) used a laboratory experiment with more general or generic system warning messages that also demonstrated warning message habituation. We theorize that technology users who are already habituated to security warning messages will perceive low risk when presented with the messages since habituation can cause the warning to not even rise to an individual’s attention level (Amer & Maris, 2007) and because it is likely that the warning never resulted in a negative consequence (also described as the warning "crying wolf" by Sunshine, Egelman, Almuhimedi, Atri, and Cranor (2009)).

We further predict that, after suffering a security incident (a fear-arousing stimulus) in relation to a threat, the security warning message warned against, such as having one's computer become infected with computer malware, an individual’s perceived threat susceptibility will become higher than before the security incident, which will increase self-reported measures of risk perception. Good, Dhamija, Muligan, and Konstan (2005) found that users who had a recent past negative experience in their computer usage were more cautious compared with other users. We also predict, consistent with the findings of Johnston and Warkentin (2010), a decrease in intended risk-taking behavior. In line with this logic, we hypothesize that:

H1: *Security warning disregard before a security incident will be higher than security warning disregard after a security incident screen.*

H2: *Pre-test self-reported measures of risk perception will be lower before a security incident than post-test self-reported measures of risk perception after a security incident.*

3.3. Effectiveness of Self-reported Measures

The theory of planned behavior (TPB) (Ajzen, 1991) posits that an individual's beliefs and intentions correlate with their actual behavior associated with those beliefs and intentions. Since raising perceptions of risk or a threat is an integral goal of a fear-arousing stimulus and since a security warning message can be considered fear-arousing, we posit that self-reported measures of risk perception will predict security warning disregard.

However, the pre-test self-reports will likely be confounded by the effects of habituation to seeing security warning screens regularly (Amer & Maris, 2007). Consequently, while some users may report high levels of perceived risk in general, they may act contrarily in their actual behavior. Thus, pre-test risk perception measurements would not be as strong a predictor for actual security warning disregard before a security incident compared with parallel post-test measurements after a security incident since the exposure to the security incident will likely break the habituation and desensitization to the security warning messages (see Bansal, Zahedi, & Gefen, 2010; Good et al., 2005; Ng & Feng, 2006). Therefore, we hypothesize that:

H3: *Pre-test self-reported measures of risk perception will negatively predict security warning disregard before a security incident is imposed.*

H4: *Post-test self-reported measures of risk perception will negatively predict security warning disregard after a security incident is received better than will pre-test self-reported measures of risk perception before a security incident is imposed.*

3.4. Effectiveness of EEG Measures

Prior studies in the Neuroscience literature have reported that amplitude measurements of the P300 component of the ERP during risk-taking laboratory experiment tasks can correlate with participants' risk-taking behavior during the experiment (Polezzi, Sartori, Rumiati, Vidotto, & Daum, 2010; Schuermann, Endrass, & Kathmann, 2012; Yeung & Sanfey, 2004). The P300 amplitude can vary depending on such factors as the valence of an outcome (i.e., whether it is a gain or loss), how frequent target stimuli are compared to non-targets, the magnitude of the gain or loss, and the personal motivation to do well (Yeung & Sanfey, 2004).

The P300 has been implicated in context updating (cf. Donchin, 1981; Donchin & Coles, 1988). According to the context-updating theory, the P300 reflects the amount of cognitive resources allocated to re-evaluating an internal model of the environment. In a task such as the Iowa Gambling Task (IGT), the internal model has to do with the probability of a reward when selecting cards from certain decks. In the case of the low-frequency, high-magnitude penalty (referred to as the B Penalty; see the task description below), a large updating is necessary because the deck has been a big winner for a number of trials but now becomes a big loser. This interpretation is supported by recent findings by San Martín, Appelbaum, Pearson, Huettel, and Woldorff (2013), who showed that the magnitude of the P300 predicted individual choices in a gain maximization/loss minimization task similar to the IGT.

We posit that the P300 will predict security warning disregard better after a security incident than it will before a security incident. While P300 amplitude measurements can still measure how individuals will respond in risk-taking situations, if the users are habituated to a particular security warning screen, they are less likely to register it as a threat until they are sensitized by a significant change, such as a security incident (Amer & Maris, 2007). After experiencing the security incident, individuals' P300 measurements will better predict their "unhabituated" security warning disregard. Given these arguments, we hypothesize that:

H5: *Pre-test P300 amplitude measures will negatively predict security warning disregard before a security incident is imposed.*

H6: *Pre-test P300 amplitude measures will negatively predict security warning disregard after a security incident is imposed better than will the same P300 amplitude measures before a security incident is imposed.*

3.5. Advantages of Neurological Measures vis-à-vis Self-reported Measures for Risk Perceptions

While self-reported measurements have been theorized to be able to predict actual behavior to a certain degree (see TPB, Ajzen, 1991), these measurements are subject to several weaknesses that decrease their explanatory power, including common methods bias, social desirability, and subjectivity bias (Dimoka et al., 2011). Moreover, constructs such as risk perceptions can be difficult to capture with self-reported measures given that they can be non-conscious, and individuals naturally cannot self-report things about themselves of which they are not cognizant (Dimoka, 2010; LeDoux, 2003). Direct neural measurement methodologies, such as EEG, overcome these weaknesses of self-reported measurements since they measure without the participants' involvement (Dimoka et al., 2011). Therefore, we predict that EEG measurements (specifically the P300 component of the ERP) will predict security warning disregard better than both pre-test and post-test self-reported measurements. Thus, we hypothesize that:

H7: *Pre-test P300 amplitude measures will negatively predict security warning disregard better than pre-test self-reported measures of risk perception* before *a security incident is imposed.*

H8: *Pre-test P300 amplitude measures will negatively predict security warning disregard better than will post-test self-reported measures of risk perception* after *a security incident is imposed.*

4. Methodology

We used a laboratory experiment to test the hypotheses. The experimental design consisted of four stages: a pre-test survey; a risk-taking experiment called the Iowa Gambling Task, during which EEG was recorded; a separate image classification computing task with simulated security warnings; and a post-test survey. We explain each of these stages below.

4.1. Pre-test Survey

Prior to taking part in the experiment, we had the participants complete a pre-test survey to gauge their general risk propensity and information security risk perceptions of malware. To measure general risk propensity, we used general risk orientation (Kam & Simas, 2010) and willingness to gamble lifetime income (Barsky, Juster, Kimball, & Shapiro, 1997) for income risk. These questions enabled us to create a general risk profile for each subject. We also used two different measures of IS security risk perception to ensure that (1) our measures were representative of IS risk perception measures used in the literature and (2) our results would not depend on any one measure. The measures we selected were those of Johnston and Warkentin (2010), who measured risk perceptions using separate items for "threat severity" and "threat susceptibility", and those of Guo et al. (2011), who measured security risk perception as a single construct. Please see Appendix B for information about how we selected these measures.

The pre-test survey contained 16 items measuring general and information security-related risk. To minimize hypothesis guessing in the experimental task, we had the participants take the pre-test survey online one week prior to the experiment. Because the survey was online, it could be taken by many students at once. However, we scheduled the other phases of the experiment at one hour each, with limitations based on laboratory and researcher availability. Consequently, some students who took the initial survey very early were not able to do the second stage of the experiment until up to four weeks later. To further obscure the objective of the pre-test survey, we added 17 unrelated personality questions to the survey and eight demographic questions. In total, the pre-test survey consisted of 44 questions, which are reported in the appendix.

4.2. Iowa Gambling Task

The first experiment consisted of the Iowa Gambling Task (IGT), a widely used technique in the Psychology and Neuroscience fields to measure individuals' decision making ability (Toplak, Sorge, Benoit, West, & Stanovich, 2010). The IGT was originally designed by Bechara, Damasio, Damasio, and Anderson (1994) at the University of Iowa as an instrument to measure risk-taking behaviors by simulating real-life decision making. The task is a gambling card game in which participants are required to choose cards from four decks for a set number of rounds. Each card draw results in participants earning a varying amount of play money, but some cards also include penalties that lose money. Certain decks are safer in that they contain smaller rewards, but the losses are also smaller, which results in overall net gains. In contrast, riskier decks contain larger rewards, but the losses are also larger, which results in overall net losses. In both the safe and risky decks, the frequency of losses is varied such that some decks have frequent, smaller losses while other decks have infrequent, larger losses (see Table 1 below). The participants' task is to learn by experience which decks are safest; that is, those that yield the most money in the long run. Participants are said to be risk seeking if, after all rounds have been completed, they have lost more money than they have earned (Weller, Levin, & Bechara, 2010).

Of particular interest to our study, the IGT has been shown to be predictive of risk behaviors outside of the experimental task (Schonberg, Fox, & Poldrack, 2011). For example, poor performance on the IGT is strongly correlated with real-world risky behaviors such as substance abuse, compulsive gambling, criminality (Buelow & Suhr, 2009), and medication non-compliance (Stewart, Acevedo, & Ownby, 2012). Additionally, IGT performance has been found to be strongly associated with sensation-seeking (Crone, Vendel, & van der Molen, 2003), disinhibition (van Honk, Hermans, Putman, Montagne, & Schutter, 2002), reward responsiveness and fun-seeking (Suhr & Tsanadis, 2007), and impulsivity (Buelow & Suhr, 2013). Given the predictive power of the IGT for real-world risky tasks, we similarly expect the IGT to be predictive of insecure computing behaviors.

In the Neuroscience field, various indexes have been used as measurements for risk-taking behavior during the IGT. Besides a simple ratio of risky to non-risky deck choices (Bechara et al., 1994), neurophysiological methods have also been used, such as skin conductance responses (SCR) (Bechara, Damasio, Tranel, & Damasio, 2005; Maia & McClelland, 2004; van Honk et al., 2002), cortisol measurements to correlate a lack of fear with higher risk-taking propensities (van Honk, Schutter, Hermans, & Putman, 2003), positron emission technology (PET), which measures normalized cerebral blood flow (rCBF) (Bolla et al., 2003; Ernst et al., 2002), and fMRI (Fukui, Murai, Fukuyama, Hayashi, & Hanakawa, 2005; Singh & Sungkarat, 2008; Tanabe et al., 2007), which also uses blood flow measures to track neural activity. In particular, EEG is a popular method of measuring the neural correlates of risk-taking behavior in the IGT (Oberg, Christie, & Tata, 2011; Schutter & Van Honk, 2005). Recent research has demonstrated that the P300 is sensitive to loss minimization, with larger amplitudes for larger than for smaller losses (San Martín et al., 2013). Furthermore, San Martín et al. demonstrated that the P300 amplitude predicted subsequent behavioral adjustment in individual subjects. Likewise, we measured P300 amplitudes during the IGT in this study.

4.3. IGT Procedures

The stimuli consisted of the four virtual decks of cards displayed on a computer monitor in an electrically shielded testing room. Participants entered each deck selection using the keyboard. On making a choice, the participant received feedback such as "You won 50" or "You won 100 but lost 50" after a 750 ms delay. This delay was used to separate the electrical activity of the motor act of pressing the keyboard button from the ERP response to the feedback message. We implemented the same reward/penalty schedule as the original IGT described by Bechara et al. (1994) over 100 trials (see Table 1). However, we modified the original IGT design to include four rounds of 100 trials per round (400 trials total). We did this to make the task more suitable for ERP measurement (Christie & Tata, 2009). The position of the four decks was randomized at the start of each round to require participants to rediscover which decks were most profitable. Win and loss subtotals were displayed in between each round. The decks were shuffled at the start of each round. Finally, we instructed participants that they would be eligible to receive a bonus extra credit point if they finished the IGT with a positive balance.

Table 1. IGT Deck Details

Deck	Gains	Losses	Frequency	Net gain/loss over 10 trials	Rank of riskiness
A	Large (+100)	-150, -200, -250, -300, -350	Frequent	(-250)	2
B	Large (+100)	-1,250	Infrequent	(-250)	1
C	Small (+50)	-25, -50, -75	Frequent	+250	4
D	Small (+50)	-250	Infrequent	+250	3

4.4. Electrophysiological Data Recording and Processing

The electroencephalogram was recorded from 128 scalp sites using a HydroCel Geodesic Sensor Net and an Electrical Geodesics Inc. (EGI; Eugene, Oregon, USA) amplification system (amplification 20K, nominal bandpass 0.10–100Hz). We referenced the EEG to the vertex electrode and digitized at 250 Hz. Impedances were maintained below 50 kΩ. EEG data were processed off-line beginning with a 0.1 Hz first-order highpass filter and a 30 Hz lowpass filter. Stimulus-locked ERP averages were derived spanning 200 ms pre-stimulus to 1,000 ms post-stimulus and segmented based on the following trial type criteria: risky deck (decks A and B) rewards, safe deck (decks C and D) rewards, deck A penalty, deck B penalty, deck C penalty, and deck D penalty. We removed eye blinks from the segmented waveforms using independent components analysis (ICA) in the ERP principal components analysis (PCA) toolkit (Dien, 2010) for Matlab (Mathworks, Natick, MA). We removed the ICA components that correlated at 0.9 with the scalp topography of a blink template from the data (Dien, Michelson, & Franklin, 2010). We removed artifacts in the EEG data due to saccades and motion from the segmented waveforms using PCA in the ERP PCA toolkit (Dien, 2010). We marked channels as bad if the fast average amplitude exceeded 100 μV or if the differential average amplitude exceeded 50 μV. Because the structure of the IGT results in fewer trials in some conditions than in others, there is a possibility that ERP results could be biased by unequal trial counts in the conditions of interest due to the lower signal to noise ratio associated with fewer trials (Clayson, Baldwin, & Larson, 2013). To counter this, we randomly chose a subset of trials from the conditions with more trials to match the number of good trials following artifact correction in the condition with the lowest trial count (the deck B penalty in almost every case). We excluded data from three participants (one female, two males) from our ERP analyses due to low trial counts or excess bad channels. We average re-referenced data from the remaining participants and baseline corrected waveforms using a 200 ms window prior to feedback stimulus presentation. The participants spent about 15 minutes on the IGT task, after which we removed the EEG cap.

4.5. Image Classification Task

After participants completed the IGT, we took them into another testing room to perform an online image classification task. The purpose of this task was to observe how participants responded to security warnings when working under time pressure (simulating real-world working conditions). However, we concealed the purpose for this task from the participants, who were led to believe that classifying online images was the sole purpose of the task. In doing so, we followed a deception protocol approved by the university's institutional review board.

