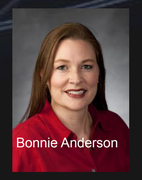
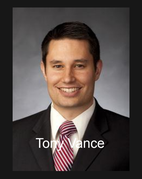
Using Neuroscience to Explain User Responses to Malware Warnings

Bonnie Anderson,[1] Anthony Vance[1], Brock Kirwan,[2] James Hansen[1]
Department of Information Systems,[1] Department of Psychology,[2] Brigham Young University





Introduction

- Malicious software (or malware) constitutes a serious threat to information security of organizations.
- There is a need to understand how malware warnings can be made more effective to alert end users of potential threats.
- Our research will look at the neuroscience associated with subjects' identification, recall, and reaction to malware warnings.

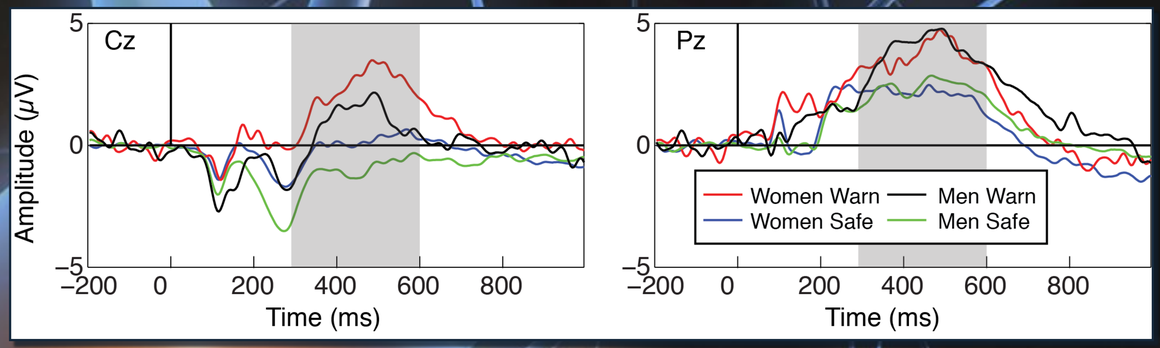


Figure 1: Grand average ERP waveforms at the Cz and Pz electrode sites. At the Cz electrode site, women had greater amplitudes for the P300 for both safe websites and the warning screen. The P300 amplitude was enhanced for the warning screen across genders at both electrode sites.

Preliminary Study

- We performed a NeuroIS study, employing electroencephalography (EEG) and measuring the P300 wave to examine whether men and women process malware warnings in the brain.
- P300 is a measurement of voltage changes within the brain associated with attention and memory operations
- Stronger measurements relating to novel stimuli.

Results

- **H1:** P300 is higher for all participants when viewing malware warning screenshots than when viewing legitimate website screenshots.

H2: P300 is higher for women than for men when viewing malware warning screenshots when examining the Cz region, not supported with Pz region.

H3: P300 latency was not slower for women or men.

H4: P300 did not vary when viewing red malware warning screenshots than when viewing grayscale malware warning screenshots. Nor when viewing red legitimate website screenshots than when viewing non-red legitimate website screenshots.

1-Motivation and risk aversion

- We will examine behavioral economics insights to why users disregard warnings.
- Participants will be required to perform a computer-based classification task.
- Periodic malware warning screens will interrupt the task.
- The quantity of malware warning screens will be adaptive based on behavior.
- Consequences of ignoring the warnings may be simulated.
- Motivation and risk aversion will be evaluated.

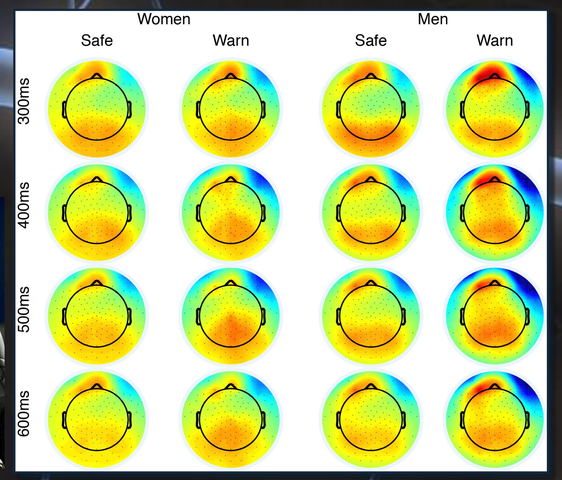
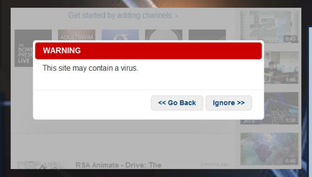


Figure 2: Topographical distribution of ERP potentials for the 300-600 ms post-stimulus periods. There was a centrally distributed positivity at 300ms post-stimulus onset.

2-Memory and memetics regarding malware warning design

- We will examine color, animation, and a variety of other design characteristics and evaluate memorability.
- We will use eye-tracking equipment to track participants' eye movements.



3- Neuroscience and security

- Develop and perform fMRI studies examining areas of the brain associated with fear, risk, trust.
- Participants will be given tasks to complete while in an fMRI machine.
- Malware warnings will interrupt the tasks.
- Blood flow in the brain will show areas of activity.
- We will look specifically at the areas of the brain associated with fear, risk, trust.

NSF Secure and Trustworthy Cyberspace
Inaugural Principal Investigator Meeting
27-29 November 2012
National Harbor, MD

Meet the principal investigators! Place your Post-its® here!

Our gratitude to research assistants Dixon Brown, David Eargle, Lee J Hinkle, and Arthur Weagel