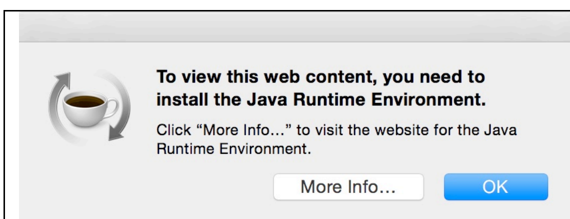
It All Blurs Together: How the Effects of Habituation Generalize Across System Notifications and Security Warnings

Abstract. Habituation to security warnings—the diminished response to a warning with repeated exposures—is a well-recognized problem in information security. However, the scope of this problem may actually be much greater than previously thought because of the neurobiological phenomenon of generalization. Whereas habituation describes a diminished response with repetitions of the same stimulus, *generalization* occurs when habituation to one stimulus carries over to other novel stimuli that are similar in appearance. Because software user interface guidelines call for visual consistency, many notifications and warnings share a similar appearance. Unfortunately, generalization suggests that users may already be deeply habituated to a warning they have never seen before because of exposure to other notifications. In this work-in-progress study, we propose an eye tracking and fMRI experiment to examine how habituation to frequent software notifications generalizes to infrequent security warnings, and how security warnings can be designed to resist the effects of generalization.

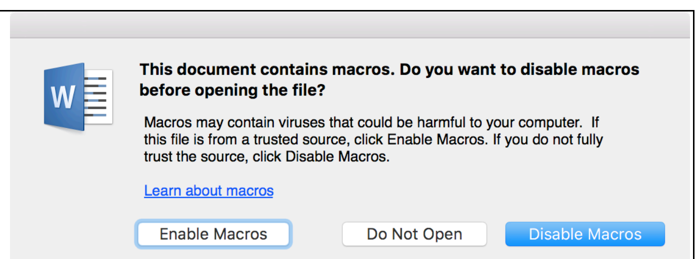
Keywords: Security warnings · habituation · generalization · fMRI · eye tracking · NeuroIS.

1 Introduction

Habituation is a major factor in users' failure to adhere to security warnings [1, 2]. Through this neurological phenomenon, repeated warnings that once were salient become virtually unnoticeable. *Generalization* occurs when the effects of habituation to a repeated stimulus carry over to other novel stimuli that are similar in appearance [3]. Applied to the domain of information security, generalization suggests that users not only habituate to individual security warnings, but also to whole classes of notifications and warnings that share a similar appearance and user interaction (UI) paradigm (see Figure 1). If true, then the threat and potential impact of habituation is much broader than previous work has suggested [4-6], as users may already be deeply habituated to a security warning that they have never seen before.



System-generated notification



Security warning

Figure 1. A notification and security warning. Note the similarities in UI and mode of interaction.

We outline an experiment using eye tracking and fMRI to (1) measure the extent to which the effects of habituation generalize across similar types of notifications and security warnings, and (2) determine warning designs that can reduce the occurrence of generalization.

2 Literature Review

2.1 Habituation and Generalization

Habituation is defined as the decreased response to a repeated stimulation [7]. Habituation is fundamentally a neurobiological phenomenon [8], and is evident "in every organism studied, from single-celled protozoa, to insects, fish, rats, and people" [9, p. 125]. Not surprisingly, humans also exhibit habituation to a wide variety of stimuli as early as infancy [10].

Habituation with novel stimuli can have one of two responses: stimulus specificity, or stimulus generalization. With stimulus specificity, a novel stimulus similar to the habituated stimulus, will not show a habituated response [11]. However, the fields of neuroscience and neurobiology show that the effects of habituation to a familiar stimulus can generalize to other, novel (i.e., never before experienced) stimuli that share similar characteristics [3, 12]. This occurs for animals with simple stimuli [13] as well as for people with complex, real-world stimuli such as video games [11]. This generalization means that although the stimulus is novel, the response is similar to a habituated response to a repeated stimulus. When habituation occurs, generalization is also likely to occur.

