THE UNIVERSITY OF ALABAMA AT BIRMINGHAM

Knowledge that will change your world

Neural Signatures of User-Centered Security: An fMRI Study of Phishing, and Malware Warnings

Ajaya Neupane

Department of Computer and Information Sciences
University of Alabama at Birmingham

Co-authors: Nitesh Saxena, Keya Kuruvilla, Michael Georgescu, and Rajesh Kana

Outline

- Introduction
- Our Work
- Research Questions
- Contributions
- Design of Experiments
- Data Collection
- Data Analysis and Results
- Conclusions

Introduction and Motivation

- Study of user-centered security
- Several lab-based studies on security warnings and security indicators are done
  - **Conclusion:** users hardly perform well at these tasks.
- Akhawe and Felt [Usenix Sec'13] large scale field study of modern browser's phishing, SSL and Malware Warnings
  - **Conclusion:** users actually heed these warnings with high likelihood

- **Motivation**:
  - Need a better understanding of user-centered security behavior

Our Work

- What we studied?
  - User-centered security from a neuropsychological standpoint
  - Measured user's security performance and the underlying neural activity with respect to two critical security tasks
    - **Phishing Detection** distinguishing between a legitimate and a phishing website
    - **Malware Warnings**: Heeding security (malware) warnings
- How we did it?
  - Using fMRI (functional Magnetic Resonance Imaging)
  - What fMRI does?
    - Blood Oxygen Level Dependent function measure.
    - Measures brain activity by detecting related changes in local cerebral blood flow
    - Better spatial resolution

Our Work (Contd.)

- What we were looking for ?
  - Brain Areas that might be controlling user's performance in phishing and warnings tasks – **Neural Signatures**

Fig: A Pilot Subject being prepared for the scan

Research Questions

- Fundamental questions driving our research
  - Whether or not users actively engage in security tasks?
  - What brain regions get activated while performing these tasks?
  - Does certain personality traits influence users’ security behavior and task performance ?
  - Is users’ behavior in one task related to their behavior in other tasks ?

Contributions

- Novel methodology to study user-centered security
  - Interdisciplinary innovation across **Computer Science, Psychology, and Neuroscience**
- Designed and developed in-scanner fMRI experiments
- Comprehensive analysis of neural and behavioral data
- Limitations
  - Tasks performed inside fMRI scanner
  - Constrained interface

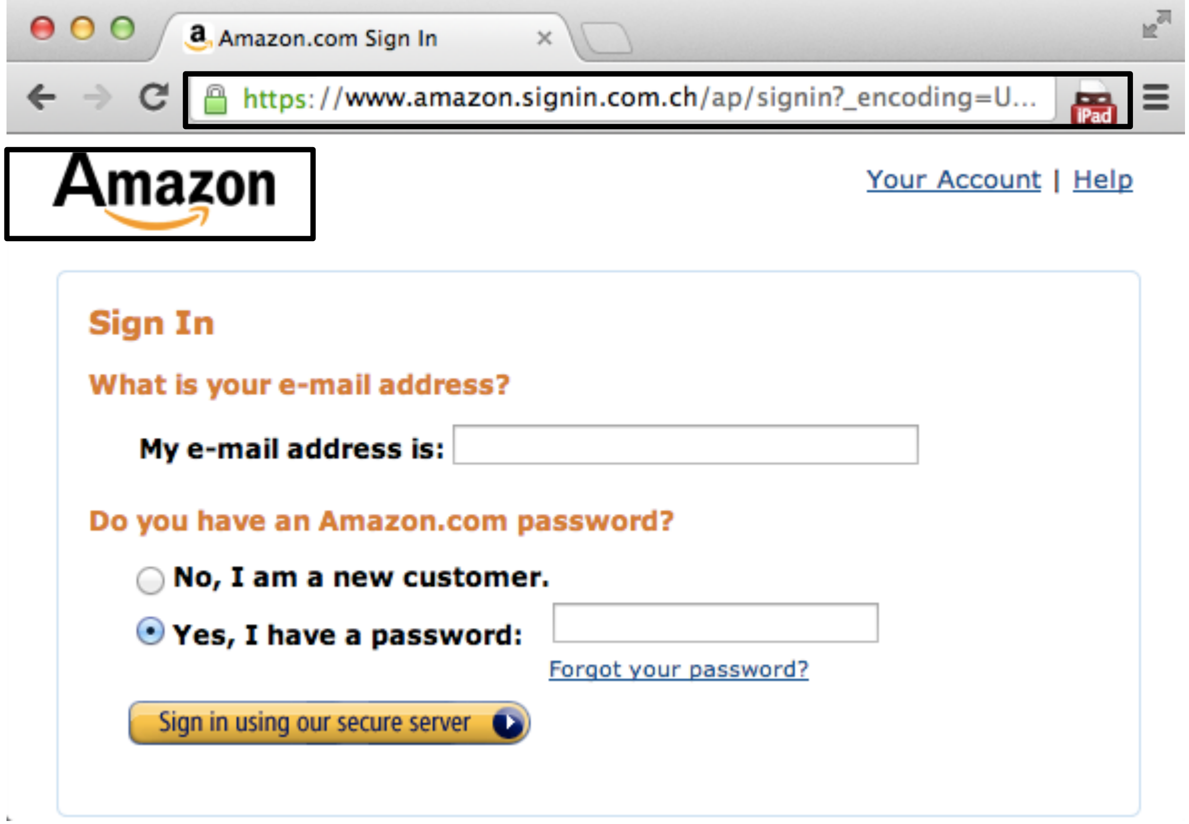
Design of Experiments (1/8)

