

How Polymorphic Warnings Reduce Habituation in the Brain—Insights from an fMRI Study

Bonnie Brinton Anderson,* C. Brock Kirwan,* Jeffrey L. Jenkins,* David Eargle,†
Seth Howard,‡ Anthony Vance*

*Brigham Young University
{bonnie_anderson, jeffrey_jenkins,
kirwan, anthony.vance}@byu.edu

†University of Pittsburgh
dave@daveeargle.com

‡Google Inc.
sethhoward@google.com

ABSTRACT

Research on security warnings consistently points to habituation as a key reason why users ignore security warnings. However, because habituation as a mental state is difficult to observe, previous research has examined habituation indirectly by observing its influence on security behaviors. This study addresses this gap by using functional magnetic resonance imaging (fMRI) to open the “black box” of the brain to observe habituation as it develops in response to security warnings. Our results show a dramatic drop in the visual processing centers of the brain after only the second exposure to a warning, with further decreases with subsequent exposures. To combat the problem of habituation, we designed a polymorphic warning that changes its appearance. We show in two separate experiments using fMRI and mouse cursor tracking that our polymorphic warning is substantially more resistant to habituation than conventional warnings. Together, our neurophysiological findings illustrate the considerable influence of human biology on users’ habituation to security warnings.

Author Keywords

Security warnings; habituation; functional magnetic resonance imaging (fMRI); mouse cursor tracking.

ACM Classification Keywords

H.5.2. Information interfaces and presentation: User Interfaces; J.4. Social and Behavioral Sciences.

INTRODUCTION

Warning messages are one of the last lines of defense in computer security and are fundamental to users’ security interactions with technology. Consequently, researchers have actively sought to understand how users interact with

security warnings and why warnings are so pervasively ignored [3]. A key contributor to the disregard of security warnings is *habituation*—i.e., the diminishing of attention because of frequent exposure to warnings [16]. Although habituation has been inferred as a factor in many security-warning studies [e.g., 24], little research has examined habituation in the context of security directly because habituation as a mental state is difficult to observe using conventional methods. Therefore, there is a gap in our understanding regarding how habituation to security warnings occurs in the brain, limiting researchers’ efforts to design warnings that can mitigate its effects.

This study addresses this gap by using functional magnetic resonance imaging (fMRI) to open the “black box” of the brain to observe habituation to security warnings. Specifically, we point to the *repetition suppression* (RS) effect, the reduction of neural responses to stimuli that are repeatedly viewed, as the neurological manifestation of habituation. By investigating how RS occurs in the brain, we can take a more precise approach to the design of security warnings that are resistant to the effects of habituation. Accordingly, the research objectives of this paper are: to (1) investigate how and where in the brain habituation develops in response to security warnings and (2) design a security warning that is substantially more resistant to habituation than the status quo.

We investigate these research objectives in two experiments. In Study 1, we designed a polymorphic warning that demonstrated substantially more resistance to habituation as compared to conventional warnings in the relevant brain regions. In Study 2, we use mouse tracking as a surrogate for attention to cross-validate that the polymorphic warning is more resistant to habituation in a more ecologically valid scenario.

BACKGROUND INFORMATION

Research frequently points to the role of habituation in the failure of warnings [e.g., 23]. For example, users click through 50% of SSL warnings in 1.7 seconds, a finding that “is consistent with the theory of warning fatigue” [1]. In another study, habituation was proposed to explain why only 14% of participants recognize changes in the content of confirmation dialogs [4]. These studies and others, however, do not directly measure habituation as a mental

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

Copyright © is held by the owner/author(s).
CHI 2015, Apr 18-23, 2015, Seoul, Republic of Korea
ACM 978-1-4503-3145-6/15/04.
<http://dx.doi.org/10.1145/2702123.2702322>

state; rather, they indirectly measure habituation by observing users' responses to warnings. Therefore, measuring the occurrence of habituation in the brain through neuroscience instruments (e.g., fMRI) presents an opportunity to quantify the antecedents, onset, and impacts of habituation on security-warning behavior.

Neuroscience allows researchers to open the "black box" of cognition by directly observing the brain. Using neuroscience methodologies holds promise for "providing a richer account of user cognition than that which is obtained from any other source, including the user himself" [20]. We define *neurosecurity* as applying neuroscience to behavioral information security to better understand and improve users' security behaviors. One ultimate goal of neurosecurity is to design more effective user interfaces (UIs) that can help users make informed decisions [26].

fMRI is a method of choice in neuroscience because of its superior ability to identify areas of the brain that are activated during decision making. fMRI measures neural activity by tracking changes in the level of blood oxygenation, which are driven by changes in the metabolic demands of active neural populations. This phenomenon is known as the *blood oxygen level dependent* (or BOLD) effect. By measuring changes in blood flow to different brain regions, researchers can identify distinct brain regions where activity is correlated with specific cognitive processes. Previous work has successfully used fMRI to examine cognitive processing of malware warnings [22]. Thus, fMRI is well suited to investigate how habituation to security warnings occurs in the brain.

Repetition Suppression: A Neuroscience View of Habituation

The neural underpinnings of habituation have been well studied in the context of simple behavior in model organisms. For example, a series of now-classic studies using sea slugs demonstrated that neural responses to a given stimulus decrease with repeated exposures to that stimulus [17]. This kind of RS to repeated stimuli has also been observed in humans across a variety of brain regions and experimental situations [for review, see 13].

The exact cause and purpose of RS may vary according to brain region. In the visual processing system, RS appears to be related to repetition priming and may reflect the facilitation of stimulus processing with repeated exposures [9]. This type of response is highly influenced by the similarity of the repeated stimuli, where a more robust RS is associated with more similar stimuli [18].