Our goal in this study was to determine the effectiveness of the self-reported risk perception measures and the IGT measure to predict risky security behavior—specifically, security warning disregard. For this reason, it was critical that participants perceived actual risk to their data when performing the task. Accordingly, we required participants to bring their personal laptops to the experiment to use during the image classification task. In a few instances, participants failed to bring a laptop, in which case they were provided with a laptop that belonged to one of the researchers. Debriefing interviews with participants subsequent to the task uniformly confirmed that participants perceived higher risk due to using their (or the researcher's) personal laptop rather than a laboratory machine.

4.6. Image Classification Task Procedure

Participants used their laptops to browse to a URL for the image classification task and signed in using a participant number. They were left alone in the room to complete the task. Participants read instructions stating that their task was to classify images of Batman on the Web as either animated or photographic versions of the character. The ostensible purpose for doing so was to compare a computer algorithm's performance in the classification task to that of a human.

During the task, the experimental website displayed in an HTML frame websites found through a Goo gle Image search for "Batman" (see Figure 1 below).

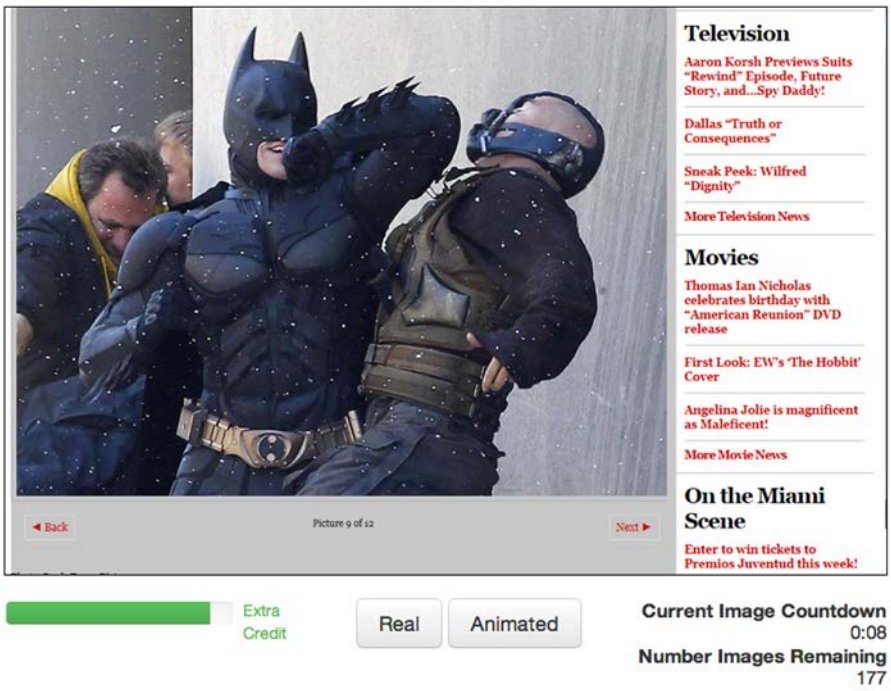


Figure 1. The Image Classification Experimental Website

For each website, participants were required to click a button labeled "real" or "animated" to classify the images. Additionally, participants were under time pressure to complete the task. For each website, participants had ten seconds in which to classify the image. Failure to classify the image was counted as an incorrect answer.

A performance bar in the bottom-left corner of the screen provided participants with live feedback on their performance. Initially, the bar was green, and it remained so as long as participants classified images correctly. A green bar communicated to participants that they were on track to receive a bonus extra credit point given as an incentive. However, if the bar turned red due to a misclassification, participants knew that they were no longer eligible for the bonus extra credit point. This design was chosen because loss aversion research (cf. Kahneman & Tversky, 1984) indicates that people are more passionate about not losing something relative to the possibility of gaining something.

The penalty for failing to classify an image correctly was a 40 percent reduction of the performance bar. However, each correct classification increased the performance bar by 10 percent (if not already full). At the 90-pecent level, the bar became green again, which thus built forgiveness into the task so that participants could recover with good performance.

Unbeknownst to the participants, the experimental website was programed to periodically display Web browser security warnings (see Figure 2). We modeled the security warning after the one used

in Google Chrome. As such, participants could click on the security warning by choosing either “No, don’t proceed” or “Yes, proceed anyway”.

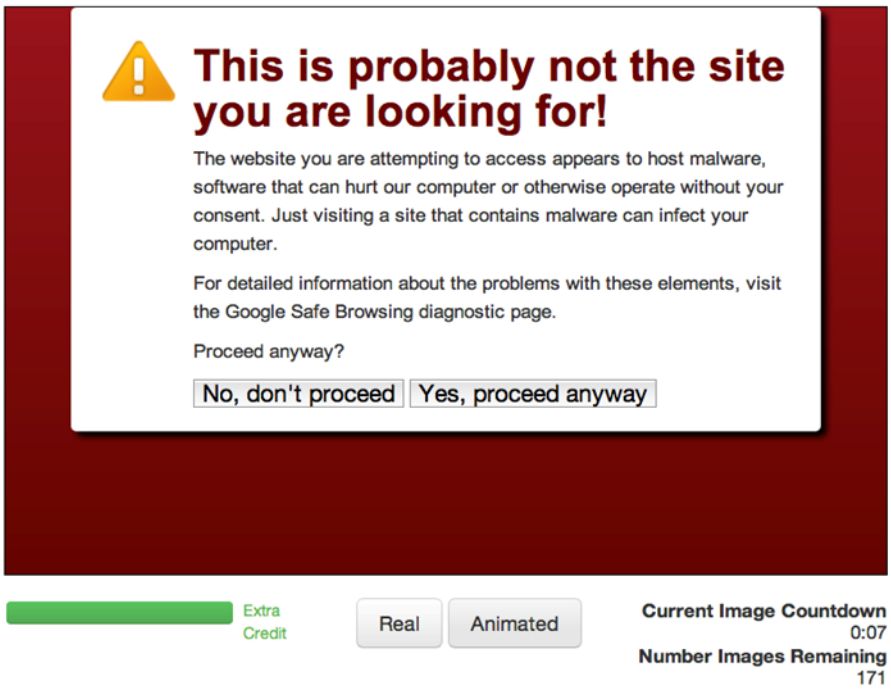


Figure 2. The Security Warning Screen

If the participants clicked “no”, the Web frame redirected to Google.com, and they were penalized for failing to classify the image correctly. If participants clicked “yes”, they were taken to the Google Image search result and allowed to classify the image. If participants failed to take action in ten seconds, the experimental website displayed the next Google Image search result and penalized participants for failing to classify the image. Thus, participants were under pressure to work quickly and perform well on the task. Heeding the security warning therefore came at a cost of productivity, simulating the real-world costs of observing security warnings (Herley, 2009).

The configuration for the algorithm's penalty mechanism was as follows: the performance bar had a range from zero to 100 and an initial value of 100. If the bar dropped below a predetermined 90-point threshold, the bar would turn red. The penalty for a failed classification was 40 points. If the performance bar was not already full, the reward for a correct answer was 10 points.

In addition to examining how participants would respond to security warnings initially, we also wished to observe whether a security incident could raise perceptions of information security risk and change behavior (H2, H4, H6, H8). Therefore, we imposed a simulated security incident without warning midway through the image classification task. The security incident consisted of a message from an “Algerian hacker” that displayed a ten-second countdown timer and the words “Say goodbye to your computer” (Figure 3).

Figure 3. The Simulated Security Incident Hacker Screen

The message was displayed full-screen to maximize impact and was modeled after actual website defacements archived on Zone-H.org. Participants reported a relatively high degree of concern as a result of the hacker screen (an average of 7.5 on a scale of 0—"not concerned at all"—to 10—"100% concerned")—a result that was significantly higher than the neutral response of 5 (one-sample t-test, $t = 3.752$, $p < .001$). Additionally, we corroborated this finding in the debriefing interviews[1].

The frequency at which security warnings were displayed varied between every 10th and 15th website viewed to prevent participants' detection of a fixed pattern. Before the security incident, every participant saw a total of seven security warnings. After the security incident, the warnings appeared at the same variable frequency as before until all 182 experimental websites had been viewed, which equated to approximately six to eight warning screens displayed.

4.7. Post-survey

On completing the image classification task, we asked participants to complete a brief post-survey (see the appendix), which allowed us to compare whether self-reported measures of information security risk increased after the security incident (H2). Accordingly, we included the same measures of information security risk given on the pre-test ("perceived risk", "threat susceptibility", and "threat severity"). We also quizzed participants about how much they recalled from the pre-test to assess whether the pre-test survey influenced their behavior during the image classification task. No participant perfectly recalled the pre-test, although 3.7 percent of respondents correctly identified all of the general topics of the pre-test. Approximately 27 percent reported that the pre-test did influence their behavior, but of these, the average reported influence was moderate (an average of 3.3 on a five-point Likert-type scale). Finally, a t-test showed no difference in security warning disregard between those who claimed that the pre-test influenced their behavior and those who did not (before

[1] At least seven participants shut their laptop lids, powered off their laptops, unplugged the network cable, and/or otherwise stopped the experiment after seeing the hacker screen.

a security incident: t = 1.08; after: t = 1.10; both not significant). Therefore, we conclude that the pre-test had minimal impact on the results of the image classification task.

4.8. Pilot Tests

In preparation for our study, we conducted two pilot tests. The first pilot test (N = 25) consisted of the pre-test survey and the image classification task. The second pilot test (N = 30) also included the IGT. After each pilot test, we made adjustments to the experimental protocol based on participant feedback and analysis of the data. For example, after the second pilot test, we found that electrical activity in the brain from the motor act of pressing the keyboard button masked the signal of the ERP in response to the IGT feedback. As a result, we instituted a 750 ms delay between deck selection and win/loss feedback, which substantially reduced the noise in the data.

4.9. Primary Data Collection

Sixty-two healthy volunteers (16 females, 46 males) at a large private university in the western United States were recruited to participate. The average age was 21.84 (std. 1.96). These participants were part of a research pool that gave extra credit points toward a variety of university courses. Each extra credit point corresponded to .025 percent of the participants' course grades. We told participants (as a part of the sign-up process) that they would receive two extra credit points for completing all four steps of the experiment. Prior to the second step (the IGT), we told participants that they could receive a bonus extra credit point if they finished the IGT with a positive balance and completed the image classification task in the "green zone". However, all students who completed all four steps were given three extra credit points.

Participants reported demographic variables such as age, gender, handedness (right: 55, left: 7), normal or corrected vision (yes: 59, no 3), colorblindness (no: 58, yes: 4), whether they were a native English speaker (yes: 55, no: 7), and whether they had been treated for a neurological or psychiatric condition (yes: 3, no: 59). These variables are known to affect neural processing (Luck, 2005) and were later used as controls in our analysis.

5. Analysis

We chose linear regression to test our hypotheses because it is a common form of analysis for both EEG and field survey studies. Thus, regression provided a shared method to assess both the self-reported and EEG-related hypotheses. Second, our models were simple: they had one or two independent variables depending on the information security risk measure used. Therefore, a multivariate technique such as structural equation modeling was unnecessary.

The dependent variable in our analysis was participants' security warning disregard during the image classification task. We operationalized this variable as the ratio of the number of times participants actively chose to ignore the security warning (by clicking "Yes, proceed anyway" on the security warning) over the total number of security screens displayed. If a participant either clicked the "No, don't proceed" button or took no action before the timeout period, then we recorded that the security warning was not disregarded. We calculated this ratio both before and after the participant received the security incident.

5.1. Control Variable Analysis

We examined whether the control variables influenced security warning disregard using stepwise regression. We found that, for "security warning disregard (before-incident)", the control variables had no significant influence. However, for "security warning disregard (after-incident)", whether or not participants were native English speakers (seven non-native speakers, 55 native) had a significant effect (-.281 standardized beta, t = -2.272 one-tailed, $p < .05$). Accordingly, we included this variable in our regressions involving "security warning disregard (after-incident)".

5.2. Validation of Self-reported Measures

We validated our self-reported measures as follows. First, we tested the reliabilities of our risk measures (general risk orientation, perceived security risk, perceived threat susceptibility, and

perceived threat severity) for both pre- and post-test measures. All items exhibited a Cronbach's alpha greater than .70, which indicates good reliability (Nunnally, 1970). We then summed the items for each construct to create a single independent variable to be used in the regression analysis. Second, we performed an exploratory factor analysis (EFA) for "perceived threat susceptibility" and "perceived threat severity" to ensure that these constructs functioned as distinct independent variables in the same model (Straub, Boudreau, & Gefen, 2004)[2]. The EFA showed a clear pattern of loading onto two factors, with all items loading onto the appropriate factor, consistent with Johnston and Warkentin (2010). Therefore, we conclude that instrument validation was sufficient to support statistical testing of our hypotheses.

5.3. Iowa Gambling Task Behavioral Performance

To assess performance in the IGT, we calculated the ratio of choices from the "risky" decks (decks A and B) to choices from the "safe" decks (decks C and D) for each block of 100 trials. Consistent with previous studies employing the IGT (e.g., Bechara et al., 1994), participants switched from choosing more from the risky decks to choosing from the safe decks (mean ratio of 1.06, .83, .69, and .71 for blocks 1, 2, 3, and 4, respectively). A repeated-measures ANOVA on the risky- to safe-choice ratio revealed a main effect of block ($F = 12.68$, $p < 0.001$) and a significant linear trend across blocks ($F = 21.42$, $p < 0.001$).

5.4. Hypothesis Testing

5.4.1. Testing the Impact of the Security Incident in the Experimental Task

First, we tested whether participants disregarded the security warning screens less frequently after the hacker screen was received (hereafter referred to as the security incident). A paired-sample t-test showed that, on average, participants disregarded the security warning screen significantly less after a security incident (ratio of .66 of warnings disregarded over warnings received) than before (.73), which indicates a significant decrease in security warning disregard ($t = 2.192$ one-tailed, $p < 0.05$). We confirmed the impact of the security warning screen in the post-test survey, in which participants reported that the security warning screen was both realistic and concerning (6.76 and 8.47 respectively, measured on a 0 to 10 scale). This supports H1 because participants changed their security warning disregard after a security incident (see Table 2).

