Association for Information Systems
AIS Electronic Library (AISeL)

WISP 2012 Proceedings

Pre-ICIS Workshop on Information Security and Privacy (SIGSEC)

Winter 12-14-2013

Is Your Susceptibility to Phishing Dependent on Your Memory?

Bonnie Anderson
Brigham Young University - Utah, bonnie_anderson@byu.edu

Anthony Vance
Brigham Young University - Utah

David Eargle
University of Pittsburgh, dave@daveeargle.com

Is Your Susceptibility to Phishing Dependent on Your Memory?

Anderson Bonnie Brinton[1]
Brigham Young University, USA

Vance Anthony
Brigham Young University, USA

Eargle David
University of Pittsburgh, USA

ABSTRACT

Phishing has become a major attack vector for hackers and cost victims $687 million in the first half of 2012 alone. Additionally, despite technical solutions to defend against this threat, reports show that phishing attacks are increasing. There is therefore a pressing need to understand why users continue to fall victim to phishing, and how such attacks can be prevented. In this research-in-progress paper, we argue that the cognitive neuroscience of memory provides a useful lens through which to study the problem of phishing. A commonly reported finding from the field of memory is the *eye movement-based memory effect*, the phenomenon of people paying less visual attention to images that have been previously viewed. We aim to show in this paper that this effect holds in the context of email processing, and that the eye movement-based memory effect is a significant contributing factor to users' susceptibility to phishing. We propose an experimental design that uses a memory task involving simulated phishing emails, and measures users' behavioral responses and eye tracking data in response to our phishing manipulations. We further propose to show how training can be designed to help users overcome the eye movement-based memory effect and become less prone to phishing attacks.

[1] Corresponding author: bonnie_anderson@byu.edu

KEYWORDS: eye movement-based memory effect, phishing emails, memory, eye tracking, NeuroIS, behavioral information system security

INTRODUCTION

Phishing has become a major attack vector for hackers and cost victims $687 million in the first half of 2012 alone (Kessem, 2012). Additionally, despite technical solutions to defend against this threat, reports show that phishing attacks are increasing. There is therefore a pressing need to understand why users continue to fall victim to phishing, and how such attacks can be prevented.

In this research-in-progress paper, we argue that the cognitive neuroscience of memory provides a useful lens through which to study the problem of phishing. A commonly reported finding from the field of memory is the *eye movement-based memory effect*, the phenomenon of people paying less visual attention to images that have been previously viewed. We investigate whether the EMM effect holds in the context of email processing, and whether it is a significant contributing factor to users' susceptibility to phishing. Our research question for this paper therefore is:

RQ: How does the EMM effect influence users' susceptibility to phishing?

We propose an experimental design that uses two memory tasks: one involving household objects, and the other involving simulated phishing emails, and measures users' behavioral responses and eye tracking data in response to our phishing manipulations.

BACKGROUND

This research applies decision neuroscience to the information systems (NeuroIS), which provides a methodology whereby the "black box" of cognition can be opened to more directly

observed. For example, one NeuroIS technique—measuring eye-movements—allows researchers to understand what is consciously processed by the participant (Benbasat et al., 2010). Eye movement behavior provides important insights into cognitive processes that may not be available to conscious introspection.

Neuroscientists have established the utility of tracking eye movements as an indirect measure of memory (Hannula et al., 2010). In Smith et al. (2008), participants were asked to identify scenes as "new," "identical," or "changed." The researchers found that eye movements into and out of critical regions are dependent on an individual's awareness that a region has been altered. Those who were unaware or the alteration viewed the image similarly to how they viewed the image before they were shown the altered version. Thus they did not fixate on those areas unless they were conscious a change occurred.

We anticipate the EMM effect will also hold in the context of processing email. Specifically, that people will pay less attention to emails purportedly from the same sender with a similar appearance. Further, we expect that this reduction in attention for familiar looking emails makes individuals more susceptible to phishing.

HYPOTHESES

The *eye movement-basedmemory effect* (EMM), first documented by Althoff et al. (1999) and numerous cognitive neuroscience of memory studies since (see Hannula et al., 2010 for a review), is a phenomenon in which people pay less attention to images that they have previously viewed. We theorize that the EMM effect will hold in the context of email processing. Smith et al. (2006; 2008) examined how participants classified identical, novel, and manipulated stimuli

in gauging the EMM effect. In our context, we will examine whether participants can distinguish, (1) previously seen emails, (2) novel emails, and (3) manipulated phishing emails. Accordingly, we hypothesize:

H1. Participants will exhibit lower viewing activity of previously seen emails as compared with novel emails.