Moreover, research has found that neural measures can be more sensitive than behavioral measures to RS effects [21]. Therefore, the RS effect is an effective way to assess the underlying neural computational processes of behavior. In this study, we utilized the differential RS effect in various brain regions to map the sensitivity of those regions to stimulus repetition of security warnings.

Our objective is to design a security warning implementation that is substantially more resistant to habituation than the status quo. Several approaches have been suggested to increase attention to security warnings and thus reduce habituation. For example, following design guidelines found in the warning literature, researchers have designed and tested a variety of warning attractors [e.g., 2]. These studies demonstrated that habituation can be reduced through UI design.

Wogalter states, "habituation can occur even with well-designed warnings, but better designed warnings with salient features can slow the habituation process. Where feasible, changing the warning's appearance may be useful in reinvigorating attention switch previously lost because of habituation" [28]. Accordingly, polymorphic warnings—i.e., warnings that change their appearance to reduce habituation—have been suggested as a way to reduce habituation. For example, [5] found that study participants who received software warning dialogs with randomized options opened risky e-mail attachments significantly less than those who received conventional warnings.

Despite this finding, it is not clear whether polymorphic warnings actually mitigate habituation, as past studies have not measured habituation directly. Furthermore, research has only tested polymorphic warnings that randomize options in dialog boxes, acknowledging that "the design space for polymorphic dialogs is vast" [5]. It is therefore unknown which types of polymorphic warnings are most effective at reducing the effects of habituation. Thus, our research objectives are to (1) use fMRI to determine whether a polymorphic warning design is able to reduce RS in the brain in response to repeated exposure to security warnings and (2) identify the polymorphic variations that are most effective in reducing the RS effect.

Hypothesis

Polymorphic warnings garner more attention over time due to the novelty of their changing appearance. Changing appearance of the warning reinvigorates attention [28], especially in brain regions that have been shown to demonstrate RS to exact repetitions of visual stimuli [e.g., ventral visual processing stream; 27]. For this reason, polymorphic warnings that continually change their appearance will slow the rate of habituation. Accordingly, we hypothesize

For polymorphic warnings, brain regions associated with visual processing, such as the occipital lobe, will have a higher BOLD response upon successive viewings as compared with static warnings.

STUDY 1: FMRI

We conducted an fMRI laboratory experiment to test our hypothesis following the guidelines of Dimoka [10]. fMRI was selected because of its high spatial resolution, which makes it an ideal choice for determining brain regions involved in specific cognitive functions. fMRI uses the same equipment as clinical MRIs and is equally

noninvasive; it also measures activity in the whole brain in a grid comprised of three-dimensional pixels (or *voxels*) that are 3 mm to a side, approximately every 2 seconds.

In order to test our hypothesis, we first developed a polymorphic warning implementation. To do so, we relied on the warning science literature [e.g., 28] to develop 12 graphical variations to a warning dialog expected to capture attention. Our polymorphic warning artifact rotated through the graphical variations on each subsequent exposure. The variations are displayed in Figure 3.

fMRI Procedures

Participants were given a verbal briefing about the MRI procedures and the task. Participants laid down on their backs on the fMRI table and had a volume coil placed over their heads to allow imaging of the entire brain (see Figure 1). They were then moved via the fMRI table into the MRI scanner. Participants viewed the experimental images on a large monitor at the opening of the MRI scanner by use of a mirror placed over their face. The monitor was configured to display images in reverse so that they appeared normal when viewed through the mirror. The technical details of the fMRI scans, experimental procedures, and data analyses are documented in the appendix.



Figure 1. The fMRI scanner. A volume coil (with mirror to view the computer monitor) is placed over the participant's head.

Experimental Design

We created a unique set of 560 images for each participant. Forty images of general software applications (such as Microsoft Word or Mozilla Firefox) were displayed one time each. These images created a baseline of unique presentations throughout the task that served as our fMRI contrast. Of a pool of 40 warning images, 20 were selected randomly by computer algorithm to receive the polymorphic treatment and the other 20 were given the static treatment. The polymorphic warnings were shown one time each per variation (for a total of 13 unique images per warning), and the same static warnings were shown 13 times. The 560 images were randomized for each participant across five blocks of 6.5 minutes each (with a 2-minute break in between) in order to relieve participants' fatigue during the task. Participants were shown images for 3 seconds each with a 0.5-second inter-stimulus interval.

Experimental Task

Before being placed inside the fMRI scanner, participants were given an fMRI button box (see Figure 2) to indicate if each image shown was (1) identical to one seen before the task, (2) similar yet different from one seen previously in the task, or (3) new, never seen before in the task.

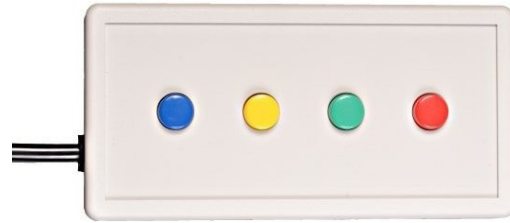


Figure 2. fMRI button box.

This process was used to provide behavioral performance data that ensured that participants were appropriately engaged in the task. Our subsequent analysis of the responses (“old”, “similar”, and “new”) to each stimulus type (novel, static, polymorphic manipulations) revealed that participants performed the task as expected. At the conclusion of the fMRI scan, participants were led to an adjoining room to complete a brief posttest survey. To ensure manipulation validity, the posttest survey included a manipulation check asking if the participants noticed the variations in the images. All but one of the participants reported that they noticed the experimental treatment, indicating that they were on the whole successfully manipulated.