Table 2. Paired-sample T-test Comparing Security Warning Disregard Before and After the Security Incident (H1)

Mean of SWD (before Incident)	Mean of SWD (after Incident)	Mean of difference	Std. deviation of difference	Std. error mean of difference	95% confidence interval		*t*
					Lower	Upper	
.733	.656	.077	.274	.035	.007	.146	2.192*

* $p < .05$; degrees of freedom = 61; SWD = security warning disregard.

Next, we tested whether perceptions of information security risk increased after a security incident. We would expect that the impact of the security incident would make information security risks more salient for participants, which would lead to a higher perception of information security risks. Again, a paired-sample t-test analysis showed that "threat severity" and "threat susceptibility" perceptions increased significantly by approximately 15 percent after participants had a security incident (an increase of 3.31 and 3.11, respectively, in a range of 21; $t = 6.104$ one-tailed, $p < .001$). Likewise, "perceived risk" also increased in the post-test, but not significantly (an increase of .57 out of a range of 21). As such, while our findings did not support H2a for "perceived risk", they did support H2b for "threat severity" and H2c for" threat susceptibility" (see Table 3).

[2] We did not perform an EFA was for perceived security risk, general risk perceptions, or risk income preferences because we ran these constructs in separate models as the only independent variable or factor.

Table 3. Paired-sample T-test Comparing Perceptions of Information Security Risk Before and After the Security Incident

H2a: for perceived security risk of malware (PSRM)

Mean of PSRM (before incident)	Mean of PSRM (after incident)	Mean of difference	Std. deviation of difference	Std. error mean of difference	95% confidence interval		*t*
					Lower	Upper	
14.19	14.76	.565	3.911	.497	-.429	1.558	1.558 ns

H2b: for threat severity of malware (TSEV)

Mean of TSEV (before incident)	Mean of TSEV (after incident)	Mean of Difference	Std. deviation of difference	Std. error mean of difference	95% confidence interval		*t*
					Lower	Upper	
11.51	14.82	3.311	4.237	.542	2.226	4.397	6.104*

H2c: for threat susceptibility of malware (TSUS)

Mean of TSUS (before incident)	Mean of TSUS (after incident)	Mean of difference	Std. deviation of difference	Std. error mean of difference	95% confidence interval		*t*
					Lower	Upper	
12.42	9.31	3.113	3.725	.473	2.167	4.059	6.581*

*** p < .001; ns = not significant; degrees of freedom = 61.

5.4.2. Testing the Predictive Validity of Self-reported Measures of IS Risk Perception

To test H3, we examined whether self-reported measures of information security risk perceptions predicted security warning disregard before a security incident. To do this, we ran three separate regression equations with "security warning disregard (before incident)" as the dependent variable and one of the three self-reported information security risk perception measures ("threat severity", "threat susceptibility", and "perceived risk") as a single independent variable (see Table 4). This allowed us to examine the effect of each measure independently[3]. As an additional test, we also examined whether general risk orientation (i.e., general risk appetite and willingness to gamble lifetime income) also predicted behavior. However, none of the above regression tests were significant. Therefore, H3 was not supported—none of the self-reported security risk measures (H3a for "perceived risk", H3b for" threat severity", H3c for "threat susceptibility", H3d for "general risk appetite", and H3e for "willingness to gamble lifetime income") predicted security warning disregard before the security incident.

[3] We also tested threat severity and threat susceptibility together since these measures are designed to predict jointly (Johnston & Warkentin, 2010). However, in this model, both factors remained insignificant.

Table 4. Regression Results for the Effects of Pre-test Risk Perception on Security Warning Disregard (Before Security Incident)

H3a: Perceived security risk of malware—pre-test				
Model	**ß**	**Std. error**	**Standardized ß**	***t***
Intercept	.851	.189	—	4.493***
Perceived security risk of malware—pre-test	-.008	.013	-.083	-.647 ns
Model statistics: R^2 = .007; f = .419, p = .520				
H3b: Threat severity of malware—pre-test				
Model	**ß**	**Std. error**	**Standardized ß**	***t***
Intercept	.673	.127	—	5.306***
Threat severity of malware—pre-test	.006	.010	.080	.614 ns
Model statistics: R^2 = .006; f = .377, p = .542				
H3c: Threat susceptibility of malware—pre-test				
Model	**ß**	**Std. error**	**Standardized ß**	***t***
Intercept	.760	.127	—	6.004***
Threat susceptibility of malware—pre-test	-.003	.012	-.031	-.236 ns
Model statistics: R^2 = .001; f = .056, p = .814				
H3d: General risk appetite—pre-test				
Model	**ß**	**Std. error**	**Standardized ß**	***t***
Intercept	1.015	.254	—	4.002***
General risk appetite—pre-test	-.010	.009	-.140	-1.088 ns
Model statistics: R^2 = .020; f = 1.184, p = .281				
H3e: Willingness to gamble lifetime income—pre-test				
Model	**ß**	**Std. error**	**Standardized ß**	***t***
Intercept	.793	.087	—	9.106***
Willingness to gamble lifetime income	-.037	.043	-.110	-.854 ns
Model statistics: R^2 = .012; f = .729, p = .397				

*** p < .001; ns = not significant; one-tailed tests.

We also hypothesized that self-reported measures of risk would be more effective at predicting security warning disregard after a security incident compared to before a security incident (H4). To test this hypothesis, we followed the same procedure as described above for testing H3, with the difference that we tested "security warning disregard after incident" as the dependent variable and used post-test measurements for "threat severity", "threat susceptibility", "perceived risk", and both "threat severity" and "threat susceptibility" in the same model. Our results showed that the post-test measurement of "perceived risk" (-.252 standardized beta, t = -2.069 one-tailed, p < .05) and "threat susceptibility" significantly (-.294 standardized beta, t = -2.448 one-tailed, p < .05) reduced security warning disregard (see Table 5). Thus, H4a for "perceived risk" and H4c for "threat susceptibility" were supported[4]. However, H4b for "threat severity" was not. As such, the results support our hypothesis that self-reported measures would be more predictive immediately after subjects experienced a salient security incident only for "threat susceptibility".

[4] We also tested whether pre-test measures for self-reported risk predicted post-consequence security warning disregard. In this case, perceived risk had a significant negative effect (-.22 standardized beta, p < .05), but threat susceptibility had no effect.

Table 5. Regression Results for the Effects of Post-test Risk Perception on Security Warning Disregard (After Security Incident)

H4a: Perceived security risk of malware—post-test				
Model	**ß**	**Std. error**	**Standardized ß**	***t***
Intercept	1.436	.262	—	5.475***
Native English speaker	-.435	.168	-.316	-2.595**
Perceived security risk of malware—Post-test	-.028	.013	-.252	-2.069*
Model statistics: R^2 = .142; f = 4.863 , p = .011				
H4b: Threat severity of malware—post-test				
Model	**ß**	**Std. error**	**Standardized ß**	***t***
Intercept	1.092	.268	—	4.077***
Native English speaker	-.401	.175	-.291	-2.298*
Threat severity of malware—post-test	-.005	.013	-.055	-.432 ns
Model statistics: R^2 = .082; f = 2.640, p = .080				
H4c: Threat susceptibility of malware—post-test				
Model	**ß**	**Std. error**	**Standardized ß**	***t***
Intercept	1.423	.232	—	6.142***
Native English speaker	-.446	.166	-.324	-2.692**
Threat susceptibility of malware—post-test	-.030	.012	-.294	-2.448*
Model statistics: R^2 = .164; f = 5.793, p = .005				

*** p < .001; ** p < .01; * p < .05; ns = not significant; one-tailed tests.

5.4.3. Testing of the Predictive Validity of EEG Measures of Risk Perception

Next, we tested the predictive validity of our EEG measures of risk perception. To do this, we used the P300 ERP component as the independent variable. We extracted the P300 amplitudes as the mean amplitude in the 300–600 ms post-stimulus window (Fjell & Walhovd, 2001). We calculated latencies as the 50 percent area latency (Bashore & Ridderinkhof, 2002; Polich & Corey-Bloom, 2005) for the 300–600 ms post-stimulus window.

We calculated each participant's P300 responinformationses to gain/loss feedback subsequent to deck selections in the IGT. The deck selections of special interest for our context were the highest-risk deck (deck B with high-penalty and low-frequency) and the lowest-risk deck (deck C with low-penalty and high-frequency). These decks provided the greatest contrast to participants' responses to penalties incurred in the IGT. In the course of the experiment, participants chose from deck B an average of 94.7 times (SD=29.9, min=53, max=160). Penalties in this deck are high-magnitude and low-frequency; participants received an average of 6.5 B penalty trials (SD=3.8, min=1, max = 16). Participants chose from deck C an average of 104.9 times (SD=29.1, min=38, max=160). Penalties in this deck are of lower magnitude and higher frequency; participants received an average of 34.5 C penalty trials (SD=17.5, min=1, max=73; see Appendix E for more information). As we note in Section 4.4, we randomly selected trials from conditions with higher trials counts to reduce bias due to differential trial counts.

Next, to highlight the contrast between risky and safe feedback, we calculated a difference curve or score by subtracting the activity observed for the C penalty response from that of the B penalty response (Falkenstein, Hoormann, Christ, & Hohnsbein, 2000) (see Figure 4). This technique suppresses activity that is common between two experimental conditions and thus leaves only the difference of what is truly distinct (Hoormann, Falkenstein, Schwarzenau, & Hohnsbein, 1998).

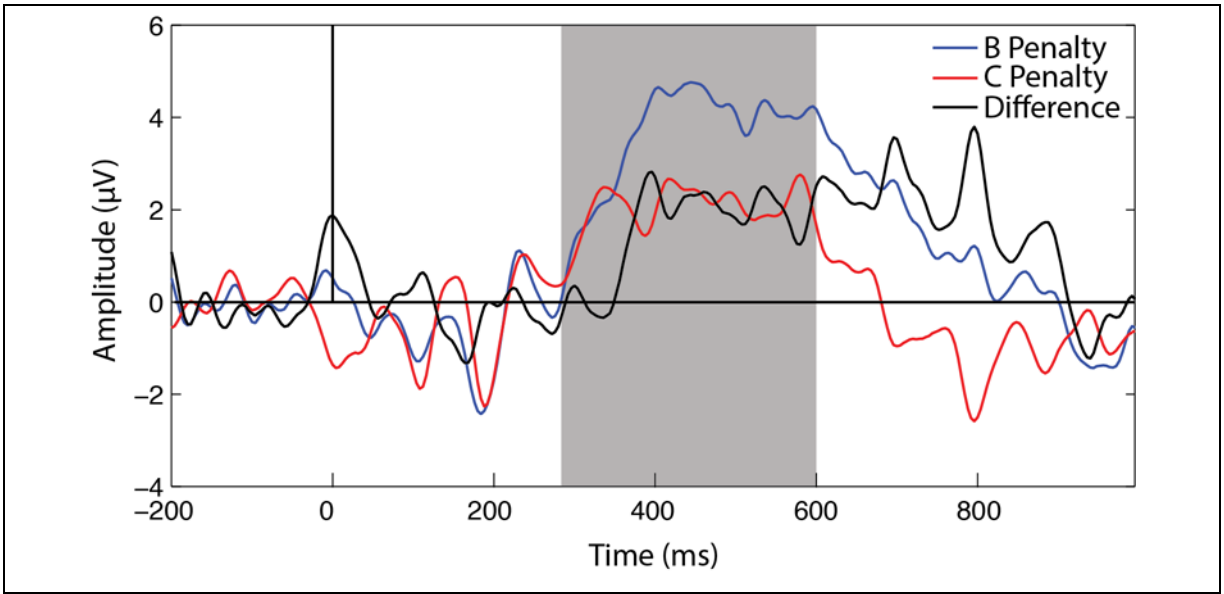


Figure 4. Event-related Potentials between 300 and 600 ms for Responses to B and C Deck Penalties and Their Difference at the Pz Electrode Site[5]

Consistent with previous literature examining feedback-related ERPs (San Martín et al., 2013), we analyzed the P300 using the mean ERP amplitude 300-600ms after stimulus onset (Figure 4 shaded area; see Appendix F for further detail). A topographical analysis of our difference data showed two potential peaks in the 300-600ms time window; one that was situated more frontally and peaked earlier, probably representing the P3a subcomponent of the P300, and another that was situated more posteriorly and peaked slightly later, probably reflecting the P3b subcomponent of the P300. We observed the greatest P300 amplitude differences in the parietal (Pz) region of the scalp (see Figure 5). Therefore, we used measures from this region in our analyses. Hence, our calculated independent variable was the difference in mean activity between the B penalty and the C penalty at the Pz electrode site (i.e., a B penalty minus C penalty difference score).

[5] We extracted mean amplitude for the P300 during the 300–600 ms post-stimulus epoch (shaded) deck penalties.