- **Phishing Experiment**
  - Presented snapshots of real & fraudulent versions of popular websites
  - Instruction:
    - Figure out if the website was real or fake

- **Phishing-Control**
  - Baseline for Phishing Experiment to capture the effect of visual activation
  - Presented snapshots of popular websites
  - Instructions:
    - View each website (no response needed)

Design of Experiments (2/8)

- Sample snapshot used in Phishing-Control & Phishing Experiments



Amazon.com Sign In

https://www.amazon.signin.com.ch/ap/signin?_encoding=U...

Amazon

Your Account | Help

Sign In

What is your e-mail address?

My e-mail address is:

Do you have an Amazon.com password?

○ **No, I am a new customer.**

● **Yes, I have a password:**

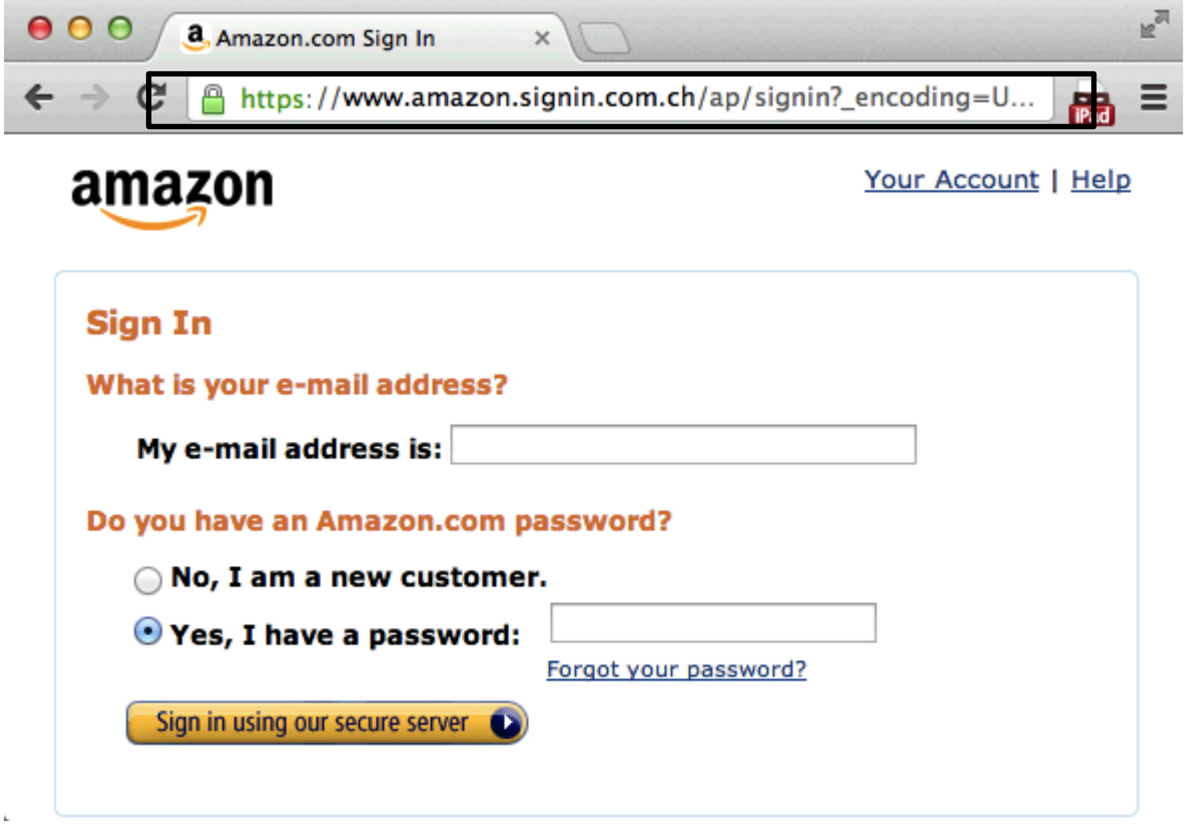
Forgot your password?