Participants

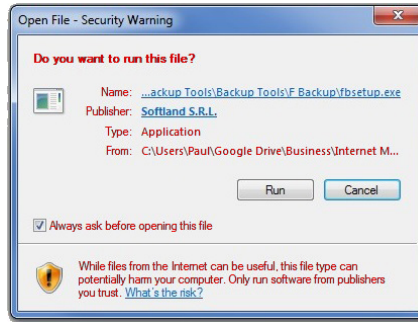
Twenty-five participants were recruited from the university community. We screened participants to require MRI compatibility, native English speakers, corrected-normal visual acuity, and right-handedness and excluded those with color-blindness or who were taking psychotropic medications. All participants were given an informed consent form to sign in accordance with the university's institutional review board (IRB) protocol. Of the 25 participants, 21 were male and 4 were female. Participant age ranged from 20–27 years of age with a mean age of 23.68 years. Participants were paid \$25 for approximately 1 hour in the scanner.

Data Analysis

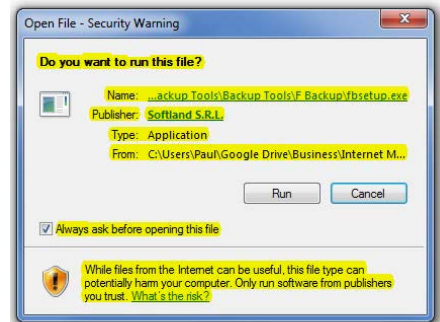
We first performed an ANOVA on the fMRI data with stimulus type (polymorphic, static) and repetition number as factors. We found four regions where there was a stimulus type \times repetition number interaction, indicating that these regions reacted differently to repeated exposures to static and polymorphic warnings. Two regions, the left and right superior parietal cortex (Figure 4, upper panels) have been implicated in attentional processing [7]. In these regions, activation was higher for polymorphic than for static warnings, consistent with sustained attention and reduced repetition suppression for the polymorphic warnings. The other two regions, bilateral medial prefrontal cortex and the left retrosplenial cortex (Figure 4, lower



(a) Original Warning Screenshot



(b) Color of Text Variation



(c) Highlight of Text Variation



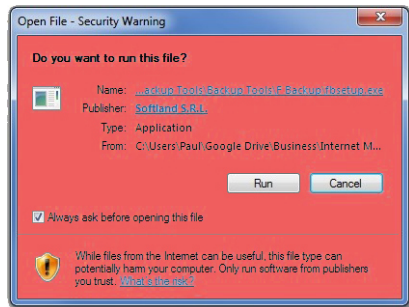
(d) Signal Word Variation



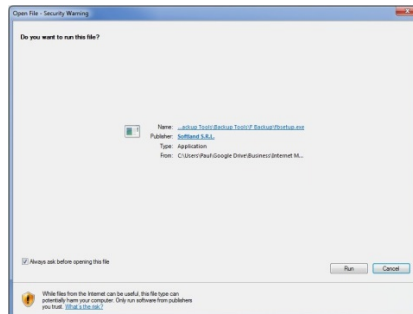
(e) Pictorial Signals Variation



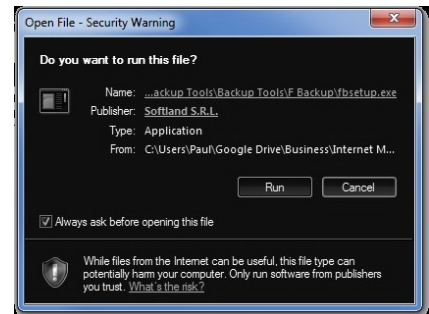
(f) Ordering of Options Variation



(g) Color Variation



(h) Size Variation (3X Larger)



(i) Contrast Variation



(j) Border Variation

Figure 3. Polymorphic warning design variants.

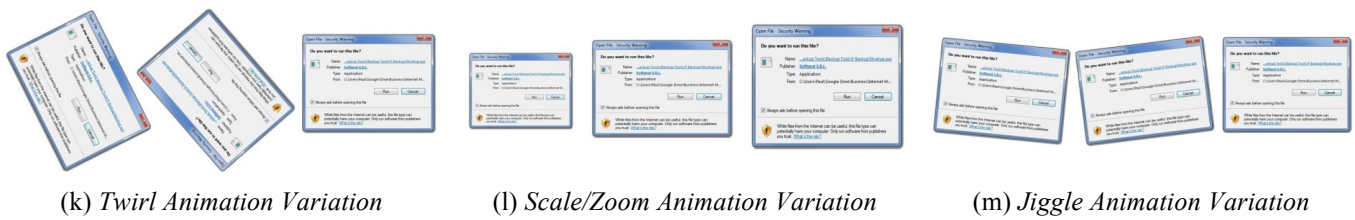


Figure 3 (continued). Polymorphic warning design variants.

panels), have both been implicated in memory retrieval processes and as nodes in the default mode network (DMN) [6]. In both regions, activation was higher for static than for polymorphic images for later presentations, consistent with either increased recognition memory processes or increased default mode activation for static but not for polymorphic warning images. Thus, our hypothesis was supported.

We also examined the specific polymorphic warnings collapsing across presentation order to determine which polymorphic variations were the most resistant to the RS effect. Accordingly, we examined the specific polymorphic warnings collapsing across presentation order to determine which polymorphic variations were most resistant to the RS effect. To quantify the most resistant variations for use in Study 2, we performed a *t*-test contrasting polymorphic and

static warning images. Resulting statistical parameter maps were thresholded using a false-discovery rate of .05 and a spatial extent threshold of 1,080 mm³. The largest cluster of activation comprised bilateral visual processing streams (including occipital, dorsal parietal, and inferior temporal lobes). Mean fMRI activations in this cluster for each polymorphic variation were then compared against the mean activation for static images (collapsed across repetitions) and ranked. Figure 5 depicts the mean beta fMRI activation for each polymorphic variation. Each manipulation had significantly greater fMRI activation than static images (Bonferroni-corrected *p*-value = .0038) with the exception of the “border” manipulation (*p* = .009). The highest scoring variations (“jiggle”, “scale”, “window color”, and “symbol”) were selected for Study 2.