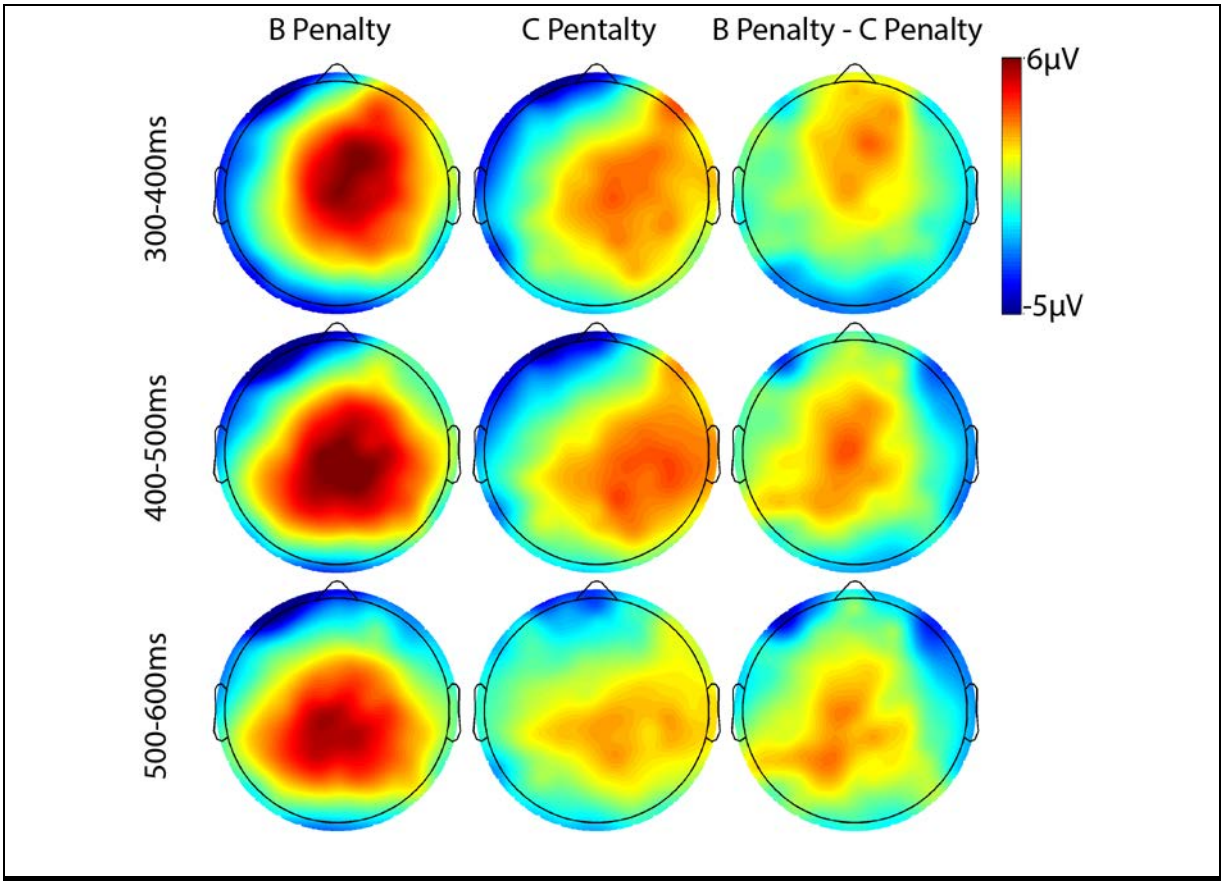


Figure 5. Topographical Heat Map of P300 Amplitudes for Responses to B and C Deck Penalties and Their Difference

First, we tested whether our "difference score" measure of risk predicted security warning disregard in the before-incident phase of the image classification task (H5). We found that the "difference score" had a significant influence on security warning disregard *before* the incident (standardized beta of -.277, t = -2.235 one-tailed, p < .05). This indicates a medium effect size following Cohen's classification of effect sizes (where small, medium, and large effects correspond to .10, .30, and .50, respectively (Cohen, 1992). Therefore, H5 was supported (see Table 6). This was in contrast to the self-reported measures, which had no effect before the security incident.

Table 6. Regression Results for the Effects of P300 Difference Score on Security Warning Disregard (Before a Security Incident)

H5: P300 difference score

Model	ß	Std. error	Standardized ß	*t*
Intercept	.766	.051	—	14.940***
P300 difference score	-.028	.013	-.277	-2.235*
Model statistics: R^2 = .077; f = 4.997, p = .029				

*** p < .001; ** p < .01; * p < .05; ns = not significant; one-tailed tests.

Next, we tested whether the "difference score" predicted security warning disregard better after a security incident compared to before a security incident (H6). A regression analysis showed that the "difference score" did predict security warning disregard after-incident (standardized beta of -.324, t = -2.750 one-tailed, p < .01) (see Table 7).

Table 7. Regression Results for the Effects of P300 Difference Score on Security Warning Disregard (After Security Incident)

H6: P300 difference score

Model	ß	Std. error	Standardized ß	*t*
Intercept	1.014	.153	—	6.649***
Native English speaker	-.355	.162	-.258	-2.185*
P300 difference score	-.036	.013	-.324	-2.750**
Model statistics: R^2 = .184; f = 6.644, p = .002				

*** p < .001; ** p < .01; * p < .05; ns = not significant; one-tailed tests.

The difference in the size of the standardized path coefficients before and after the security incident was -.047 (-.324 less -.277). To test whether this difference was significant, we used a SAS 9.2 macro to bootstrap our sample following the procedure described in (Hayes, 2009). In this approach, we randomly selected, with replacement, 62 observations from our dataset of 62. We then ran two separate regression models: the "difference score" regressed on security warning disregard "before" and "after" the security incident, and saved the resulting path coefficients of the models. We repeated this process to obtain 1,000 resamples and associated path coefficients because 1,000 or more resamples are recommended (Hayes, 2009). Next, we calculated the difference between the coefficients in each resample (e.g., standardized $\beta_{Resample1}$ "difference score" [after-incident] - standardized $\beta_{Resample1}$ "difference score" [before-incident], standardized $\beta_{Resample2}$ "difference score" [after-incident] - standardized $\beta_{Resample2}$ "difference score" [before-incident], etc.).

We next sorted the difference values of the resampled coefficients from largest to smallest to create a percentile-based confidence interval ci% using the formula k(.5 - ci/200) for the lower bound and the formula 1 + k(.5 + ci/200) for the upper bound, where k is the number of resamples (Hayes, 2009). In our case, we obtained 1,000 resamples and specified a 95 percent confidence interval. For the sorted values of the differences of the coefficients, the lower bound of the confidence interval was represented by the value in the 25th position, whereas the upper bound was denoted by the value in the 975th position. For the constructed confidence interval, if zero is not between the lower and upper bound, then one can state with ci% confidence that the difference between the coefficients is not zero (MacKinnon, 2008). Table 8 reports the results of the 95 percent confidence interval test. Because zero was included in the confidence interval, we conclude that the effect of "difference score" on after-incident behavior was not significantly greater than its effect on before-incident behavior. Therefore, H6 was not supported. The effect of the P300 "difference score" on behavior was essentially the same before as it was after the security incident.

Table 8. Bootstrapped Confidence Interval to Test for an Increase in the Strength of the P300 Difference Score After a Security Incident (H6)

Variable	Confidence interval		Zero in interval?	H6 supported?
	2.5% lower bound	97.5% upper bound		
Std. β difference score (after-incident) - Std. β difference score (before-incident)	-.230	.029	Yes	No

5.5. Comparing the Relative Predictive Strength of Self-reported and EEG Measures of Risk

Having tested the effects of the self-reported and EEG measures of risk separately, we then compared the relative strength of the self-reported risk perception measures and the EEG P300 "difference score" measure. Consistent with the results of H3 and H5, only the difference score had a significant effect on the security warning disregard *before* the security incident. Thus, H7 was

supported—that is, in every case (H7a for "perceived risk", H7b for "threat severity", and H7c for "threat susceptibility"), EEG measures of risk were more predictive of security warning disregard before incident than were the self-reported measures.

Finally, to test H8, we followed the same process as for testing H7, with the difference that we now examined "security warning disregard after-incident" as the dependent variable. The variables "perceived risk (post-test)", "threat susceptibility (post-test)", and "difference score" independently had significant negative effects on "security warning disregard after-incident" (see testing for H4 and H6 above). To test whether the path coefficient of the "difference score" was significantly greater than those of the self-reported measures, we followed the same bootstrapping procedure described for our tests of H6 above. We individually bootstrapped the effects of "perceived risk (post-test)", "threat susceptibility (post-test)", and "difference score" on after-incident behavior and saved the coefficients. This resulted in 1,000 resamples for each coefficient. We then calculated the difference between each resampled pair and sorted the resulting difference scores to create a 95 percent confidence interval. In both cases, zero was inside the 95 percent interval (see Table 9).

Table 9. Bootstrapped Confidence Interval to Compare the Strength of the P300 Difference Score with Self-reported Risk Measures (After Incident) (H8)

Variable	Confidence interval		Zero in interval?	H8 supported?
	2.5% lower bound	97.5% upper bound		
Std. β P300 difference score - std. β perceived risk (after-incident)	-.418	0.169	Yes	H8a: No
Std. β P300 difference score - std. β threat susceptibility (after-incident)	-.356	.290		

Thus, H8b for "threat severity" was supported, but H8a for "perceived risk" and H8c for "threat susceptibility" were not. The EEG measures were no more effective in predicting security warning disregard after the adverse incident was received than were the post-test measures of perceived risk and threat susceptibility. Table 10 summarizes the results of our hypothesis testing.

Table 10. Summary of Hypothesis Testing	
Hypothesis	**Supported?**
H1. Security warning disregard before a security incident will be higher than security warning disregard after a security incident.	Yes
H2. Pre-test self-reported measures of risk perception will be lower before a security incident than post-test self-reported measures of risk perception after a security incident.	
a. for perceived risk	No
b. for threat severity	Yes
c. for threat susceptibility	Yes
H3. Pre-test self-reported measures of risk perception will negatively predict security warning disregard before a security incident is imposed.	
a. for perceived risk	No
b. for threat severity	No
c. for threat susceptibility	No
d. for general risk appetite	No
e. for willingness to gamble lifetime income	No
H4. Post-test self-reported measures of risk perception will negatively predict security warning disregard after a security incident is imposed better than pre-test measures negatively predict security warning disregard before a security incident is imposed.	
a. for perceived risk	No
b. for threat severity	Yes
c. for threat susceptibility	No
H5. Pre-test P300 amplitude measures will negatively predict security warning disregard before a security incident is imposed.	Yes
H6. Pre-test P300 amplitude measures will negatively predict security warning disregard after a security incident is imposed better than will the same P300 amplitude measures before a security incident is imposed	No
H7. Pre-test P300 amplitude measures will negatively predict security warning disregard better than pre-test self-reported measures of risk perception before a security incident is imposed.	
a. for EEG superior to “perceived risk”	Yes
b. for EEG superior to “threat severity”	Yes
c. for EEG superior to “threat susceptibility”	Yes
H8. Pre-test P300 amplitude measures will negatively predict security warning disregard better than will post-test self-reported measures of risk perception after a security incident is imposed.	
a. for EEG superior to “perceived risk”	No
b. for EEG superior to “threat severity”	Yes
c. for EEG superior to “threat susceptibility”	No

6. Discussion

This study’s results provide several important contributions to research on information security risk perceptions and their measurement, which Table 11 summarizes below. In this section, we elaborate on each of these contributions.

Table 11. Research Contributions	
Element of research	**Contributions**
P300 measure of risk propensity	The P300 difference score proved the strongest predictor: it significantly predicted security warning disregard both before and after a security incident.
Self-reported measures of risk perceptions	Self-reported measures of information security risk perception did not predict security warning disregard before a security incident, which indicates a poor correspondence with behavior in this experimental setting.
P300 measure of risk propensity and self-reported measures of risk perceptions	After a security incident, "perceived risk" and "threat susceptibility" significantly predicted security warning behavior to essentially the same degree as the P300 difference score. This suggests that self-reported measures are better predictors when information security risks are salient. In contrast, the P300 difference score was a strong predictor even when information security risks were not salient.
Security warning behavior and self-reported measures of risk perceptions	Security warning disregard and self-reported measures of risk perception change with the introduction of an adverse consequence.

First, we found that the P300 difference score, derived from participants' P300 amplitudes in response to losses in the IGT, was the strongest predictor of security warning disregard in our study (H5, H7). It was also the most robust measure because it predicted security warning disregard consistently before and after a security incident (H6). Accordingly, with this study, we provide evidence that NeuroIS measures of risk propensity can predict security behavior. In doing so, we paper respond to the call to use NeuroIS methods to study information security behaviors (Anderson et al., 2012; Crossler et al., 2013).

Second, we found that a variety of self-reported risk measures—five different measures in all—failed to predict security warning disregard before the security incident was imposed (H3). This was a surprising finding and counter to our hypotheses. The levels of risk perception of the self-reported measures were moderate to high in all cases. However, despite this, these measures were weakly and insignificantly correlated with security warning disregard. This finding is consistent with previous studies on security and privacy risk that have showed that participants reported high levels of concern about their privacy and online security but later behaved contrarily to their stated apprehensiveness (Acquisti & Grossklags, 2004; Belanger et al., 2002; Norberg et al., 2007). This has important implications for research involving information security risk perceptions.

Third, this paper contributes the interesting finding that self-reported measures of information security risk were insignificant before a security incident; however, after an incident, "perceived risk" and "threat susceptibility" predicted security warning disregard more or less equally with the P300 difference score (H4, H8). This finding has been anticipated by Dimoka et al. (2012, 2011), who observed that many emotions—such as fear and uncertainty—are not processed consciously and are therefore difficult to measure using self-reported measures. Thus, initially, the self-reported measures did not accurately reflect participants' actual attitudes toward information security risk. However, after a security incident, attitudes toward information security risk became salient and were processed consciously by the participants. This conclusion is supported by the fact that measures of "threat severity" and "threat susceptibility" significantly increased after a security incident (H2).

In contrast, the P300 difference score measured a correlate of the neural response as early as 300 ms after receiving the IGT loss stimulus. The amplitude of the P300 ERP component has been shown to be sensitive to target probability and is influenced by participants' expectations (for a review, see Polich, 2007). As the probability of a loss in the various decks in the IGT was the same for each participant, the target probability (i.e., the probability of a loss in a particular deck) cannot account for variations in the difference scores between participants. Thus, in our experiment, the P300 difference

score appeared to be more in line with participants' actual security warning disregard. This measure was unrelated to the perceptions of information security risk participants espoused before the incident and was therefore a far more accurate predictor of actual security warning disregard (H5). After a security incident, the P300 difference score continued significantly to predict security warning disregard, which indicates robust predictive validity (H6; Straub et al., 2004).

Interestingly, "perceived risk" and "threat susceptibility", but not "threat severity", became significant after a security incident. Although this was contrary to our hypothesis, this result is reasonable given that the severity of the security incident for each participant was apparently nil—contrary to the dire claim reported by the hacker screen that data on the participant's laptop would be erased. Therefore, while participants recognized that they were more susceptible to malware than they initially thought, they had no cause to change their attitudes about the severity of such attacks. This explanation also holds for the measure of "perceived risk" because its items combine the concepts of susceptibility and severity into the same measure. Thus, because perceptions of susceptibility increased, the predictive power of "perceived risk" increased as well.

Finally, we showed that people's behavior and attitudes do change after a security incident—security warning disregard significantly decreased while risk perceptions significantly increased (H1, H2). This finding is also consistent with research showing that users substantially alter their computing behaviors to be more cautious after being compromised by malware (Fox, 2005; Good et al., 2005). This result, though expected, nonetheless demonstrates that information security risk perceptions are not static but change as people gain experience with information security threats.