Sign in using our secure server

Fig: Sample website easy fake

Design of Experiments (3/8)

- Sample snapshot used in Phishing-Control & Phishing Experiment



Amazon.com Sign In

https://www.amazon.signin.com.ch/ap/signin?_encoding=U...

amazon

Your Account | Help

Sign In

What is your e-mail address?

My e-mail address is:

Do you have an Amazon.com password?

○ **No, I am a new customer.**

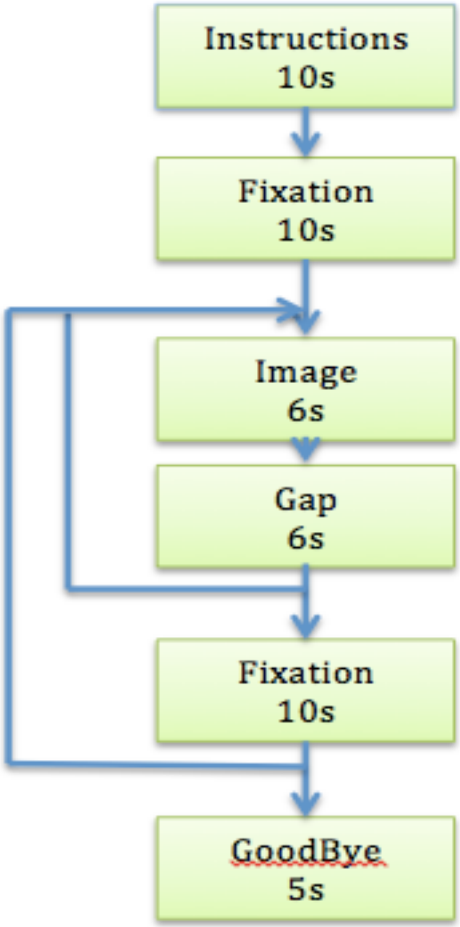
● **Yes, I have a password:**

Forgot your password?

Fig: Sample website difficult fake

Design of Experiments (4/8)

➢ Timing Structure Phishing and Phishing-Control Experiment



Fixation: Baseline when not doing anything

Phishing:

No of trials= 36

Real : 12, Easy Fake: 12, Difficult Fake: 12

Phishing Control:

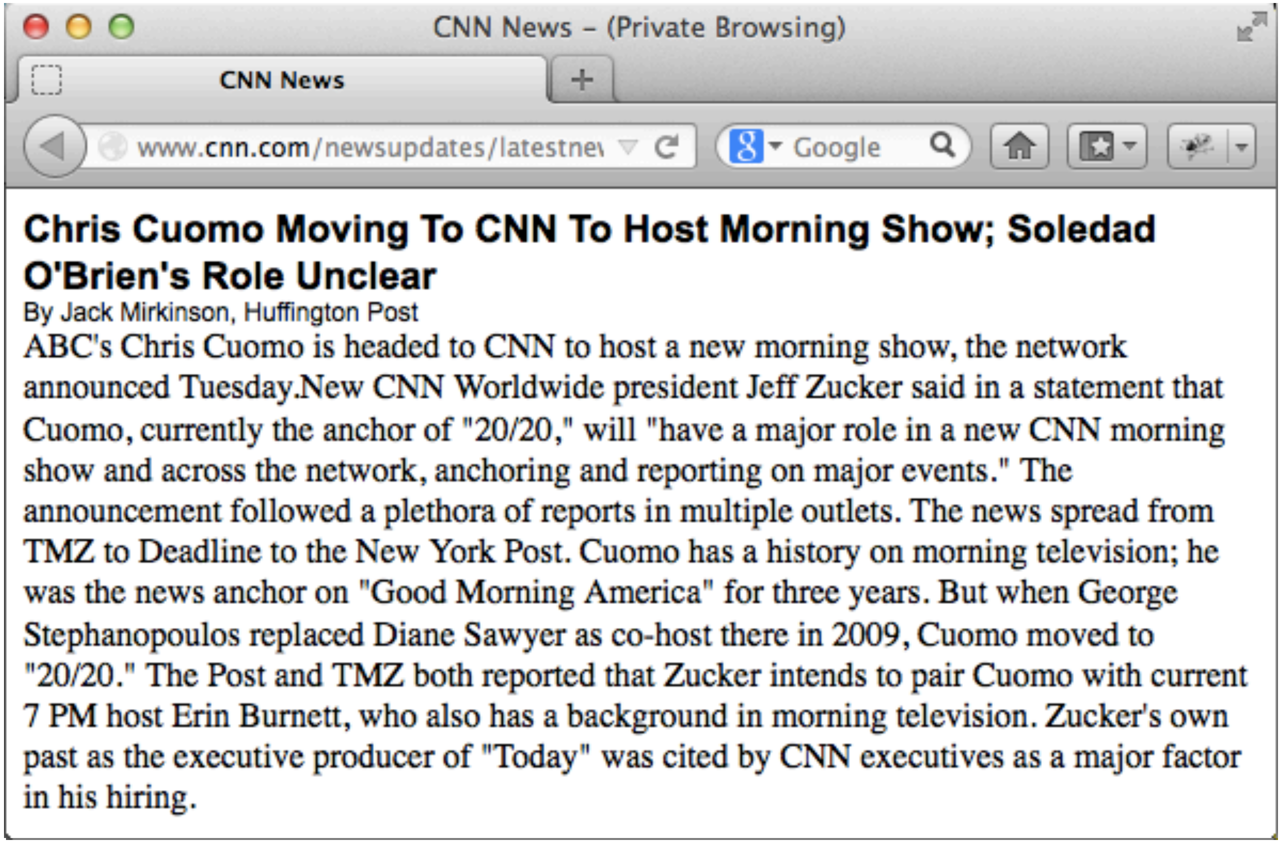
No of trials= 20

Design of Experiments (5/8)

- **Malware Experiment**
  - Participants asked to read abstract of news item
  - Interrupted by Pop-Ups while reading
  - Pop-up Types:
    - Warnings
    - Non-Warnings – Casual Pop-Ups
  - Constrained interface – only able to display rudimentary versions of warnings

Design of Experiments (6/8)

- Malware Experiment


CNN News – (Private Browsing)
CNN News
+
www.cnn.com/newsupdates/latestnev
Google