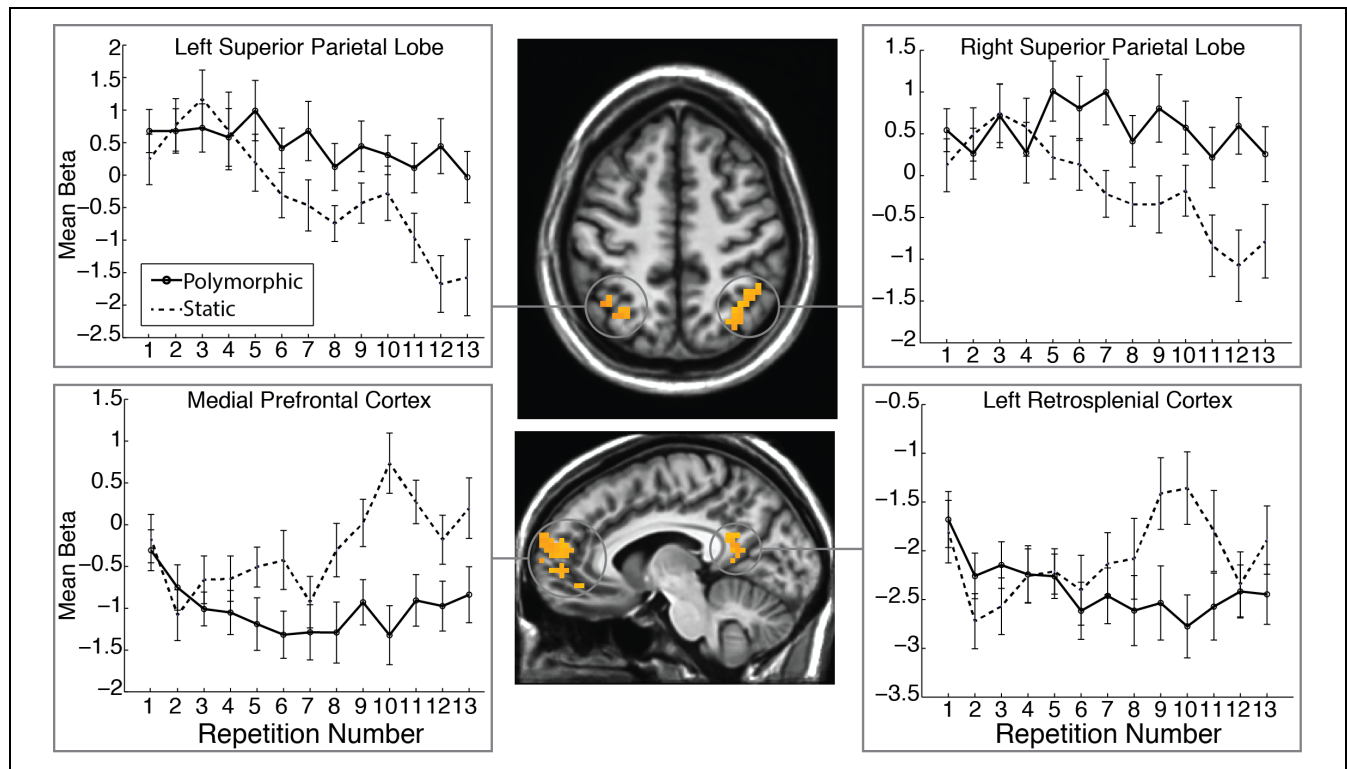


Figure 4. Brain regions that had a significant stimulus type × repetition number interaction. In the upper panels, the superior parietal lobes had elevated levels of activation for the polymorphic warnings, consistent with sustained attention. In the lower panels, the medial prefrontal cortex and the retrosplenial cortex had greater activation for static than for polymorphic images, consistent with memory retrieval.

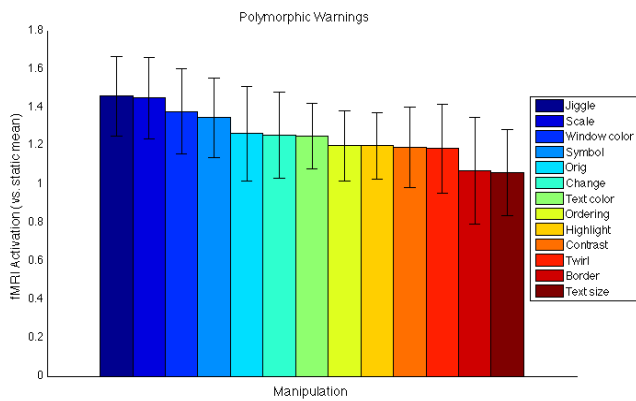


Figure 5. Mean fMRI activation for each polymorphic warning variation. Zero on the Y-axis indicates mean activation for static images (collapsed across repetitions). Error bars represent standard error of the mean.

STUDY 2: MOUSE CURSOR TRACKING

The results of Study 1 demonstrated that polymorphic warnings substantially reduced the occurrence of RS compared with conventional static warnings. However, constraints inherent in the fMRI method necessarily introduced artificiality into Study 1 that limited its ecological validity (e.g., participants must lie still in a supine position, view images rather than naturally interact with a computer). To address this limitation, we designed a behavioral laboratory experiment in which participants responded to security warnings in the natural course of completing a task on their personal laptops. To measure attention given to the warning, we analyzed users' mouse cursor movement (i.e., how a person responded using the laptop touchpad) through an embedded JavaScript library, which collected the x, y coordinates and timestamp of each movement at a millisecond precision rate and sent them to a web service for post hoc analysis. Although there is no causal link between the neural activation in Study 1 and the mouse tracking behavior in Study 2, both neurophysiological methods serve to corroborate each other.