6.1. Methodological Implications

From the aforementioned findings, two primary methodological implications are evident. First, the P300 difference score from the Iowa Gambling Task is a good measure of risk propensity and a significant predictor of security warning disregard. This indicates that researchers who wish to measure information security risk perceptions should consider using an EEG measure of risk because of its superior predictive power. It also suggests that other NeuroIS methods may be similarly effective in predicting information security behavior because of their ability to avoid measurement biases. Moreover, our findings demonstrate the value of capturing information security risk at an unconscious level, which is possible using a variety of NeuroIS methods, such as fMRI and galvanic skin response (Dimoka et al., 2011).

Second, our findings show that self-reported measures of information security risk and risk generally were not effective predictors of security behavior in this experimental setting until a security incident was salient in a person's recent experience. This suggests that researchers might profitably use self-reported measures in a post-test after an experimental treatment that simulates a security incident. Alternatively, researchers might try measuring past experiences with information security incidents (as per Anderson & Agarwal, 2010) to help qualify self-reported measures of risk. Additionally, self-reported measures of information security risk might be used to triangulate data collected using NeuroIS or other behavioral methods (Dimoka et al., 2011).

However, researchers should use caution when attempting to use self-reported measures of information security risk as a predictor of information security intentions. Although information security risks have been shown to be significant predictors of security-related intentions in the past (e.g., Guo et al., 2011; Malhotra et al., 2004), this and other research suggests that this predictive ability may not translate into actual security behavior (Acquisti & Grossklags, 2004; Belanger et al., 2002; Crossler et al., 2013).

6.2. Limitations and Future Research

This study has several limitations that point to future research opportunities. For example, this study examined security warning disregard as the only form of security behavior. It is possible that other forms of security behavior may be more amenable to prediction using self-reported measures of risk.

Similarly, this study used only one NeuroIS method—EEG. It is possible that using different NeuroIS methods such as fMRI may yield different results. Future research should attempt to measure

information security risk perceptions using other techniques to determine the most effective methods to measure the construct of information security risk. Further, one or more NeuroIS methods in combination could be used to produce more reliable measurements (Dimoka et al., 2012).

Another limitation is that we did not use EEG to measure information security risk per se. Rather, we used EEG to measure participants' responses to gains and losses in the IGT, a widely used technique in the Psychology and Neuroscience fields that has been shown to be strongly correlated with real-world risky behaviors (Buelow & Suhr, 2009). Nevertheless, there are two aspects of the IGT that deserve further consideration. The first is that, due to the structure of the IGT, there are relatively few B penalty trials, which could lead to noisier ERP measures associated with the B penalty trials. Noisy ERP signals are particularly problematic for peak measures (such as peak amplitude or latency to peak) and, as such, we used average ERP amplitude measures, which are more resistant to noisy measures (Luck, 2005).

The second consideration is that the relative frequency of penalty types differed between the B penalty and C penalty conditions. The P300 has been observed widely in ERP studies and seems to be generated any time a task requires stimulus discrimination (for a review, see Polich, 2007). Furthermore, the different subcomponents of the P300 (namely the P3a and P3b) may reflect different underlying neural computations. Previous P300 research has indicated that the amplitude of the P300 is influenced by the local probability of a target stimulus occurring (Polich & Margala, 1997). Because of how the IGT is structured, this almost certainly has an influence on the amplitude of the P300 in response to the B penalty (relatively infrequent) and the C penalty (relatively frequent). However, while stimulus probability explains the difference in the amplitudes in these conditions, it does not explain why this difference predicts subsequent security warning disregard. Indeed, further analysis found no evidence of either mediation or moderation of participants' selections of B or C decks on subsequent security warning disregard (see Appendix D). From these results, we conclude that the behavioral data in the IGT had no influence (either as a mediator or moderator) on the effect of the P300 difference score on security warning disregard in the image classification task.

Similarly, it has been suggested that the amplitude of the P300 is related to attentional processes (Kok, 1990) or neural inhibition (Polich, 2007). If participants devote fewer attentional resources to losses of different magnitudes or frequencies, or, alternatively, if they are less likely to inhibit non-related processes following a high-magnitude loss, they may be less likely to attend to security warnings in another context.

These disparate explanations are accounted for by another influential theory of the purpose of the P300—the context-updating theory (Donchin, 1981; Donchin & Coles, 1988). As we discuss in Section 2.5, the P300 reflects changing mental representations of the ongoing task structure in response to incoming stimuli. The context-updating theory can account for several findings in the P300 literature, including those listed above (Polich, 2007). For example, infrequent targets may demand more cognitive resources to update ongoing task representations and thus produce a larger P300. In the context of risk perception, the P300 may be reduced in cases of individuals who are less responsive to negative outcomes and are therefore more risk-seeking. Previous research has demonstrated that risk perceptions measured using EEG during the IGT do predict other risk behaviors (Bianchin & Angrilli, 2011; Schuermann, Kathmann, Stiglmayr, Renneberg, & Endrass, 2011). Our study's results are consistent with this previous body of work in that it shows that the P300 difference score does significantly predict security behavior. Thus, it exhibited strong predictive validity (Straub et al., 2004) and demonstrated its value as an information security risk measure.

Another limitation of our study is that it may not be very generalizable given the necessary artificiality of the laboratory environment and the use of student subjects. However, the raison d'être for laboratory experimentation is not external validity but precision and control (Dennis & Valacich, 2001; McGrath, 1981). Similarly, student subjects typically represent a homogenous sample that reduces noise and thereby provides the strictest test of the hypotheses (Calder, Phillips, & Tybout, 1982). Regardless, as young people spend a proportionately large amount of time online compared with the general population (Pew Research Center, 2012), students frequently encounter threats from security and are therefore a valid sample to study information security risk perceptions and security warning disregard.

Additionally, there is a chance that our design choice may have introduced some reciprocal causation. That is, in the interest of completing the task, participants may have disregarded security warnings as quickly as possible to continue with the classification. Thus, those who sought to optimize the image classification task may have exhibited the highest level of security warning disregard (the dependent variable). However, we considered this possibility in our design of the experiment for two reasons. First, we intentionally tried to emulate real life, in which an individual is typically striving to complete a task when a pop-up warning or message is received and therefore tries to remove an interfering item as quickly as possible. Dual-task interference theory explains the difficulty people have when trying to perform two or more tasks, even relatively simple ones (Pashler, 1994). Further, recent research (Jenkins & Durcikova, 2013) asserts that is it possible to predict a disconnect between security intentions and behaviors due to the cognitive load of two simultaneous tasks. Consequently, we designed the experiment so that the participants would have to deal with a time-sensitive competing task when responding to the security warning. Second, as we note in Section 3.1, people routinely sacrifice promised information security for some other utility. Thus, the temptation to ignore the security warnings received on their personal laptops in exchange for better performance in the image classification also mirrored real life.

Finally, our sample consisted of 62 participants. While more data is generally better, our regression models involved only one or two predictors at one time. As the "rule of ten" sample size heuristic for regression suggests ten observations per predictor (Chin, 1998), this indicates that our sample size was sufficient. Further, Dimoka (2012) points out that, although sample sizes tend to be smaller for NeuroIS studies due to the expense and time commitment required per subject, NeuroIS methods generally provide many data points per subject. In our case, we collected 400 behavioral observations per participant while recording EEGs at 250 Hz. Thus, the amount of data captured and used in our analysis was actually much greater than a sample size of 62 would suggest.

7. Conclusion

With this study, we show that participants' EEG P300 amplitudes in response to losses in a risk-taking experimental task strongly predicted security warning disregard in a subsequent and unrelated computing task using participants' own laptop computers. By comparison, self-reported measures of information security risk did not predict security warning disregard. However, after secretly simulating a malware incident on the participants' own laptops, post-test measures of information security risk perception did predict participants' security warning disregard after a security incident. This suggests that self-reported measures of information security risk can significantly predict security behavior when security risks are salient. In contrast, the P300 risk measure is a significant predictor of security behavior both before and after a security incident is imposed, which highlights the robustness of NeuroIS methods in measuring risk perceptions and their value in predicting security behavior.

Acknowledgments

We thank James V. Hansen for collaborating with us on an early version of this study. Support for this research was provided by the David Berg Center for Ethics and Leadership at the Joseph M. Katz Graduate School of Business and College of Business Administration at the University of Pittsburgh through a grant from the BNY Mellon Foundation of Western PA.

References

Abraham, S., & Chengalur-Smith, I. (2010). An overview of social engineering malware: Trends, tactics, and implications. *Technology in Society, 32*(3), 183-196.

Acquisti, A., & Grossklags, J. (2004). Privacy attitudes and privacy behavior. In L. J. Camp & S. Lewis (Eds.), *Economics of information security* (Vol. 12, pp. 165-178). Boston, MA: Springer.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179-211.

Amer, T. S., & Maris, J.-M. B. (2007). Signal words and signal icons in application control and information technology exception messages—hazard matching and habituation effects. *Journal of Information Systems, 21*(2), 1-25.

Anderson, B., Vance, A., Hansen, J., Kirwan, B., Eargle, D., Hinkle, L., & Weagel, A. (2012). *Neural correlates of gender differences in distinguishing malware warnings and legitimate websites: A NeuroIS study.* Paper presented at the IFIP WG8.11/WG11.13, Provo, UT.

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

Anderson, R. (2008). *Security engineering: A guide to building dependable distributed systems* (2nd ed.). Indianapolis, IN: Wiley.

Bansal, G., Zahedi, F., & Gefen, D. (2010). The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decision Support Systems, 49*(2), 138-150.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182.

Barsky, R. B., Juster, F. T., Kimball, M. S., & Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *The Quarterly Journal of Economics, 112*(2), 537-579.

Bashore, T. R., & Ridderinkhof, K. R. (2002). Older age, traumatic brain injury, and cognitive slowing: Some convergent and divergent findings. *Psychological Bulletin, 128*(1), 151-198.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition, 50*(1), 7-15.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (2005). The Iowa Gambling Task and the somatic marker hypothesis: Some questions and answers. *Trends in Cognitive Sciences, 9*(4), 159-162.

Belanger, F., Hiller, J. S., & Smith, W. J. (2002). Trustworthiness in electronic commerce: The role of privacy, security, and site attributes. *The Journal of Strategic Information Systems, 11*(3), 245-270.

Bianchin, M., & Angrilli, A. (2011). Decision preceding negativity in the Iowa Gambing Task: An ERP study. *Brain and Cognition, 75*(3), 273-280.

Bolla, K. I., Eldreth, D. A., London, E. D., Kiehl, K. A., Mouratidis, M., Contoreggi, C., Matochik, J. A., Kurian, V., Cadet, J. L., Kimes, A. S., Funderburk, F. R., & Ernst, M. (2003). Orbitofrontal cortex dysfunction in abstinent cocaine abusers performing a decision-making task. *NeuroImage, 19*(3), 1085-1094.

Brooker, G. (1984). An assessment of an expanded measure of perceived risk. *Advances in Consumer Research, 11*(1), 439-441.

Buelow, M. T., & Suhr, J. A. (2009). Construct validity of the Iowa Gambling Task. *Neuropsychology Review, 19*(1), 102-114.

Buelow, M. T., & Suhr, J. A. (2013). Personality characteristics and state mood influence individual deck selections on the Iowa Gambling Task. *Personality and Individual Differences, 54*(5), 593-597.

Calder, B. J., Phillips, L. W., & Tybout, A. M. (1982). The concept of external validity. *Journal of Consumer Research*, 240-244.

Clayson, P. E., Baldwin, S. A., & Larson, M. J. (2013). How does noise affect amplitude and latency measurement of event-related potentials (ERPs)? A methodological critique and simulation study. *Psychophysiology*, *50*(2), 174-186.

Carte, T. A., & Russell, C. J. (2003). In pursuit of moderation: nine common errors and teir solutions. *MIS Quarterly, 27*(3), 479-501.

Chin, W. W. (1998). The partial least squares approach for structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 295-336). London: Lawrence Erlbaum Associates.

Christie, G. J., & Tata, M. S. (2009). Right frontal cortex generates reward-related theta-band oscillatory activity. *NeuroImage, 48*(2), 415-422.
Cohen, J. (1992). A power primer. *Psychological bulletin, 112*(1), 155-159.
Crone, E. A., Vendel, I., & van der Molen, M. W. (2003). Decision-making in disinhibited adolescents and adults: Insensitivity to future consequences or driven by immediate reward? *Personality and Individual Differences, 35*(7), 1625-1641.
Crossler, R. E., Johnston, A. C., Lowry, P. B., Hu, Q., Warkentin, M., & Baskerville, R. (2013). Future directions for behavioral information security research. *Computers & Security, 32*(0), 90-101.
D'Arcy, J., & Herath, T. (2011). A review and analysis of deterrence theory in the IS security literature: Making sense of the disparate findings. *European Journal of Information Systems, 20*(6), 643-658.
Dennis, A. R., & Valacich, J. S. (2001). Conducting experimental research in information systems. *Communications of the Association for Information Systems, 7*.
Dholakia, U. M. (2001). A motivational process model of product involvement and consumer risk perception. *European Journal of Marketing, 35*(11/12), 1340-1362.
Dien, J. (2010). The ERP PCA toolkit: An open source program for advanced statistical analysis of event-related potential data. *Journal of Neuroscience Methods, 187*(1), 138-145.
Dien, J., Michelson, C. A., & Franklin, M. S. (2010). Separating the visual sentence N400 effect from the P400 sequential expectancy effect: Cognitive and neuroanatomical implications. *Brain Research, 1355*, 126-140.
Dimoka, A. (2010). What does the brain tell us about trust and distrust? Evidence from a functional neuroimaging study. *MIS Quarterly, 34*(2), 373-396.
Dimoka, A. (2012). How to conduct a functional magnetic resonance (fMRI) study in social science research. *MIS Quarterly, 36*(3), 811-840.
Dimoka, A., Banker, R. D., Benbasat, I., Davis, F. D., Dennis, A. R., Gefen, D., Gupta, A., Ischebeck, A., Kenning, P., Müller-Putz, G., Pavlou, P. A., Riedl, R., vom Brocke, J., & Weber, B. (2012). On the use of neurophysiological tools in IS research: Developing a research agenda for NeuroIS. *MIS Quarterly, 36*(3), 679-702.
Dimoka, A., Pavlou, P. A., & Davis, F. D. (2011). Research commentary—neuroIS: The potential of cognitive neuroscience for information systems research. *Information Systems Research, 22*(4), 687-702.
Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. *Information Systems Research, 17*(1), 61-80.
Donchin, E. (1981). Surprise!… surprise? *Psychophysiology, 18*(5), 493-513.
Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11*(3), 357-374.
Egelman, S., Cranor, L. F., & Hong, J. (2008). *You've been warned: An empirical study of the effectiveness of web browser phishing warnings.* Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy.
Ernst, M., Bolla, K., Mouratidis, M., Contoreggi, C., Matochik, J. A., Kurian, V., Cadet J. L., Kimes, A. S., & London, E. D. (2002). Decision-making in a risk-taking task: A PET study. *Neuropsychopharmacology, 26*(5), 682-691.
Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: A tutorial. *Biological Psychology, 51*(2), 87-107.
Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology, 54*(1), 297-327.
Fjell, A., & Walhovd, K. (2001). P300 and neuropsychological tests as measures of aging: Scalp topography and cognitive changes. *Brain Topography, 14*(1), 25-40.
Fox, S. (2005). *Spyware*. Pew Internet & American Life Project. Retrieved April 24, 2013, from http://www.pewinternet.org/Reports/2005/Spyware
Fukui, H., Murai, T., Fukuyama, H., Hayashi, T., & Hanakawa, T. (2005). Functional activity related to risk anticipation during performance of the Iowa Gambling Task. *Neuroimage, 24*(1), 253-259.
Furnell, S., & Clarke, N. (2012). Power to the people? The evolving recognition of human aspects of security. *Computers & Security, 31*(8), 983-988.
Gefen, D. (2002). Customer loyalty in e-commerce. *Journal of the Association for Information Systems, 27*(51), 51.