2.2 Habituation and Generalization to Security Warnings

Given its strong security implications, habituation is frequently cited as a key contributor to users' failure to heed warnings. However, many studies infer the presence of habituation, rather than empirically examine it. For example, Sunshine et al. [14] found a correlation between user disregard for warnings and user recognition of warnings as previously viewed, and attributed this correlation to habituation. An exception is Anderson et al. [4], who used fMRI to measure habituation in the brain in response to warnings. Their results showed a large drop in activity in the visual processing centers of the brain after only the second exposure to a warning, and found further decreases with additional exposures.

Generalization of habituation effects has important implications for security warnings. Although users are regularly presented with a variety of warnings [5], research on human–computer interaction indicates that the frequency of warnings is small compared with the barrage of system-generated notifications that are common to the computing experience [15, 16]. The high frequency of notifications is compounded by UI guidelines that call for visual consistency across UI elements such as notifications [12]. If habituation to these notifications generalize, users are unlikely to give full attention to security warnings even if they haven't seen them before [9].

The effects of generalization of software notifications to warnings has not been empirically examined previously. Instead, researchers have observed that habituation to a single warning in one context carries over to a different context. For example, Sunshine et al. [14] observed that users who correctly identified the risks of an SSL warning in a library context inappropriately identified these same risks in a banking context. Similarly, Amer et al. [17] found that users who habituated to exception notifications in an accounting information system in one context were habituated to a different through visually identical exception notification in a different context. However, in each of these cases, users habituated to the same type of security warning or notification. As a result, it is unclear to what extent software notifications generalize to security warnings.

2.3 Hypotheses

Replicating the findings of previous studies [18], we hypothesize that users will habituate to software notifications. When users repeatedly see software notifications, the brain creates a mental model of these notifications. Rather than giving attention to future exposures to the notifications, the brain increasingly relies on this mental model. As a result, users' responses to future notifications decrease (i.e., habituate) in response to repeated exposures of notifications [19]. In summary, we predict:

H1: Users will habituate to repeated exposures of software notifications.

We predict that habituation to software notifications will also decrease users' responses to security warnings. Habituation to stimuli extends beyond the specific sensory characteristics of a repeated stimulus. Rather, habituation is also "happening centrally rather than in primary sensory afferents" [3, p. 137]. In other words, when a person sees and habituates to repeated software notifications, this decreases one's response to other stimuli in the environment. People's responses to security warnings may be particularly vulnerable to the effects of software notification habituation because they share several of the same properties (e.g., both popup on a computer screen, they often share the same shape or size, and some even have the same look and feel). As such, the brain may rely on the mental models of software notifications when interpreting security messages, which will result in generalized habituation to the security messages [19]. In summary, we predict:

H2: Habituation to software notifications will generalize to security warnings.

To test for generalization, we will present several stimuli equally capable of eliciting a response. We will compare whether the response to the first novel stimulus is similar to the first exposure original stimulus. If the novel stimulus response is weaker than the first exposure of the original stimulus, then there is evidence of generalization. However, if the novel stimulus response is greater than the first exposure of the original stimulus, that provides evidence of stimulus specificity habituation [11].

3 Experimental Design

3.1 Methodology

We plan to use fMRI and eye-tracking tools simultaneously. We are interested in capturing data that will allow us to examine the levels of cognitive processing, thus requiring a neurophysiological tool such as fMRI that has high spatial resolution of brain activity. While fMRI has the highest spatial resolution of any non-invasive volumetric brain imaging technology [20], tools such as eye-tracking have a much higher temporal resolution. In addition, by using these two tools concurrently, we can examine the eye fixations and pair them with the neural responses. We will use the Siemens 3T Tim-Trio scanner at our university's MRI research facility. In addition, we will collect eye-tracking data during each scan using an MRI-compatible SR Research EyeLink 1000 Plus long-range eye tracker.