We selected Google Chrome extension permission warnings as the warning type. During the course of the experiment, participants were required to search for, install, and evaluate 20 weather extensions for usability and aesthetics. Because a permission warning is displayed whenever a Google extension is installed, participants also received 20 permission warnings. In this way, they were naturally habituated while completing the task.

Experimental Design

Two experimental manipulations were included in the experiment. First, participants were randomly assigned the polymorphic warning treatment or the static warning treatment (a between-subject manipulation). The polymorphic treatment group randomly utilized the four most effective variations identified in Study 1 (Figure 5). The control group received conventional warnings that did not change their appearance. Second, the content of the

warnings was randomly manipulated between reasonable and unreasonable permission requests (a within-subject manipulation) to motivate participants to read the warnings.

Ethics

The university IRB approved a deception protocol in which we led participants to believe that their objective was to evaluate Google Chrome extensions. Additionally, we received permission to perform a man-in-the-middle attack to spoof Google.com search results in order to ensure that only our experimental Chrome extensions were installed. After the experiment, participants were debriefed and told the true purpose of the experiment and assured that the extensions did not gather any personal information.

Experimental Task

To heighten perceived risk, participants were asked to use their personal laptop and were required to read the following disclaimer: "The researchers have not tested all possible weather extensions that you may encounter. The university is not responsible for any potential malicious software that may be installed as a result of this study. Each participant has the responsibility to use good judgment in deciding which extensions to install and evaluate."

Next, participants were asked to use Chrome and search Google.com to find 20 different weather extensions and evaluate them usability and aesthetics. However, as previously stated, we performed a man-in-the-middle attack to manipulate the Google search results. If a query related to Chrome weather extensions was shown, the search returned our spoofed results. Each of the links in the spoofed results pointed to sites under our control with legitimate URLs. Our Chrome extension transparently redirected HTTPS traffic to HTTP to avoid SSL certificate errors. Finally, we caused traffic to the Chrome Web Store to redirect to a "down for maintenance" page. Consequently, participants were required to use the in-line Chrome extension installation process on individual web sites.

The manipulated Google search results included more than 40 different weather extensions. Participants were only required to evaluate 20 of these, and only 3 of these at random presented an unreasonable permission warning. If a participant received a permission warning requesting unreasonable permissions, the participant was free to not install the extension. Participants had ample time to go back to the Google search results to find an alternative extension.

Participants

Eighty students from the university community participated in the study. Technical limitations on a few participants' laptops inhibited data collection of the mousing data. As a result, we had analyzable data from 76 participants across 1,466 warnings (40 in the treatment group and 36 in the control group). Approximately 63% of the participants were male, and the average age was 22.35 years. The participants were paid \$7.50 for the 45-minute task.

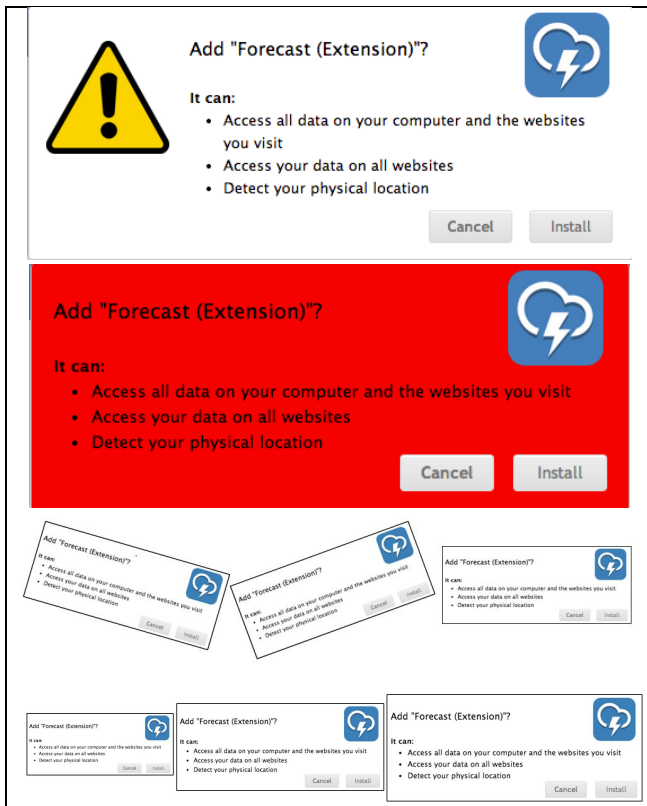


Figure 6. Pictorial, red, jiggle, and zoom variations.

Data Analysis

We analyzed users' mouse cursor movements to measure attention given to the warning, and thereby habituation. Research has repeatedly shown that the mouse cursor denotes where users' devote their attention, following the eye when navigating a webpage [14]. Attention given to information (e.g., warning information) can be inferred through a collection of mouse cursor movement statistics. For example, area-under-the-curve (AUC) is often used to denote attention [12]. AUC is the geometric area between a user's movement trajectory and a straight line connecting the user's starting point when a warning is shown and the final destination point (e.g., the button the user clicks to dismiss the warning) [see 12 for calculations]. When information from a warning catches users' attention, it causes them to deviate from this straight line as the cursor moves toward the warning information. AUC is a measure of total deviation from the idealized response trajectory or, in other words, total attention given to other information [12]. Other indicators of attention include people moving greater distance as they explore the warning, initially accelerating more slowly as they process information, changing directions more often while exploring a warning, moving more slowly indicating increased cognitive processing, and having greater click latency indicating increased cognitive processing. Although inferring attention through one statistic alone can be challenging, interpreting all the statistics together provides strong support of attentional changes.