Chris Cuomo Moving To CNN To Host Morning Show; Soledad O'Brien's Role Unclear

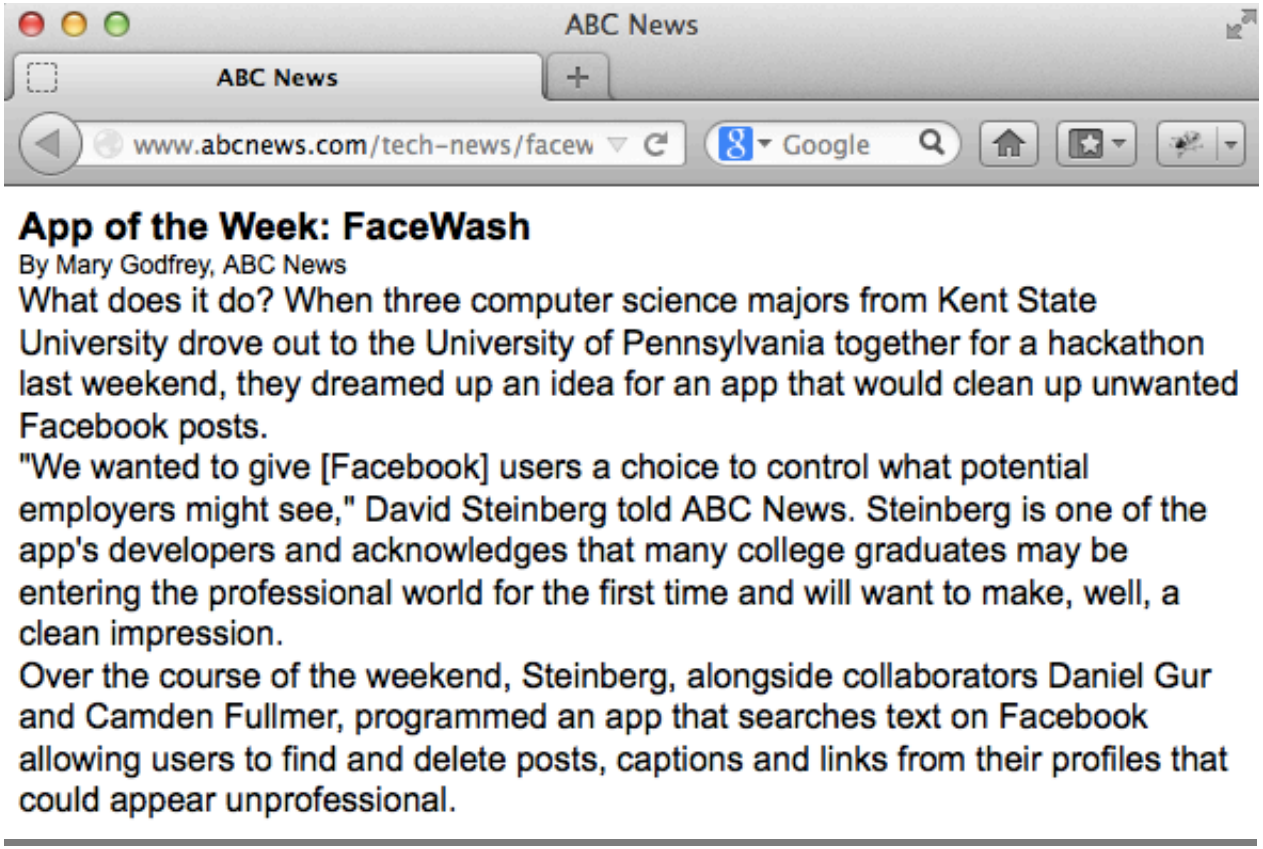
By Jack Mirkinson, Huffington Post

ABC's Chris Cuomo is headed to CNN to host a new morning show, the network announced Tuesday.New CNN Worldwide president Jeff Zucker said in a statement that Cuomo, currently the anchor of "20/20," will "have a major role in a new CNN morning show and across the network, anchoring and reporting on major events." The announcement followed a plethora of reports in multiple outlets. The news spread from TMZ to Deadline to the New York Post. Cuomo has a history on morning television; he was the news anchor on "Good Morning America" for three years. But when George Stephanopoulos replaced Diane Sawyer as co-host there in 2009, Cuomo moved to "20/20." The Post and TMZ both reported that Zucker intends to pair Cuomo with current 7 PM host Erin Burnett, who also has a background in morning television. Zucker's own past as the executive producer of "Today" was cited by CNN executives as a major factor in his hiring.

Fig: Sample Trial Malware Experiment

Design of Experiments (7/8)

- Malware Experiment



App of the Week: FaceWash

By Mary Godfrey, ABC News

What does it do? When three computer science majors from Kent State University drove out to the University of Pennsylvania together for a hackathon last weekend, they dreamed up an idea for an app that would clean up unwanted Facebook posts.