Good, N., Dhamija, R., Muligan, D., & Konstan, J. (2005). Stopping spyware at the gate: A user study of privacy, notice and spyware. *Proceedings of the 2005 symposium on usable privacy and security* (pp. 43-52). New York, NY: ACM.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological review, 102*(1), 4-27.

Grewal, D., Gotlieb, J., & Marmorstein, H. (1994). The moderating effects of message framing and source credibility on the price-perceived risk relationship. *Journal of Consumer Research*, 145-153.

Guo, K. H., Yuan, Y., Archer, N. P., & Connelly, C. E. (2011). Understanding nonmalicious security violations in the workplace: A composite behavior model. *Journal of Management Information Systems, 28*(2), 203-236.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs, 76*(4), 408-420.

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.

Herley, C. (2009). *So long, and no thanks for the externalities: The rational rejection of security advice by users.* Paper presented at the Proceedings of the 2009 New Security Paradigms Workshop.

Hong, W., & Thong, J. Y. (2013). Internet privacy concerns: An integrated conceptualization and four empirical studies. *MIS Quarterly, 37*(1), 275-298.

Hoormann, J., Falkenstein, M., Schwarzenau, P., & Hohnsbein, J. (1998). Methods for the quantification and statistical testing of ERP differences across conditions. *Behavior Research Methods, Instruments, & Computers, 30*(1), 103-109.

Jarvenpaa, S. L., Tractinsky, N., & Saarinen, L. (1999). Consumer trust in an Internet store: A cross-cultural validation. *Journal of Computer-Mediated Communication, 5*(2), 1-36.

Jenkins, J. L., & Durcikova, A. (2013). *What, I shouldn't have done that? The influence of training and just-in-time reminders on secure behavior.* Paper presented at the International Conference for Information Systems, Milan, Italy.

Johnston, A. C., & Warkentin, M. (2010). Fear appeals and information security behaviors: An empirical study. *MIS Quarterly, 34*(3), 549-566.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Pyschologist, 39*(4), 341-350.

Kam, C. D., & Simas, E. N. (2010). Risk orientations and policy frames. *Journal of Politics, 72*(2), 381-396.

Kaplan, L. B., Szybillo, G. J., & Jacoby, J. (1974). Components of perceived risk in product purchase: A cross-validation. *Journal of Applied Psychology, 59*(3), 287-291.

Kirwan, C. B., Shrager, Y., & Squire, L. R. (2009). Medial temporal lobe activity can distinguish between old and new stimuli independently of overt behavioral choice. *Proceedings of the National Academy of Sciences of the United States of America, 106*(34), 14617-14621.

Kok, A. (1990). Internal and external control: a two-factor model of amplitude change of event-related potentials. *Acta psychologica, 74*(2-3), 203-236.

LeDoux, J. (2003). The emotional brain, fear, and the amygdala. *Cellular and Molecular Neurobiology, 23*(4-5), 727-738.

Liang, H., & Xue, Y. (2010). Understanding security behaviors in personal computer usage: A threat avoidance perspective. *Journal of the Association for Information Systems, 11*(7), 394-413.

Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. New York, NY: Erlbaum.

Maia, T. V., & McClelland, J. L. (2004). A reexamination of the evidence for the somatic marker hypothesis: What participants really know in the Iowa Gambling Task. *Proceedings of the National Academy of Sciences of the United States of America, 101*(45), 16075-16080.

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research, 15*(4), 336-355.

Mandiant. (2013). *APT1: Exposing one of China's cyber espionage units*. Mandiant. Retrieved April 26, 2013, from http://intelreport.mandiant.com/Mandiant_APT1_Report.pdf

Massar, S. A. A., Kenemans, J. L., & Schutter, D. J. L. G. (2013). Resting-state EEG theta activity and risk learning: Sensitivity to reward or punishment? *International Journal of Psychophysiology*.

Massar, S. A. A., Rossi, V., Schutter, D. J. L. G., & Kenemans, J. L. (2012). Baseline EEG theta/beta ratio and punishment sensitivity as biomarkers for feedback-related negativity (FRN) and risk-taking. *Clinical Neurophysiology, 123*(10), 1958-1965.

McGrath, J. E. (1981). Dilemmatics: The study of research choices and dilemmas. *American Behavioral Scientist, 25*, 179-210.
McGraw, G., & Felten, E. W. (1999). *Securing Java: Getting down to business with mobile code*. New York, NY: John Wiley & Sons.
Ng, B. Y., & Feng, A. E. (2006). An exploratory study on managerial security concerns in technology start-ups. *PACIS 2006 Proceedings*.
Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs, 41*(1), 100-126.
Nunnally, J. C. (1970). *Introduction to psychological measurement*. New York: McGraw-Hill.
Oberg, S. A., Christie, G. J., & Tata, M. S. (2011). Problem gamblers exhibit reward hypersensitivity in medial frontal cortex during gambling. *Neuropsychologia, 49*(13), 3768-3775.
Orne, M. T., & Whitehouse, W. G. (2000). Demand characteristics. In A. E. Kazdin (Ed.), *Encyclopedia of* psychology (Vol. 2, pp. 469-470). Washington, DC: American Psychological Association.
Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychological Bulletin, 116*(2), 220-244.
Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes*. San Diego: Academic Press.
Pew Research Center. (2012). Internet use and home broadband connections (mobile survey). *Pew Internet & American Life Project*. Retrieved from http://www.pewinternet.org/Infographics/2012/Internet-Use-and-Home-Broadband-Connections
Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879-903.
Polezzi, D., Sartori, G., Rumiati, R., Vidotto, G., & Daum, I. (2010). Brain correlates of risky decision-making. *NeuroImage, 49*(2), 1886-1894.
Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology, 118*(10), 2128-2148.
Polich, J., & Corey-Bloom, J. (2005). Alzheimer's disease and P300: Review and evaluation of task and modality. *Current Alzheimer Research, 2*(5), 515-525.
Polich, J., & Margala, C. (1997). P300 and probability: Comparison of oddball and single-stimulus paradigms. *International Journal of Psychophysiology, 25*(2), 169-176.
Prentice-Dunn, S., & Rogers, R. W. (1986). Protection motivation theory and preventive health: Beyond the health belief model. *Health Education Research, 1*(3), 153-161.
Rogers, R. W. (1975). A protection motivation theory of fear appeals and attitude change. *The Journal of Psychology, 91*(1), 93-114.
San Martín, R., Appelbaum, L. G., Pearson, J. M., Huettel, S. A., & Woldorff, M. G. (2013). Rapid brain responses independently predict gain maximization and loss minimization during economic decision making. *The Journal of Neuroscience, 33*(16), 7011-7019.
Schneier, B. (2004). *Secrets and lies: Digital security in a networked world*. Hoboken, NJ: Wiley.
Schonberg, T., Fox, C. R., & Poldrack, R. A. (2011). Mind the gap: Bridging economic and naturalistic risk-taking with cognitive neuroscience. *Trends in Cognitive Sciences, 15*(1), 11-19.
Schuermann, B., Endrass, T., & Kathmann, N. (2012). Neural correlates of feedback processing in decision-making under risk. *Frontiers in Human Neuroscience, 6*.
Schuermann, B., Kathmann, N., Stiglmayr, C., Renneberg, B., & Endrass, T. (2011). Impaired decision making and feedback evaluation in borderline personality disorder. *Psychol Med, 41*(9), 1917-1927.
Schutter, D. J. L. G., & Van Honk, J. (2005). Electrophysiological ratio markers for the balance between reward and punishment. *Cognitive Brain Research, 24*(3), 685-690.
Singh, M., & Sungkarat, W. (2008). Dynamic fMRI of a decision-making task. In X. P. Hu & A. V. Clough (Eds.), *Medical Imaging 2008: Physiology, Function, and Structure from Medical Images*, *6916,* 691608-691608.
Skowronski, J. J., & Lawrence, M. A. (2001). A comparative study of the implicit and explicit gender attitudes of children and college students. *Psychology of Women Quarterly, 25*(2), 155-165.
Slovic, P. (1987). Perception of risk. *Science, 236*(4799), 280-285.

Spangler, W. D. (1992). Validity of questionnaire and TAT measures of need for achievement: Two meta-analyses. *Psychological Bulletin, 112*(1), 140-154.

Stewart, J., Acevedo, A., & Ownby, R. (2012). Examining the influence of executive function on medication adherence in individuals with HIV-Type 1. *Archives of Clinical Neuropsychology, 27*(6), 576-685.

Straub, D., Boudreau, M.-C., & Gefen, D. (2004). Validation guidelines for IS positivist research. *Communications of the Association for Information Systems, 13*, 380-427.

Suhr, J. A., & Tsanadis, J. (2007). Affect and personality correlates of the Iowa Gambling Task. *Personality and Individual Differences, 43*(1), 27-36.

Sunshine, J., Egelman, S., Almuhimedi, H., Atri, N., & Cranor, L. F. (2009). *Crying wolf: An empirical study of SSL warning effectiveness.* Paper presented at the SSYM'09 Proceedings of the 18th conference on USENIX Security Symposium, Montreal, Canada.

Tanabe, J., Thompson, L., Claus, E., Dalwani, M., Hutchison, K., & Banich, M. T. (2007). Prefrontal cortex activity is reduced in gambling and nongambling substance users during decision-making. *Human Brain Mapping, 28*(12), 1276-1286.

Theorell, T., & Hasselhorn, H. M. (2005). On cross-sectional questionnaire studies of relationships between psychosocial conditions at work and health—are they reliable? *International Archives of Occupational and Environmental Health, 78*(7), 517-522.

Toplak, M. E., Sorge, G. B., Benoit, A., West, R. F., & Stanovich, K. E. (2010). Decision-making and cognitive abilities: A review of associations between Iowa Gambling Task performance, executive functions, and intelligence. *Clinical Psychology Review, 30*(5), 562-581.

van Honk, J., Hermans, E. J., Putman, P., Montagne, B., & Schutter, D. J. L. G. (2002). Defective somatic markers in sub-clinical psychopathy. *Neuroreport, 13*(8), 1025-1027.

van Honk, J., Schutter, D. J. L. G., Hermans, E. J., & Putman, P. (2003). Low cortisol levels and the balance between punishment sensitivity and reward dependency. *Neuroreport, 14*(15), 1993-1996.

Van Slyke, C., Shim, J., Johnson, R., & Jiang, J. (2006). Concern for information privacy and online consumer purchasing. *Journal of the Association for Information Systems, 7*(6), 415-444.

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.

Weller, J. A., Levin, I. P., & Bechara, A. (2010). Do individual differences in Iowa Gambling Task performance predict adaptive decision making for risky gains and losses? *Journal of Clinical and Experimental Neuropsychology, 32*(2), 141-150.

Willison, R., & Warkentin, M. (2013). Beyond deterrence: An expanded view of employee computer abuse. *MIS Quarterly, 37*(1), 1-20.

Winkielman, P., & Berridge, K. C. (2004). Unconscious emotion. *Current Directions in Psychological Science, 13*(3), 120-123.

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

Workman, M., Bommer, W. H., & Straub, D. (2008). Security lapses and the omission of information security measures: A threat control model and empirical test. *Computers in Human Behavior, 24*(6), 2799-2816.

Xu, H., Luo, X. R., Carroll, J. M., & Rosson, M. B. (2011). The personalization privacy paradox: An exploratory study of decision making process for location-aware marketing. *Decision Support Systems, 51*(1), 42-52.

Yee, K.-P. (2004). Aligning Security and Usability. *Security & Privacy, IEEE, 2*(5), 48-55.

Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *The Journal of Neuroscience, 24*(28), 6258-6264.

Appendices

Appendix A: Survey Instrumentation

Table A-1. Survey Constructs

Construct	Pre/post (mean, std.)	Questions	Source
Willingness to gamble lifetime income	Pre-test (1.65, 1.19)	Suppose that you are the only income earner in the family, and you have a good job guaranteed to give you and your current (family) income every year for life. You are given the opportunity to take a new and equally good job, with a 50-50 chance that it will double your (family) income and a 50-50 chance that it will cut your (family) income by a third. Would you take the new job?	Barsky et al., 1997
		- If yes, suppose the chances were 50-50 that it would double your (family) income, and 50-50 that it would cut it in half. Would you still take the new job?	
		- If no to the first question, suppose the chances were 50-50 that it would double your (family) income and 50-50 that it would cut it by 20 percent. Would you then take the new job?	
General risk appetite	Pre-test (27.56, 5.58)	Some people say you should be cautious about making major changes in life. Suppose these people are located at 1. Others say that you will never achieve much in life unless you act boldly. Suppose these people are located at 7. Others have views in between. Where would you place yourself on this scale?	Kam & Simas, 2010
		I would like to explore strange places.	
		I like to do frightening things.	
		I like new and exciting experiences, even if I have to break the rules.	
		I prefer friends who are exciting and unpredictable.	
		In general, it is easy for me to accept taking risks.	
Perceived security risk of malware†	Pre-test (14.19, 3.99) Post-test (14.76, 4.08)	Ignoring malware warning screens can cause damages to computer security.	Guo et al., 2011
		Ignoring malware warning screens can put important data at risk.	
		Ignoring malware warning screens will most likely cause security breaches.	
Threat susceptibility†	Pre-test (9.31, 4.13) Post-test (12.42, 4.32)	My computer is at risk for becoming infected with malware.	Johnston & Warkentin, 2010
		It is likely that my computer will become infected with malware.	
		It is possible that my computer will become infected with malware.	
Threat severity†	Pre-test (11.51, 5.01) Post-test (14.77, 4.43)	If my computer were infected by malware, it would be severe.	Johnston & Warkentin, 2010
		If my computer were infected by malware, it would be serious.	
		If my computer were infected by malware, it would be significant.	
Bias‡	Post-test (3.29, .77)	How much did the pre-study survey influence your behavior on the image classification task?	None
Malware warning screen realism*	Post-test (6.76, 3.03)	On a scale of 0 to 10, how realistic do you think the following screen is? (malware warning screen)	None
Hacker screen realism*	Post-test (4.76, 3.44)	On a scale of 0 to 10, how realistic do you think the following screen is? (hack screen)	

Table A-1. Survey Constructs (cont.)