To explore the influence of polymorphic warnings on mouse cursor movement indicators of attention and thereby habituation, we constructed a linear mixed-effects model, accounting for the repeated nature of the data nested within each participant. To model habituation across multiple warning exposures, we included a fixed effect of the warning order number in the model. The participant identifier was included as a random effect to allow variability across participants. The model included a fixed effect for whether participants were in the treatment group. In addition, the model contained a fixed effect for whether the warning cautioned that the add-on was requesting reasonable information or unreasonable information (coded as *isMalicious*). As participants were allowed to naturally complete the task for ecological validity, they could potentially revisit websites, reinstall extensions, and see the warning for those extensions again. Hence, a fixed effect was also included regarding whether the warning was being reinstalled. The model included interaction terms for *treatment × Order* and *isMalicious × Order*. AUC, initial acceleration, direction changes, movement speed, and click latency were included as dependent variables.

The results are shown in Table 1 and support the results of the fMRI study. The warning order number was significantly correlated with all of the mousing statistics in the predicted directions, suggesting that habituation increased with subsequent exposures (i.e., with warning order number). The use of polymorphic warnings was also significantly correlated with all of the mousing statistics in the predicted directions, suggesting that people habituate less when polymorphic warnings are shown.

Mousing Statistic	treatment	isMalicious	order	isReinstalled	treatment × order	isMalicious × order
AUC	60.53 ***	0.87 ns	902.71 ***	71.68 ***	156.81 ***	0.71 ns
Initial Acc.	9.81 ***	0.26 ns	110.26 ***	42.53 ***	12.46 ***	0.48 ns
Direction Changes	112.50 ***	0.20 ns	10.90 ***	15.08 ***	0.04 ns	8.55 **
Speed	16.06 ***	0.47 ns	155.58 ***	23.55 ***	8.62 **	0.31 ns
Distance	21.70 ***	1.15 ns	17.45 ***	118.83 ***	3.94 *	0.88 ns
Click Latency	85.26 ***	1.29 ns	8.26 **	35.84 ***	0.21 ns	0.89 ns

Chi-squared shown in cells, df = 1 for all tests
* p < .05, ** p < .01, ***p < .001, ns = not significant

Table 1: Study 2 Chi-Squared Test Results

In addition, the results suggest that people habituate more slowly when polymorphic warnings were shown (the interaction between treatment and order was significantly correlated with all of the statistics except for click latency).

Three additional variables were included in the model. First, we explored whether the content of the warning influenced habituation (isMalicious). Interestingly, the content of the warning did not influence habituation for any of the statistics. This finding suggests that changing the content of a warning may not be enough to deter the influence of habituation. Likewise, the interaction between isMalicious and Order of the warning did not significantly influence any of the statistics, suggesting that the content of the message may not decrease habituation even during early encounters. Finally, if an extension was being reinstalled, it influenced all of the statistics.

To further elaborate on how polymorphic warnings influence attention in terms of users' mouse cursor movements, Table 2 displays the population means, standard deviations, and effect sizes of each mousing statistic for the static warning treatment group and the polymorphic warning treatment group.

GENERAL DISCUSSION

This paper makes several contributions. First, in Study 1, we extended previous research on habituation by using neuroscience to observe habituation as it occurs in the context of human-computer interaction (HCI). Additionally, this study demonstrated how the RS effect can be directly measured in the brain using fMRI. Whereas previous HCI research measured habituation indirectly by observing its effects, such as inattentive behaviors [4], this study measured habituation directly as it occurs in the brain. Specifically, we showed how using a simple, repeated exposure, experimental design can permit researchers to detect the existence and size of the RS effect using the BOLD response. Using this method, we illustrated the precipitous drop in visual processing after only one repeated exposure and a large overall drop after 13 exposures. These results can provide researchers with a useful baseline of the RS effect in response to security warnings for future research.

Second, we demonstrated in Study 1 that polymorphic warnings are more resistant to RS than static warnings. Although polymorphic warnings have been proposed before [5], previous polymorphic variations were limited to the repositioning of options on the warning dialog. In contrast, this study utilized warning science literature to derive 12 polymorphic variations that can be generically applied to a wide variety of security warnings (see Figure 3).

Additionally, past research has not examined the influence of polymorphic warnings on habituation (either directly or indirectly) but rather measured frequency of risk-taking behavior. Therefore, it was previously unknown whether polymorphic warnings could indeed reduce habituation. We

addressed this research gap by establishing that polymorphic warnings reduce the occurrence of RS to be significantly below that of conventional static warnings.

Third, we analyzed the fMRI data to select the polymorphic variations that were most effective in reducing RS. In this way we were able to identify the four most effective variations (see Figure 5) from our set.

Fourth, the results of the mouse cursor tracking experiment corroborated our fMRI findings. Specifically, we found that that polymorphic warnings resulted in both *reduced* habituation (as evidenced by the main effect) and *slower* habituation (as indicated by the interaction effect of the polymorphic treatment and the warning display order) compared with static warnings. According to Dimoka et al., "no single neurophysiological measure is usually sufficient on its own, and it is advisable to use many data sources to triangulate across measures" [11]. Using two complimentary neurophysiological measures allowed us to compensate for weaknesses inherent in each method. For example, the mouse cursor tracking method allowed us to implement the polymorphic warning UI artifact and test it within a natural task on participants' own personal computers, significantly enhancing the ecological validity of this research.