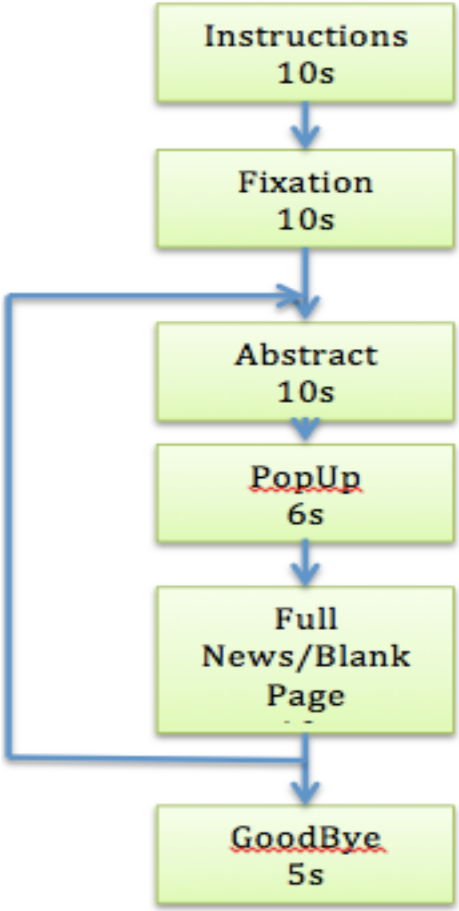
"We wanted to give [Facebook] users a choice to control what potential employers might see," David Steinberg told ABC News. Steinberg is one of the app's developers and acknowledges that many college graduates may be entering the professional world for the first time and will want to make, well, a clean impression.

Over the course of the weekend, Steinberg, alongside collaborators Daniel Gur and Camden Fullmer, programmed an app that searches text on Facebook allowing users to find and delete posts, captions and links from their profiles that could appear unprofessional.

Fig : Sample Trial Malware Experiment

Design of Experiments (8/8)

➢ **Timings Structure of Malware Experiment**



Instructions
10s

Fixation
10s

Abstract
10s

PopUp
6s

Full
News/Blank
Page

GoodBye
5s

Malware:

No of trials = 18

Warnings Pop-Up = 9

Non-Warnings Pop-Up=9

Data Collection

- Participant Recruitment
  - Study Approved by UAB's IRB
  - 25 Participants recruited
- Pre-Scanning Phase
  - Participants answered (BIS) Barratt's Impulsivity questionnaire. Why BIS?
    - To determine the trait impulsivity level of the participants
- Scanning Phase
  - fMRI scan, E-Prime to display stimulus, response and response time recorded.
  - Phishing-Control-Experiment
  - Phishing Experiment
  - Malware Experiment