Construct	Pre/post (mean, std.)	Questions	Source
Malware warning screen concern**	Post-test (4.47, 2.95)	On a scale of 0 to 10, how concerned did the following screen make you feel during the image classification task? (malware warning screen)	None
Hacker screen concern**	Post-test (6.47, 2.91)	On a scale of 0 to 10, how concerned did the following screen make you feel during the image classification task? (hack screen)	
Demographic questions	Pre-test	What is your age? (mean 21.84, std. 1.96)	Control variables
		What is your gender? (male: 46, female 16)	
		What is your handedness? (right: 55, left: 7)	
		Are you a native English speaker? (native: 55, not native: 7)	
		Do you have normal/corrected to normal vision? (yes: 59, no: 3)	
		Have you ever had an EEG? (yes: 6, no: 56)	
		Have you ever been treated for a neurological or psychiatric condition? (yes: 3, no: 59)	
		Are you color blind? (yes: 4, no: 58)	

† These questions used a 7-point Likert scale with a range from 1 (strongly disagree) to 7 (strongly agree).

‡ This question used a 5-point Likert scale with a range from 1 (not at all) to 5 (very strongly).

* These questions had a scale from 0 (not realistic) to 10 (100% realistic) and from 0 (not concerned at all) to 10 (extremely concerned).

** This question had a scale from 0 (not concerned) at all to 10 (extremely concerned).

Table A-2. List of Acronyms	
DTPB	Decomposed theory of planned behavior
EEG	Electroencephalography
EFA	Exploratory factor analysis
EMF	Electric magnetic fields
ERP	Event-related potentials
fMRI	Functional magnetic resonance imaging
ICA	Independent components analysis
IGT	Iowa Gambling Task
ISP	Information security policy
PCA	Principal components analysis
PET	Positron emission technology
PLS	Partial least squares
PMT	Protection motivation theory
PSRM	Perceived security risk of malware
Pz	Parietal region of scalp
TPB	Theory of planned behavior
TSEV	Threat severity of malware
TSUS	Threat susceptibility of malware
TTAT	Technology threat avoidance theory
SCR	Skin conductance responses
SWD	Security warning disregard

Table A-3. Construct Correlation Matrix

Construct	Mean	Std.	1	2	3	4	5	6	7	8	9	10	11
Security warning disregard—before SI (1)	0.74	0.39	N/A										
Security warning disregard—after SI (2)	0.65	0.45	.70**	N/A									
P300 difference score (3)	3.43	2.85	-.18*	-.27**	N/A								
Willingness to gamble lifetime income (4)	1.51	1.18	.01	.09	.09	N/A							
General risk appetite (5)	27.86	5.84	-.13	-.13	.06	.27**	.76						
Perceived security risk of malware-pre-test (6)	14.33	3.89	-.04	-.04	-.04	-.14	.05	.84					
Perceived security risk of malware-post-test (7)	14.98	3.86	.07	-.04	-.02	.00	.05	.49**	.87				
Threat susceptibility of malware—pre-test (8)	9.45	4.04	-.04	-.08	.04	-.13	-.14	.07	.10	.79			
Threat susceptibility of malware—post-test (9)	12.42	4.04	-.06	-.17†	.07	-.17	-.08	.21*	.54**	.58**	.81		
Threat severity of malware—pre-test (10)	11.04	4.69	.02	.00	.11	.04	.00	.32**	.39**	.35**	.48**	.95	
Threat severity of malware—post-test (11)	14.55	4.37	.12	.01	.10	-.08	-.04	.25**	.59**	.03	.35**	.50**	.95

N.B. Cronbach's α provided on the diagonal where applicable.
* $p < .05$; ** $p < .01$; † This correlation is significant when the variable native English speaker is controlled for.

Appendix B: Criteria for Selection of IS Security Risk Perceptions Measures

To identify measures of IS security risk perception, we performed a literature review using the following criteria. First, we searched all issues of the AIS Basket of six journals from their inception to February 2013 for papers that mentioned security in the title, abstract, or keywords. This resulted in 128 paper. From this set, we narrowed the paper to those that also mentioned the word "risk" anywhere in their body. We then searched each of the resulting papers for survey items related to IS risk perceptions, which yielded nine papers. Of these, two addressed privacy rather than security issues and so we excluded them. From this final set of seven, three papers measured risk perceptions using a combination of threat severity and threat susceptibility measures (Herath & Rao, 2009; Johnston & Warkentin, 2010; Liang & Xue, 2010). We chose the Johnston and Warkentin (2010) measures because they are representative of this set of papers and because they also measured risk perceptions of malware specifically. The other four papers (Anderson & Agarwal, 2010; Gefen, 2002; Guo et al., 2011; Van Slyke, Shim, Johnson, & Jiang, 2006) used items to measure risk perceptions as a single construct. We therefore chose the Guo et al. (2011) measures because they are representative of this set of papers and because they were most naturally adaptable to the context of malware. In summary, we chose two sets of items that are representative of IS security risk perception measures in use in the IS literature. Further, selecting two different sets of IS security risk perception items ensured that our results were not dependent on any one set of items.

Table B-1. Security-Related Risk Perception Items

Reference	Measures
Guo et al. (2011)	Perceived Security risk of NMSV (non-malicious security violations) • Risk 1: the action can cause damages to computer security. • Risk 2: the action can put important data at risk. • Risk 3: the action will most likely cause security breaches.
Anderson & Agarwal (2010)	Concern (adapted from Ellen and Wiener 1991; Ho 1998; Obermiller 1995) Anchors 1 = not at all concerned, 7 = very concerned Some experts have warned that hackers may try to cripple major American businesses or the government by breaking into their computers or by using home computers to attack other computers using the Internet. How concerned are you that hackers might...? • Harm American corporations or the government by breaking into their computers • Break into home computers and use them to attack computers owned by American corporations or the government • Break into your home computer and use your e-mail account to send spam to others • Use home computers to spread a virus over the Internet that harms other computers • Steal or change data stored on your computer • Gain access to your personal financial information • Gain access to your personal health/medical information • Gain access to other personal data (such as family photos, hobby information, shopping preferences, and/or school data)
Johnston & Warkentin (2010)	Threat severity • If my computer were infected by spyware, it would be severe (TSEV1). • If my computer were infected by spyware, it would be serious (TSEV2) • If my computer were infected by spyware, it would be significant (TSEV3). Threat susceptibility • My computer is at risk of becoming infected with spyware (TSUS1). • It is likely that my computer will become infected with spyware (TSUS2). • It is possible that my computer will become infected with spyware (TSUS3).

Table B-1. Security-Related Risk Perception Items (cont.)

Reference	Measures
Liang & Xue (2010)	Perceived susceptibility (1 = strong disagree, 7 = strongly disagree) • It is extremely likely that my computer will be infected by spyware in the future. • My chances of getting spyware are great. • There is a good possibility that my computer will have spyware. • I feel spyware will infect my computer in the future. • It is extremely likely that spyware will infect my computer. Perceived severity (1 = innocuous, 7 = extremely devastating) • Spyware would steal my personal information from my computer without my knowledge. • Spyware would invade my privacy. • My personal information collected by spyware could be misused by cyber criminals. • Spyware could record my Internet activities and send it to unknown parties. • My personal information collected by spyware could be subject to unauthorized secondary use. • My personal information collected by spyware could be used to commit crimes against me. • Spyware would slow down my Internet connection. • Spyware would make my computer run more slowly. • Spyware would cause system crash on my computer from time to time. • Spyware would affect some of my computer programs and make them difficult to use. Perceived threat (1 = strong disagree, 7 = strongly disagree) • Spyware poses a threat to me. • The trouble caused by spyware threatens me. • Spyware is a danger to my computer. • It is dreadful if my computer is infected by spyware. • It is risky to use my computer if it has spyware.
Herath & Rao (2009)	Perceived probability of security breach • How likely is it that a security violation will cause a significant outage that will result in loss of productivity? • How likely is it that a security violation will cause a significant outage to the Internet that results in financial losses to organizations? • How likely is it that the organization will lose sensitive data due to a security violation? Perceived severity of security breach • I believe that information stored on organization computers is vulnerable to security incidents. • I believe the productivity of organization and its employees is threatened by security incidents. • I believe the profitability of organizations is threatened by security incidents.
Van Slyke et al. (2006)	Risk perceptions. All anchors on 7-point scale anchored on very strongly disagree to very strongly agree. • How would you characterize the decision of whether to buy a product from this Web retailer (Amazon.com/Half.com)? (Anchors: very significant risk to very significant opportunity) • How would you characterize the decision of whether to buy a product from this Web retailer (Amazon.com/Half.com)? (Anchors: very high potential for loss to very high potential for gain) • How would you characterize the decision of whether to buy a product from this Web retailer (Amazon.com/Half.com)? (Anchors: very negative situation to very positive situation)
Gefen, D. (2002)	Perceived risk with vendor: • There is a significant threat doing business with Amazon.com. • There is a significant potential for loss in doing business with Amazon.com. • There is a significant risk in doing business with Amazon.com. • My credit card information may not be secure with Amazon.com.

Appendix C: Tests for Homogeneity of Sample Groups

We split the sample into two groups in two different ways to check for homogeneity of the sample between "yes-click" and "no-click" groups. The splits were set based on each participant's ratio of yes to no clicks compared to the means for the yes/no click ratio for the sample. We performed the first grouping based on the ratio of security warning disregard before the security incident, and we performed the second grouping on the ratio of security warning disregard after the security incident. We compared all of the demographic and control variables in each analysis.

First, we grouped the data based on their relation to the mean of security warning disregard before the security incident for the sample ($M = .749, SD = .388, N_{G1} = 19, N_{G2} = 43$). Table C-1 describes comparisons of control and survey research variables between the two groups. We found no statistically significant differences between the two groups.

Table C-1. Comparison of Yes-click and No-click Groups Before the Security Incident

Variable	Group 1 M (SD)	Group 2 M (SD)	*p*	Effect size (Cohen's d)
Age[1]	21.53 (2.038)	21.98 (1.933)	.529	0.229
Income risk sensitivity (scale of 1–4)[1]	2.74 (1.327)	2.60 (1.137)	.607	0.117
Perceived security risk of malware—pre-test[1]	15.11 (3.160)	13.86 (4.302)	.581	0.313
Threat severity of malware—pre-test[1]	10.83 (4.076)	11.79 (5.365)	.656	0.191
Threat susceptibility of malware—pre-test[1]	9.22 (3.370)	9.33 (4.492)	.759	0.026
General risk appetite—pre-test[3]	27.72 (4.496)	27.49 (6.017)	.883	0.041
Perceived security risk of malware—post-test[3]	15.22 (3.797)	14.63 (4.254)	.709	0.138
Threat severity of malware—post-test[1]	14.33 (3.985)	15.02 (4.662)	.266	0.154
Threat susceptibility of malware - post-test[3]	12.78 (4.066)	12.28 (4.506)	.704	0.114
Variable	**Group 1 proportions**	**Group 2 proportions**	***p*[2]**	**Effect size (ϕ)**
Gender	Male (68.4%)	Male (76.7%)	.538	.088
Left-handed	Yes (10.5%)	Yes (11.6%)	1.000	.016
Native English speaker	Yes (94.7%)	Yes (86.0%)	.422	.127
Normal vision	Yes (89.5%)	Yes (97.7%)	.220	.176
EEG experience	No (94.7%)	No (88.4%)	.657	.099
Mental condition	No (100%)	No (93.0%)	.546	.150
Colorblind	No (94.7%)	No (93.0%)	1.000	.032

[1] Due to a violation of the assumption of normality, this p-value was obtained using the Mann-Whitney U test. Cohen's d is included for ease of approximate interpretation of the effect size.
[2] p values taken from Fisher exact test.
[3] p values taken from t-test.

Next, we grouped the data based on their relation to the mean of security warning disregard after the security incident for the sample (M=.673,SD=.433,N_G1=22,N_G2=40). Table C-2 shows comparisons between the two groups' control and survey research variables. The only significant differences found were between the group proportions for native and non-native English speakers and in the comparison of group means for measures of threat susceptibility taken post-test. We also detected this difference during the main data analysis for the hypotheses and w included it as a control variable in H4's regression analyses. Likewise, we found that survey measures of threat susceptibility to malware taken post-test significantly predicted security warning disregard (see Table 5, H4c). Therefore, it is not surprising to see a significant difference between groups in post-test

threat susceptibility given that the groups were decided based on participants' security warning disregard ratios, which makes this examination, in a way, tautological.