Mousing Statistic	Static Warnings mean (sd)	Polymorphic Warnings mean (sd)	Difference Effect Size Cohen's D
AUC	0.09 (0.63)	1.03 (0.65)	1.47***
Initial Acc.	10.10 (1.06)	9.31 (1.33)	0.66**
Direction Changes	4.52 (4.48)	5.29 (8.18)	0.17*
Speed	3.19 (0.79)	2.65 (0.73)	0.71**
Distance	0.40 (0.51)	0.66 (0.82)	0.37*
Click Latency	124.90 (73.50)	152.71 (239.19)	0.16*
* Small effect size <.5, ** Medium effect size <.8, *** Large effect size >.8			

Table 2: Population descriptive statistics and effect sizes of polymorphic treatment

Finally our methodology is itself a contribution in that it illustrates the usefulness of applying neuroscience to the domain of security and HCI generally. Because automatic or unconscious mental processes underlie much of human cognition and decision making, they likely play an important role in a number of other security behaviors, such as security education, training, and awareness (SETA) programs, password use, and information security policy compliance. Additionally, neuroscience methods have the potential to lead to the development of more complete behavioral security theories and guide the design of more effective security interventions.

LIMITATIONS AND FUTURE RESEARCH

This research is subject to a number of limitations. First, both studies used laboratory experiments that necessarily introduced artificiality into their respective tasks. In particular, the current technology of fMRI limits the flexibility and realism of experimental tasks. Fortunately, these problems in the fMRI study were at least partially compensated for by the mouse cursor tracking study in which participants conducted a much more ecologically valid task. However, we leave to future research the application of field methodologies that can achieve greater levels of external validity.

Second, although we presented strong evidence that the polymorphic warning reduces habituation behavior, we did not demonstrate that it leads to lower click-through behavior [26]. This was due to our chosen context of permission warnings and our associated experimental design. For some security warnings, inappropriate click-through behavior can be determined objectively. For example, in the case of web browser malware warnings, the false positive rate is extremely low [1]. Thus, users' choice to ignore these warnings can be objectively labeled as inappropriate. In contrast, in the context of privacy permission warnings, labeling inappropriate click-through behavior is difficult because privacy concerns vary by individual and are subjective [19]. Therefore, clicking through a permission warning for one person may be due to carelessness, whereas for another the same choice may be the result of considered deliberation. For this reason, we selected the mouse cursor tracking measures as our dependent variables in Study 2 because they directly and objectively measure habituation—the focus of this study. We leave to future research to investigate whether polymorphic warnings reduce click-through behavior in other contexts.

CONCLUSION

Users' habituation to security warnings is pervasive, and is often attributed to users' carelessness and inattention [15]. However, we demonstrate that habituation is largely obligatory as a result of how the brain processes familiar visual stimuli. A chief implication of our results is that because habituation occurs unconsciously at the neurobiological level, interventions designed to encourage greater attention and vigilance on the part of users—such as SETA programs—are incomplete on their own. Our findings suggest that a complementary solution is to develop UI designs that are less susceptible to habituation. We show that the polymorphic warning artifact developed in this study is one such effective design. Our results point to future research opportunities for security interventions that take into account the biology of the user.

ACKNOWLEDGMENTS

We thank Lorrie Faith Cranor for her constructive comments on this manuscript. This work was supported by the National Science Foundation under Grants CNS-1422831 and DGE-1247842; the Rollins Center for

Entrepreneurship and Technology at Brigham Young University; and the Berg Center for Ethics and Leadership at the University of Pittsburgh.

APPENDIX: FMRI TECHNICAL DETAILS

Equipment: MRI scanning took place at a university MRI research facility with the use of a Siemens 3T Tim-Trio scanner. For each scanned participant, we collected a high-resolution structural MRI scan for functional localization in addition to a series of functional scans to track brain activity during the performance of the various tasks. Structural images were acquired with a T1-weighted magnetization-prepared rapid acquisition including a gradient-echo (MP-RAGE) sequence with the following parameters: TE = 2.26 ms, flip angle = 9°, slices = 176, slice thickness = 1.0 mm, matrix size = 256 × 215, and voxel size = 1 mm × 0.98 mm × 0.98 mm. Functional scans were acquired with a gradient-echo, echo-planar, T2*-weighted pulse sequence with the following parameters: TR = 2000 ms, TE = 28 ms, flip angle = 90°, slices = 40, slice thickness = 4.0 mm (no skip), matrix size = 64 × 64, and voxel size = 3.44 mm × 3.44 mm × 3 mm.

Protocol: We first performed a 10-second localizer scan, followed by a 7-minute structural scan. Following these, we started the experimental task. We used E-Prime software to display the stimuli and synchronize the display events and scanner software. Total time in the scanner was 55 minutes.

Analysis: MRI data were analyzed with the Analysis of Functional Images (AFNI) suite of programs [8]. Briefly, functional data were slice-time corrected to account for differences in acquisition time for different slices of each volume; then, each volume was registered with the middle volume of each run to account for low-frequency motion. Data from each run were aligned to the run nearest in time to the acquisition of the structural scan. The structural scan was then co-registered to the functional scans. As in previous studies [e.g., 21], spatial normalization was accomplished by first warping the structural scan to the Talairach atlas [25] followed by warping to a template brain with Advanced Neuroimaging Tools (ANTs). Behavioral vectors were created that coded for stimulus type (e.g., security warnings, general software application screenshots) and repetition number. These were then entered into single-participant regression analyses. Stimulus events were modeled using a stick function convolved with the canonical hemodynamic response. Regressors coding for motion and scanner drift were also entered into the model as nuisance variables. Resulting beta values were blurred with a 5-mm FWHM Gaussian kernel. Beta values for the conditions of interest were then entered into group-level analyses, such as ANOVAs or *t*-tests, which were used to determine functional ROIs. All tests were corrected for multiple comparisons using a false-discovery rate of 0.05 and a spatial extent threshold of 20 contiguous voxels (540 mm³). Finally, the single-presentation general computing screenshots served as the implicit baselines.