3.2 Task

The task will use a common go/no-go design [21], in which participants will be instructed to respond to items they see both as quickly and as accurately as possible by pressing a button on a button box. Go trials, or trials in which participants must press a button each time the stimulus is presented, will consist of images of notifications. No-go trials, or trials in which participants must inhibit the tendency to press a button and withhold their response, will be made up of warnings. Trial types will be randomized and split so that 80% of the trials are go trials and 20% of the trials are no-go trials. The exact timing of stimulus presentation will be determined during pilot testing, but warnings and notifications will be displayed for 500–1000 ms with a jittered inter-stimulus interval of 1000–2000 ms. All stimuli will be normalized in size along the major axis of the image to 750 pixels.

A structural scan will be acquired for each subject for functional localization after which the behavioral task will be completed over two functional runs, each lasting about 8 minutes. Trials will be randomly assigned to the runs.

Three regions of interest will be examined, namely the visual cortex and ventral visual pathway, medial temporal lobe, and lateral prefrontal cortex. The analysis of each of these regions allows for an examination of generalization at different levels of processing. First, the visual cortex and ventral visual pathway are involved in object perception [see 22 for review]. Second, the medial temporal lobe is involved memory specificity, or the detection of differences between similar stimuli [23]. Lastly, the dorso and ventrolateral prefrontal cortices, are involved in response inhibition [24].

3.3 Analysis

The analysis of these different levels of processing allows for the detection of the specific aspect of the stimuli are allowing for greater ability to recognize differences and withhold a response. For instance, if habituation to notifications does not influence generalization to warnings, one would expect to see a greater amplitude of fMRI activation for each warning because it is treated as a new stimulus. If, however, habituation to notifications does generalize to warnings, there is no reason why there would be a greater fMRI activation for warnings over notifications. A warning would be treated as if it was another notification and the habituation would continue. For example, no difference in fMRI activation for the notification and warning trials in the medial temporal lobe might suggest that the notifications and warnings are too similar and the visual design should be changed to make them easier to tell apart. Conversely, if there are differences in the medial temporal lobe, but no differences are observed in the lateral prefrontal cortex, then the problem doesn't lie in the visual differences of the stimulus. Rather the difficulty comes in inhibiting the common response and different designs in the mode of dismissal of the warning or notification (e.g., dragging a bar across the window for a warning rather than clicking a button for a notification) may be better design changes to help prevent this generalization.

4 Anticipated Contributions

We anticipate that our findings of this proposed study will complement and extend previous work that examined habituation to individual warnings. With proposed experimental design, we intend to examine how the effects of habituation to frequent software notifications generalize to novel security warnings. Specifically, our anticipated contributions are:

1. Determine how the effects of habituation to frequent notifications and warnings generalize to novel warnings.
2. Measure the severity of generalization.
3. In a future behavioral study based on these outcomes, develop and test distinctive warning designs that resist generalization.