Table C-2. Comparison of Yes-Click and No-Click Groups After the Security Incident

Variable	Group 1 M (SD)	Group 2 M (SD)	*p*	Effect size (Cohen's d)
Age[1]	21.50 (1.99)	22.03 (1.94)	.401	0.229
Income risk sensitivity (scale of 1–4)[1]	2.68 (1.21)	2.63 (1.19)	.843	0.117
Perceived security risk of malware—pre-test[3]	15.71 (2.795)	13.45 (4.356)	.148	0.582
Threat severity of malware—pre-test[1]	11.95 (4.213)	11.28 (5.411)	.502	0.133
Threat susceptibility of malware—pre-test[1]	9.67 (3.679)	9.10 (4.431)	.408	0.136
General risk appetite—pre-test[3]	28.48 (4.936)	27.08 (5.885)	.355	0.251
Perceived security risk of malware—post-test[1]	15.95 (3.232)	14.20 (4.410)	.148	0.433
Threat severity of malware—post-test[1]	15.57 (3.572)	14.43 (4.846)	.668	0.257
Threat susceptibility of malware—post-test[1]	14.19 (4.045)	11.50 (4.267)	.017*	0.642
Variable	**Group 1 proportions**	**Group 2 proportions**	***p*[2]**	**Effect size (ϕ)**
Gender	Male (63.6%)	Male (80.0%)	.226	.179
Left-handed	No (90.9%)	No (87.5%)	1.000	.052
Native English speaker	Yes (100%)	Yes (82.5%)	.044*	.265
Normal vision	Yes (95.5%)	Yes (95.0%)	1.000	.010
EEG experience	No (95.5%)	No (87.5%)	.409	.129
Mental condition	No (100%)	No (92.5%)	.546	.167
Colorblind	No (90.9%)	No (95.5%)	.610	.080

[1] Due to a violation of the assumption of normality, this *p*-value was obtained using the Mann-Whitney U test. Cohen's *d* is included for ease of approximate interpretation of the effect size.
[2] *p* values taken from Fisher exact test.
[3] *p* values taken from *t*-test.
* significant at α<.05.

Appendix D: Tests for Mediating and Moderating Effects

To better understand how the IGT behavioral data relates to the P300 and security warning disregard, we conducted an exploratory analysis of mediation (i.e., the behavioral data intervenes between the P300 amplitude and security warning disregard), and moderation (i.e., the behavioral data interacts with the P300 amplitude in influencing security warning disregard). In doing so, we used the ratio of choices from the most risky deck (Deck B) to choices from the safest deck (Deck C) for each participant during the IGT as our behavioral measure of performance.

First, we tested whether the behavioral IGT data mediated the effect of P300 on security warning disregard (see Figure D-1 below).

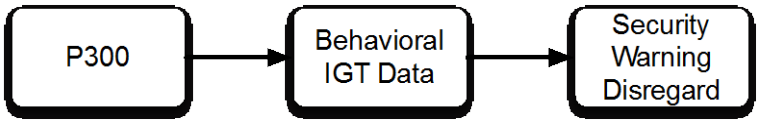


Figure D-1. Mediating Effect of Behavioral IGT Data

According to Baron and Kenny (1986), in order for mediation to occur, one must first demonstrate that the independent variable (P300) significantly predicts the mediating variable (behavioral IGT data), and that the mediating variable significantly predicts the dependent variable (security warning disregard). We tested these relationships utilizing regression, as presented in Table D-1 and D-2 below:

Table D-1. Behavioral IGT Data Regressed on the P300 Difference Score

Model	ß	Std. error	Standardized ß	*t*
Intercept	1.094	.105	—	10.378***
P300 difference score	-.007	.025	-.037	.782 ns

Model statistics: R^2 = .001; F = .078, ns
*** p < .001; ns = not significant

Table D-2. Security Warning Disregard (Before Security Incident) Regressed on Behavioral IGT Data

Model	ß	Std. error	Standardized ß	*t*
Intercept	.624	.090	—	6.971***
Behavioral IGT data	.096	.067	.184	1.427 ns

Model statistics: R^2 = .034; F = 2.037, ns
*** p < .001; ns = not significant

The results depicted in the above tables show that the P300 difference score did not predict the behavioral IGT data, nor did the behavioral IGT data predict security warning disregard (before the security incident)[6]. We therefore conclude that the behavioral data from the IGT did not mediate the effect of the P300 on security warning disregard in the image classification task.

Additionally, we tested whether the behavioral IGT data moderated the effect of P300 on security warning disregard (see Figure D-2 below).

[6] We obtained similar results when using security warning disregard (after security incident) as the dependent variable.

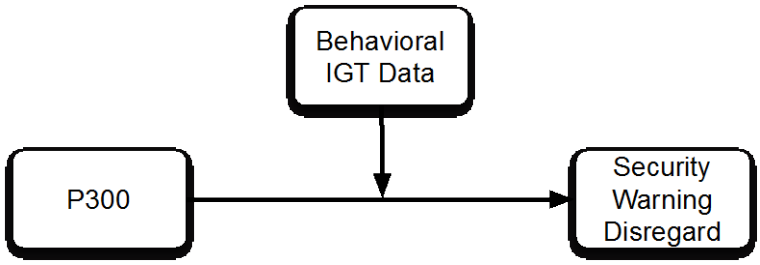


Figure D-2. Moderating Effect of Behavioral IGT Data

According to Carte and Russell (2003), in order for moderation to occur, one must first demonstrate that the product of the independent variable (P300) and the moderating variable (behavioral IGT data) significantly predicts the dependent variable (security warning disregard). We tested these relationships using hierarchical regression, with the interaction term in the second block. This allowed us to show the effect of the interaction over and above its individual components. These results are presented in Table D-3 below:

Table D-3. Security Warning Disregard (Before Security Incident) Regressed on the Interaction of Behavioral IGT Data and P300 Difference Score

Model 1	**ß**	**Std. error**	**Standardized ß**	***t***
Intercept	.664	.089	—	7.485***
P300 difference score	-.028	.013	-.272	-2.174*
Behavioral IGT data	.091	.065	.174	1.392 ns
Model statistics: R2 = .108; F = 3.446*				
Model 2	**ß**	**Std. error**	**Standardized ß**	***t***
Intercept	.663	.089	—	7.421***
P300 difference score	-.035	.020	-.348	-1.760 ns
Behavioral IGT data	.092	.066	.176	1.399 ns
Behavioral IGT data X P300 difference score	.008	.015	.098	.496 ns
Model statistics: R^2 = .112; F = 2.349 ns ΔR^2 = .004; F for ΔR^2 = .246 ns				

Because the interaction term was not significant, and because R^2 did not significantly increase when the interaction term was added to the model, we conclude that the behavioral IGT data did not moderate the effect of the P300 on security warning disregard (before security incident)[7].

Considering the above results, we conclude that the behavioral data in the IGT had no influence (either as a mediator or moderator) on the effect of the P300 difference score on security warning disregard in the image classification task. This is an example of overt behavior not being as good a predictor as neurophysiological measures (Kirwan, Shrager, & Squire, 2009).

[7] Again, the results were similar for using security warning disregard (after security incident) as the dependent variable.

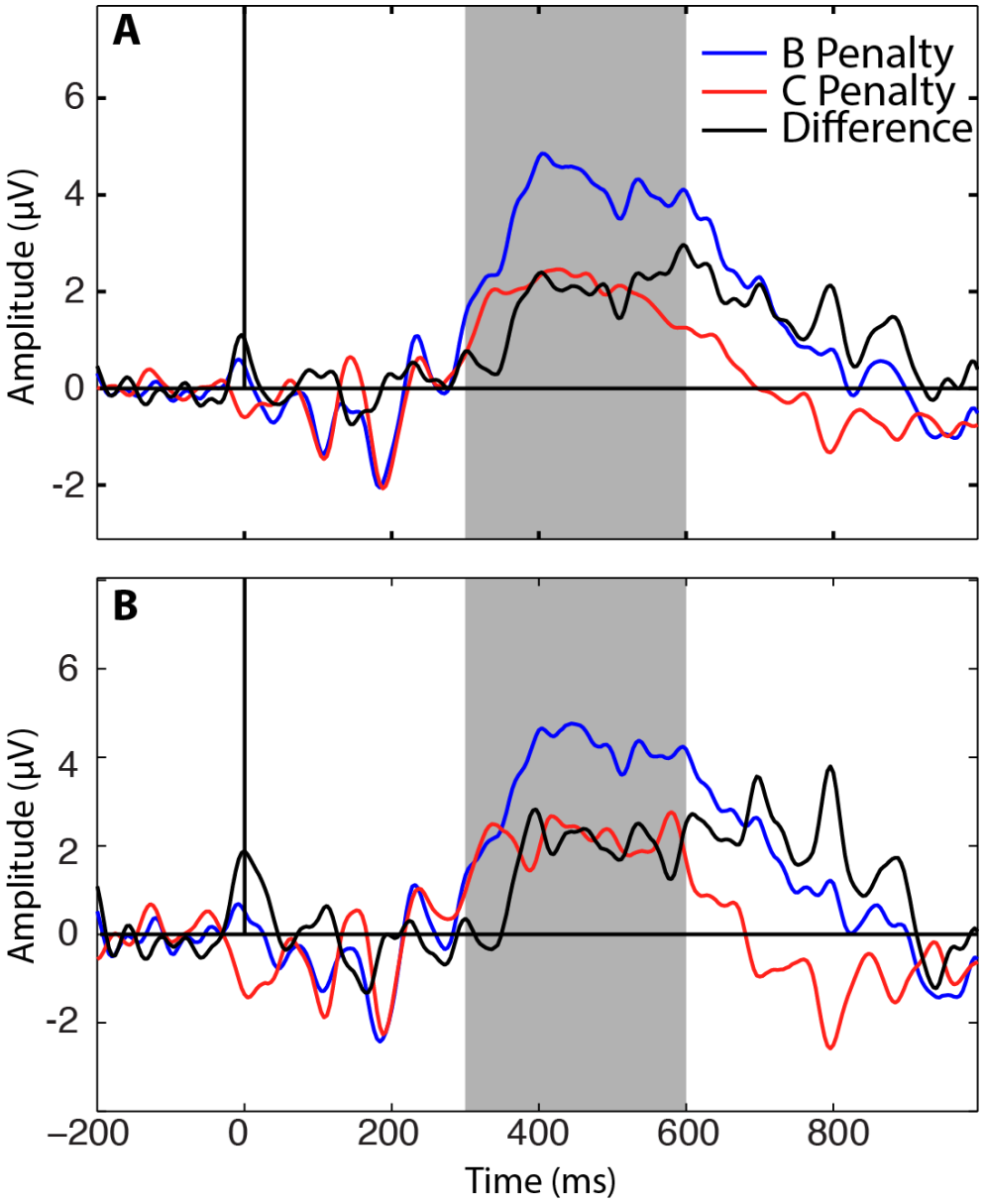
Appendix E: IGT Deck Selections

Table E-1. Participant IGT Deck Selections by Block of the Experiment

	Deck A				Deck B				Deck C				Deck D			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Block 1	19.2	6.5	10	37	30.5	7.4	9	40	20.2	8.1	4	40	30.1	8.6	14	40
Block 2	18.0	6.8	0	39	25.1	8.6	7	40	25.4	9.4	7	40	31.4	7.2	14	40
Block 3	17.7	8.4	4	40	19.7	10.4	3	40	30.2	10.5	6	40	32.4	8.8	11	40
Block 4	17.4	8.8	5	40	19.4	11.3	2	40	29.0	10.4	0	40	33.0	8.9	11	40
Total	72.4	23.2	32	129	94.7	29.9	53	160	104.9	29.1	38	160	126.9	24.5	65	160

Appendix F: Explanation of the 600-800ms Time Window

Most of the literature examining the P300 focuses on the time window 300-600ms following stimulus onset (San Martín et al., 2013), and we analyzed our data accordingly. Inspecting Figure 4 reveals a deflection from the baseline for the difference curve at the later 600-800ms time window. However, this appears to be largely due to the increased noise introduced by taking the average ERP waveform from a subset of randomly sampled trials. Inspection of an analysis that included all trials does not reveal this extreme deflection from 0 at the Pz electrode (see Figure F1A). To facilitate comparison, Figure 4 is reproduced below as Figure F-1B. We therefore conclude that the deflection from the baseline observed in Figure 4 is an artifact of noise in the resampling process, and not due to participants' actual neural activity.



Note that panel B is the same as Figure 4 from the manuscript.

Figure F-1. Mean Amplitudes When Including all Trials (A) and a Randomly Sampled Subset of Trials (B) PZ Electrode.

About the Authors

Bonnie Brinton ANDERSON is an Associate Professor of Information Systems and Director of the Master of Information Systems Management (MISM) program in the Marriott School of Management at Brigham Young University. She received her PhD from Carnegie Mellon University. Her work has been published in *Journal of the Association for Information Systems, Decision Support Systems, Electronic Commerce Research, Expert Systems with Applications; Electronic Commerce Research, Communications of the ACM, Information Sciences, IEEE Transactions: Systems, Men, and Cybernetics, The Journal of Systems and Software*, and other journals. She currently researches the intersection of Decision Neuroscience and Behavioral Information Security.

Anthony VANCE is as an Assistant Professor of Information Systems in the Marriott School of Management of Brigham Young University. He has earned PhD degrees in Information Systems from Georgia State University, USA; the University of Paris—Dauphine, France; and the University of Oulu, Finland. He received a BS in IS and Masters of Information Systems Management (MISM) from Brigham Young University, during which he was also enrolled in the IS PhD preparation program. His previous experience includes working as a visiting research professor in the Information Systems Security Research Center at the University of Oulu. He also worked as an information security consultant and fraud analyst for Deloitte. His work is published in *MIS Quarterly*, *Journal of Management Information Systems*, *Journal of the Association for Information Systems*, *European Journal of Information Systems*, *Journal of the American Society for Information Science and Technology*, *Information & Management*, and other journals. His research interests are behavioral information security and NeuroIS applications to security.

C. Brock KIRWAN is an assistant professor of Psychology and Neuroscience at Brigham Young Univ ersity. He received his PhD in Psychological and Brain Sciences from Johns Hopkins University in 20 06. Dr. Kirwan has a decade of experience conducting fMRI scans with patient populations at Johns Hopkins University, the University of California, San Diego, the University of Utah, and now BYU. He has published numerous papers reporting fMRI and neuropsychological results in journals such as *Science*, *Proceedings of the National Academy of Sciences*, *Neuron*, and the *Journal of Neuroscience*.

David EARGLE is a doctoral candidate in the Information Systems and Technology Management Area at the University of Pittsburgh in the Katz Graduate School of Business. He is currently a NSF Graduate Research Fellow. He completed a joint baccalaureate-master's program in information systems management at Brigham Young University, completing the IS PhD preparation program and graduating magna cum laude with University Honors. His research interests include human-computer interaction and information security. He has coauthored several articles in these areas using neurophysiological and other methodologies in outlets such as the *Journal of the Association for Information Systems*, the *International Conference on Information Systems*, and the *Hawaii International Conference on System Sciences*.