Unfortunately, phishers take advantage of these familiar formats in corporate emails to impersonate organizations—a key phishing tactic (Levy, 2004). We theorize based on the EMM effect that both the visual consistency and repetitive nature of corporate emails will cause users to become more familiar with emails from a certain source, and as a result, less carefully inspect emails received. We therefore hypothesize the following:

H2a. Based on eye-tracking data, participants who incorrectly classify phishing emails as previously seen will exhibit lower viewing activity of these emails than new emails.

H2b. Based on behavioral responses, participants who exhibit lower viewing activity of emails will have lower accuracy in identifying phishing emails.

Memory researchers have also found additional corollary eye movements to the EMM effect: when a previously viewed image is modified, and a person notices the modification, the modified regions of the image receive increased visual attention (Smith et al., 2008). Smith et al. (2006) found that when participants noticed modifications to images, the modified regions received significantly more fixations and repeated viewing.

H3a. Based on eye-tracking data, participants who correctly classify phishing emails will exhibit higher viewing activity of the altered portions of the emails.

H3b. Based on behavioral responses, participants who exhibit higher viewing activity of the altered portions of phishing emails will have greater accuracy in identifying phishing emails.

METHODOLOGY

We have begun our empirical assessment of our hypotheses with a two-part pilot experiment in which we collected data from 45 undergraduate students at a large private university in the U.S. For both parts of the experiment, we used the Tobii T120 monitor and software to capture all eye tracking data. In Part 1 of the experiment, we used the Amsterdam Library of Object Images (Geusebroek, et al. 2005) to populate our image database. For the first round, we showed 60 images to prime the participants. Each object was shown for 3 seconds. Participants were asked to classify each image as Pleasant or Not Pleasant. In the second round, the participants were shown each of the 60 images showed previously, 30 novel images (meaning not shown during the encoding phase) and 30 images similar to those shown in the first round. The subjects were asked to classify each image as *New (Novel), Similar,* or *Old (Identical).*

In Part 2 of the experiment, participants were shown 45 images for 12 seconds each. They were asked to classify each email image as *Aesthetically Pleasing* or *Not Aesthetically Pleasing*. The images were taken from actual emails from companies and organizations sent to the researchers, although the personal identifying information was altered to show John and Jane Doe. Next, participants were again shown 45 images for 12 seconds each. The 45 images were a combination of 15 each of new/novel emails, similar emails, and old/identical emails. The "similar" images were edited versions of the legitimate emails from the first phase. The subjects were asked to classify each image as New (Novel), Similar, or Old (Identical).

In the next stage of our research, we will analyze the data to determine what elements of the email images most affect recall. We will measure the location of the initial fixation, heat map comparisons, heat maps of similar content (size, shape, type), and the number of fixations based

on content type. We will compare the heat maps of these screenshots to determine elements (based on size and type) that people fixate on most. According to the Levels of Processing Model (Craik et al., 1972), what is fixed on most is what is most attended to, which is assumed to be what is most deeply processed cognitively. We will test the eye movements between emails that had been seen previously and emails that were similar to those previously seen (phishing, lure) emails. Eye movement over the course of the 12 seconds each email was viewed will be measured via nine fixation metrics.

PRELIMINARY ANALYSIS OF PRETEST TEST RESULTS

Our preliminary analysis of the pretest data shows that there are significant differences in the identification of similar images (see Table 1). Specifically, we found that people do not accurately recognize similar images, especially the similar images of emails.

Figure 1 presents a heat map analysis of the original alongside the phishing emails. Participants exhibited lower viewing activity for manipulated phishing version of the email as compared to the original email. This is consistent with the EMM effect. Given these initial supportive findings, we will proceed with the primary data collection of our experiment.

Table 1. Accuracy Percentage by Image Classification

Image Classification	Object Images	Email Images
New/novel	87%	81%
Old/identical	94%	83%
Similar	67%	44%

EXPECTED CONTRIBUTIONS

We expect to contribute by demonstrating how the EMMeffect influences individuals' susceptibility to phishing, a question that has not yet been investigated. In doing so, we aim to develop theory to explain why the EMM effect should hold in the context of phishing. Further,

we anticipate that our findings will have implications for other contexts in which the user must make information security decisions through the user interface of the computer, such as security warning dialog boxes.