References

1. Bravo-Lillo, C., Cranor, L., Komanduri, S., Schechter, S. Sleeper, M.: Harder to Ignore? Revisiting Pop-up Fatigue and Approaches to Prevent It USENIX Association. 105-111 (2014)
2. Kalsher, M. Williams, K.: Behavioral Compliance: Theory, Methodology, and Result, in Handbook of Warnings. Lawrence Erlbaum Associates: Mahwah NJ. p. 313-331 (2006)
3. Rankin, C.H., Abrams, T., Barry, R.J., Bhatnagar, S., Clayton, D.F., Colombo, J., Coppola, G., Geyer, M.A., Glanzman, D.L., Marsland, S., McSweeney, F.K., Wilson, D.A., Wu, C.-F. Thompson, R.F.: Habituation Revisited: An Updated and Revised Description of the Behavioral Characteristics of Habituation. Neurobiol. Learn. Mem. 92:2, 135-138 (2009)
4. Anderson, B.B., Kirwan, C.B., Jenkins, J.L., Eargle, D., Howard, S. Vance, A.: How Polymorphic Warnings Reduce Habituation in the Brain: Insights from an fMRI Study, in Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, ACM: Seoul, Republic of Korea. p. 2883-2892 (2015)
5. Bravo-Lillo, C., Komanduri, S., Cranor, L.F., Reeder, R.W., Sleeper, M., Downs, J. Schechter, S.: Your Attention Please: Designing Security-Decision UIs to Make Genuine Risks Harder to Ignore, in Proceedings of the Ninth Symposium on Usable Privacy and Security ACM: Newcastle, United Kingdom. p. 1-12 (2013)
6. Egelman, S., Cranor, L.F. Hong, J.: You've Been Warned: An Empirical Study of the Effectiveness of Web Browser Phishing Warnings, in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM: Florence, Italy. p. 1065-1074 (2008)
7. Thompson, R.F. Spencer, W.A.: Habituation: A Model Phenomenon for the Study of Neuronal Substrates of Behavior. Psychological Review. 73:1, 16-43 (1966)
8. Ramaswami, M.: Network Plasticity in Adaptive Filtering and Behavioral Habituation. Neuron. 82:6, 1216-1229 (2014)
9. Rankin, C.H.: Introduction to Special Issue of Neurobiology of Learning and Memory on Habituation. Neurobiol. Learn. Mem. 92:2, 125-126 (2009)
10. Colombo, J. Mitchell, D.W.: Infant Visual Habituation. Neurobiol. Learn. Mem. 92:2, 225-234 (2009)
11. Grizzard, M., Tamborini, R., Sherry, J.L., Weber, R., Prabhu, S., Hahn, L. Idzik, P.: The Thrill Is Gone, but You Might Not Know: Habituation and Generalization of Biophysiological and Self-Reported Arousal Responses to Video Games. Communication Monographs. 82:1, 64-87 (2015)
12. Simpson, N.: Managing the Use of Style Guides in an Organisational Setting: Practical Lessons in Ensuring UI Consistency. Interacting with Computers. 11:3, 323-351 (1999)
13. Ryan, J.D., Althoff, R.R., Whitlow, S. Cohen, N.J.: Amnesia Is a Deficit in Relational Memory. Psychol. Sci. 11:6, 454-461 (2000)

14. Sunshine, J., Egelman, S., Almuhimedi, H., Atri, N. Cranor, L.F.: Crying Wolf: An Empirical Study of SSL Warning Effectiveness, USENIX Association. 399-416 (2009)
15. Mark, G., Gudith, D. Klocke, U.: The Cost of Interrupted Work: More Speed and Stress, in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM: Florence, Italy. 107-110 (2008)
16. Mark, G., Voida, S. Cardello, A.: A Pace Not Dictated by Electrons: An Empirical Study of Work without Email, in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM: Austin, Texas, USA. 555-564 (2012)
17. Amer, T.S. Maris, J.-M.B.: 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 (2007)
18. Anderson, B., Kirwan, B., Eargle, D., Howard, S. Vance, A.: How Polymorphic Warnings Reduce Habituation in the Brain: Insights from an fMRI Study, in Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI), ACM: Seoul, South Korea (2015)
19. Groves, P.M. Thompson, R.F.: Habituation: A Dual-Process Theory. Psychological Review. 77, 419-450 (1970)
20. Riedl, R. Léger, P.-M.: Fundamentals of NeuroIS: Information Systems and the Brain. Studies in Neuroscience, Psychology and Behavioral Economics, Springer Berlin Heidelberg (2016)
21. Nosek, B.A. Banaji, M.R.: The Go/No-Go Association Task. Social Cognition. 19:6, 625-666 (2001)
22. Grill-Spector, K.: The Neural Basis of Object Perception. Current Opinion in Neurobiology. 13:2, 159-166 (2003)
23. Kirwan, C.B. Stark, C.E.L.: Overcoming Interference: An fMRI Investigation of Pattern Separation in the Medial Temporal Lobe. Learn. Memory. 14:9, 625-633 (2007)
24. Liddle, P.F., Kiehl, K.A. Smith, A.M.: Event-Related fMRI Study of Response Inhibition. Human Brain Mapping. 12:2, 100-109 (2001)