REFERENCES

1. Akhawe, D. and Felt, A.P. Alice in warningland: A large-scale field study of browser security warning effectiveness. *Proc. USENIX Security '13*, USENIX Association (2013), 257-272.
2. Bravo-Lillo, C., Cranor, L., Komanduri, S., Schechter, S., and Sleeper, M. Harder to ignore? Revisiting pop-up fatigue and approaches to prevent it. In *Proc. SOUPS 2014*, USENIX Association (2014), 105-111.
3. Bravo-Lillo, C., Cranor, L.F., Downs, J., Komanduri, S., and Sleeper, M. Improving computer security dialogs. *Proc. INTERACT 2011*, Springer-Verlag (2011), 18-35.
4. Bravo-Lillo, C., Komanduri, S., Cranor, L.F., Reeder, R.W., Sleeper, M., Downs, J., and Schechter, S. Your attention please: Designing security-decision UIs to make genuine risks harder to ignore. In *Proc. SOUPS 2013*, ACM (2013), 1-12.
5. Brustoloni, J.C. and Villamarín-Salomón, R. Improving security decisions with polymorphic and audited dialogs. In *Proc. SOUPS '07*, ACM (2007), 76-85.
6. Buckner, R.L., Andrews-Hanna, J.R., and Schacter, D.L. The brain's default network. *Ann. N. Y. Acad. Sci.* 1124, 1 (2008), 1-38.
7. Buschman, T.J. and Miller, E.K. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315, 5820 (2007), 1860-1862.
8. Cox, R.W. Afni: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 3 (1996), 162-173.
9. Desimone, R. Neural mechanisms for visual memory and their role in attention. *Proc. Natl. Acad. Sci.* 93, 24 (1996), 13494-13499.
10. Dimoka, A. How to conduct a functional magnetic resonance (fMRI) study in social science research. *MIS Quart.* 36, 3 (2012), 811-840.
11. Dimoka, A., Banker, R.D., Benbasat, I., Davis, F.D., Dennis, A.R., Gefen, D., Gupta, A., Lschebeck, A., Kenning, P.H., Pavlou, P.A., Muller-Putz, G., Riedl, R., Vom Brocke, J., and Weber, B. On the use of neurophysiological tools in IS research: Developing a research agenda for NeuroIS. *MIS Quart.* 36, 3 (2012), 679-702.
12. Freeman, J.B. and Ambady, N. Mousetracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behav. Res. Methods* 42, 1 (2010), 226-241.
13. Grill-Spector, K., Henson, R., and Martin, A. Repetition and the brain: Neural models of stimulus-specific effects. *Trends Cogn. Sci.* 10, 1 (2006), 14-23.
14. Guo, Q. and Agichtein, E. Towards predicting web searcher gaze position from mouse movements. In *Proc. CHI'10*, ACM (2010), 3601-3606.
15. Herley, C. So long, and no thanks for the externalities: The rational rejection of security advice by users. In *Proc. NSPW 2009*, ACM (2009), 133-144.
16. Kalsher, M. and Williams, K. Behavioral compliance: Theory, methodology, and result. In *Handbook of warnings* Lawrence Erlbaum Associates, Mahwah NJ (2006), 313-331, 2006.
17. Kandel, E.R. The molecular biology of memory storage: A dialogue between genes and synapses. *Science* 294, 5544 (2001), 1030-1038.
18. Lueschow, A., Miller, E.K., and Desimone, R. Inferior temporal mechanisms for invariant object recognition. *Cereb. Cortex* 4, 5 (1994), 523-531.
19. Malhotra, N.K., Kim, S.S., and Agarwal, J. Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Inf. Syst. Res.* 15, 4 (2004), 336-355.
20. Minnery, B.S. and Fine, M.S. Neuroscience and the future of human-computer interaction. *Interactions* 16, 2 (2009), 70-75.
21. Motley, S.E. and Kirwan, C.B. A parametric investigation of pattern separation processes in the medial temporal lobe. *J. Neurosci.* 32, 38 (2012), 13076-13084.
22. Neupane, A., Saxena, N., Kuruvilla, K., Georgescu, M., and Kana, R. Neural signatures of user-centered security: An fMRI study of phishing, and malware warnings. *Proc. NDSS*, (2014), 1-16.
23. Schechter, S.E., Dhamija, R., Ozment, A., and Fischer, I. The emperor's new security indicators. In *Proc. SP'07*, IEEE (2007), 51-65.
24. Sunshine, J., Egelman, S., Almuhimedi, H., Atri, N., and Cranor, L.F. Crying wolf: An empirical study of SSL warning effectiveness. In *Proc. SSYM'09*, (2009), 399-416.
25. Talairach, J. and Tournoux, P. *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: An approach to cerebral imaging*. Thieme, Stuttgart, 1988.
26. Vance, A., Anderson, B.B., Kirwan, C.B., and Eargle, D. Using measures of risk perception to predict information security behavior: Insights from electroencephalography (EEG). *J. Assoc. Inf. Syst.* 15, 10 (2014), 679-722.
27. Vuilleumier, P., Henson, R., Driver, J., and Dolan, R. Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nat. Neurosci.* 5, 5 (2002), 491-499.
28. Wogalter, M.S. Communication-human information processing (C-HIP) model. In *Handbook of warnings*, M.S. WOGALTER Ed. Lawrence Erlbaum Associates, Mahwah, NJ (2006), 51-61, 2006.