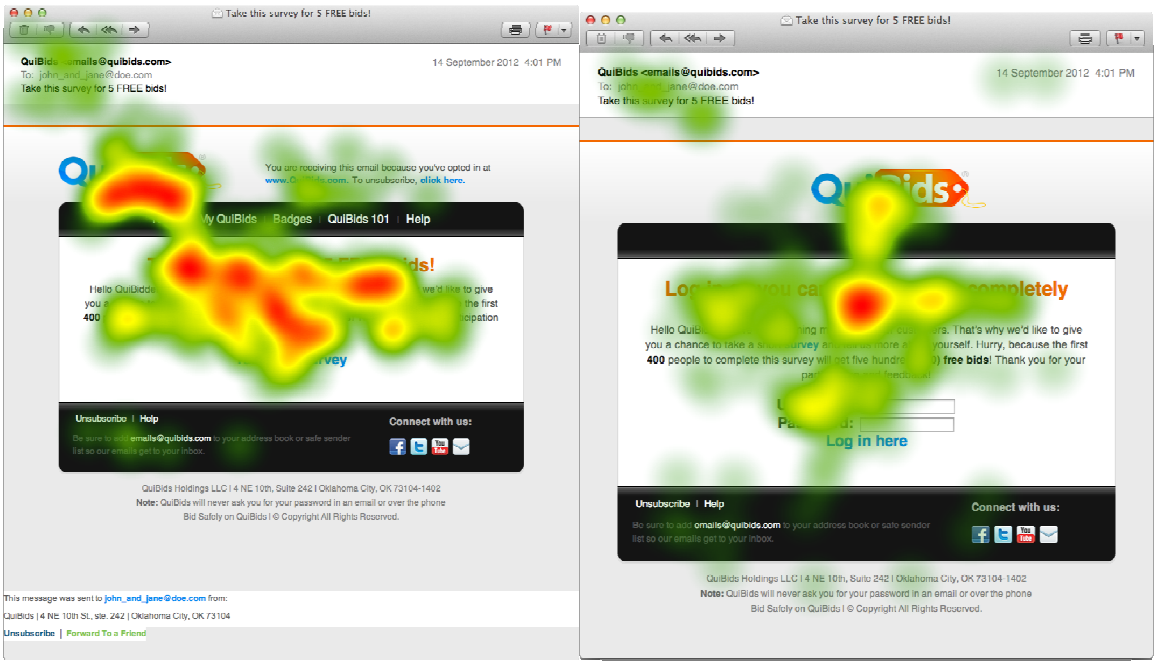

Figure 1. Heat map of the original email (left) compared with heat map of manipulated phishing email (right)

REFERENCES

Althoff, R. R., and Cohen, N. J. 1999. ”Eye-movement-basedmemoryeffect: a reprocessingeffect in face perception,”*Journal of Experimental Psychology: Learning, Memory, and Cognition*, (25:4).

Benbasat, I., Dimoka, A., Pavlou, P. A., and Qiu, L. 2010. ”Incorporating social presence in the design of the anthropomorphic interface ofrecommendation agents: insights from an fMRI study”*ICIS 2010 Proceedings*, St. Louis.

Craik, F. I. M., and Lockhart, R. S. 1972. ”Levelsofprocessing: Aframework for memory research,”*Journal of verbal learning and verbal behavior*, (11:6), pp. 671-684.

Drake, C. E., Oliver, J. J., and Koontz, E. J. 2004. ”Anatomyof a phishing email,”*In: First Conference on Email and Anti-Spam*, MountainView, CA, USA.

Geusebroek, J.M, Burghouts, G.J., and Smeulders, A. W. M. (2005), "The Amsterdam libraryofobject images", *International Journal Computational Vision*, 61(1), 103-112.

Hannula, D. E., Althoff, R. R., Warren, D. E., Riggs, L., Cohen, N. J., and Ryan, J. D. 2010. "Worth a glance: usingeyemovementstoinvestigate the cognitiveneuroscienceofmemory,"*Frontiers in Human Neuroscience*, (4:166).

Kessem, L. 2012. "Phishing in Season: A Look at Online Fraud in 2012,"*RSA: SpeakingofSecurity*. August 21 (available at http://blogs.rsa.com/phishing-in-season-a-look-at-online-fraud-in-2012/).

Levy, E. 2004. "Criminalsbecometechsavvy,"*Security&Privacy, IEEE*, (2:2), pp. 65-68.

Smith, C. N., Hopkins, R. O., and Squire, L. R. 2006. "Experience-Dependent Eye Movements, Awareness, and Hippocampus-DependentMemory,"*The Journal ofNeuroscience*, (26:44), pp. 11304-11312.

Smith, C. N., and Squire, L. R. 2008. "Experience-Dependent Eye MovementsReflect Hippocampus-Dependent (Aware) Memory,"*The Journal ofNeuroscience*, (28:48), pp. 12825-12833.