Overview of Brain Regions

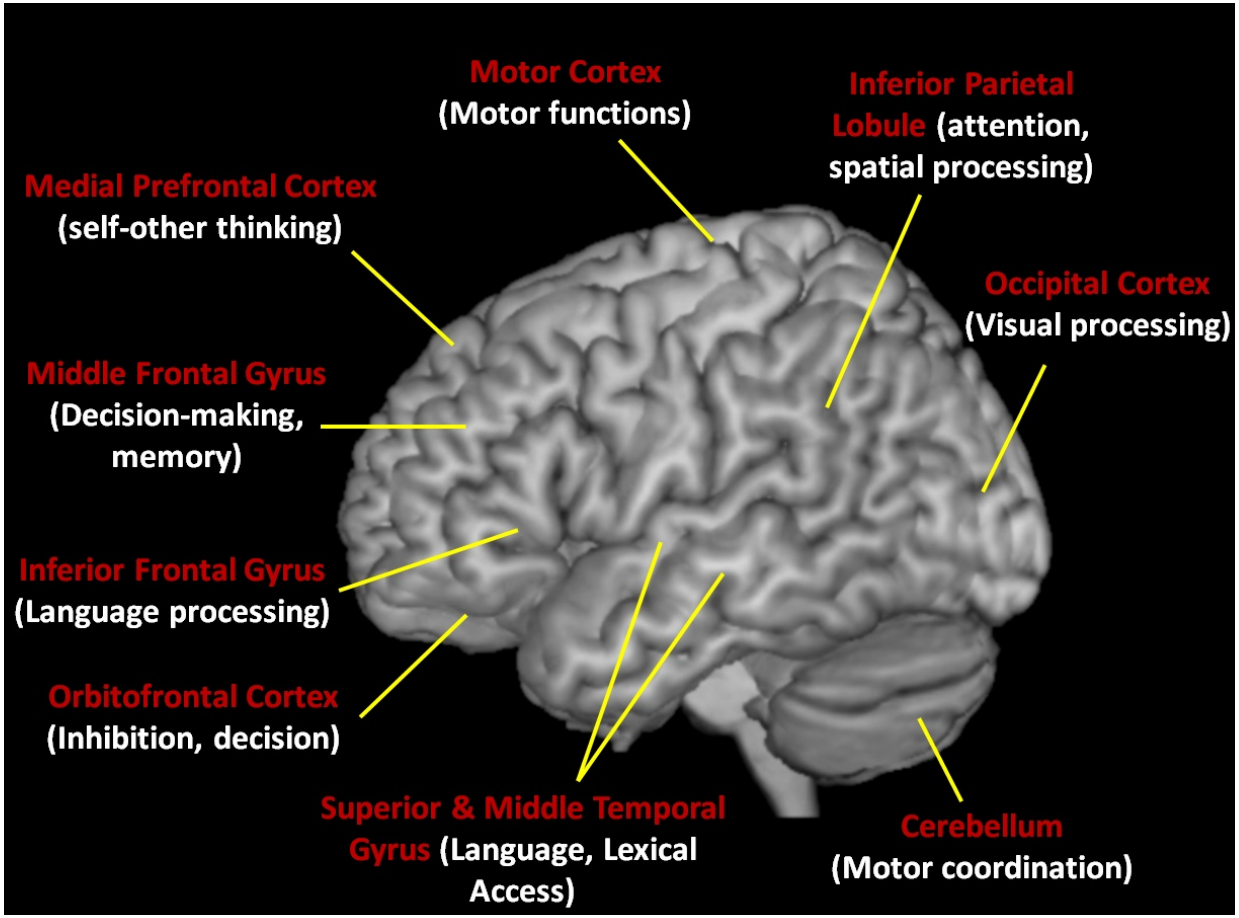
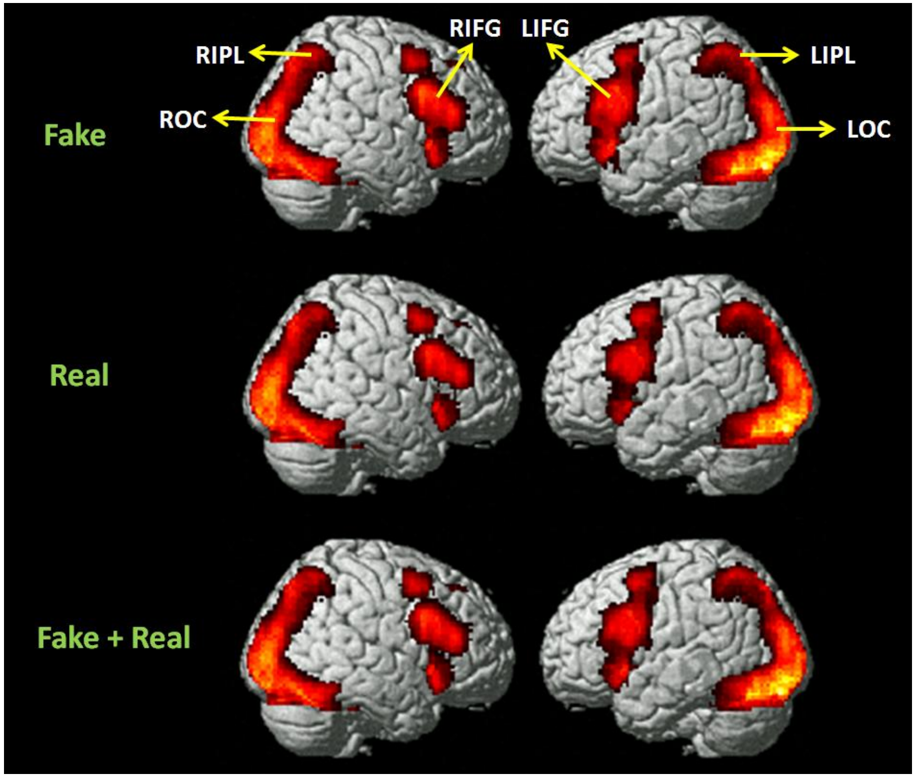

Motor Cortex (Motor functions)
Inferior Parietal Lobule (attention, spatial processing)
Medial Prefrontal Cortex (self-other thinking)
Occipital Cortex (Visual processing)
Middle Frontal Gyrus (Decision-making, memory)
Inferior Frontal Gyrus (Language processing)
Orbitofrontal Cortex (Inhibition, decision)
Superior & Middle Temporal Gyrus (Language, Lexical Access)
Cerebellum (Motor coordination)

Fig: Brain Image

Data Analysis And Results (2/10)

- Phishing Experiment
  - **Contrasts:** Fake, Real, and Real+Fake vs Fixation



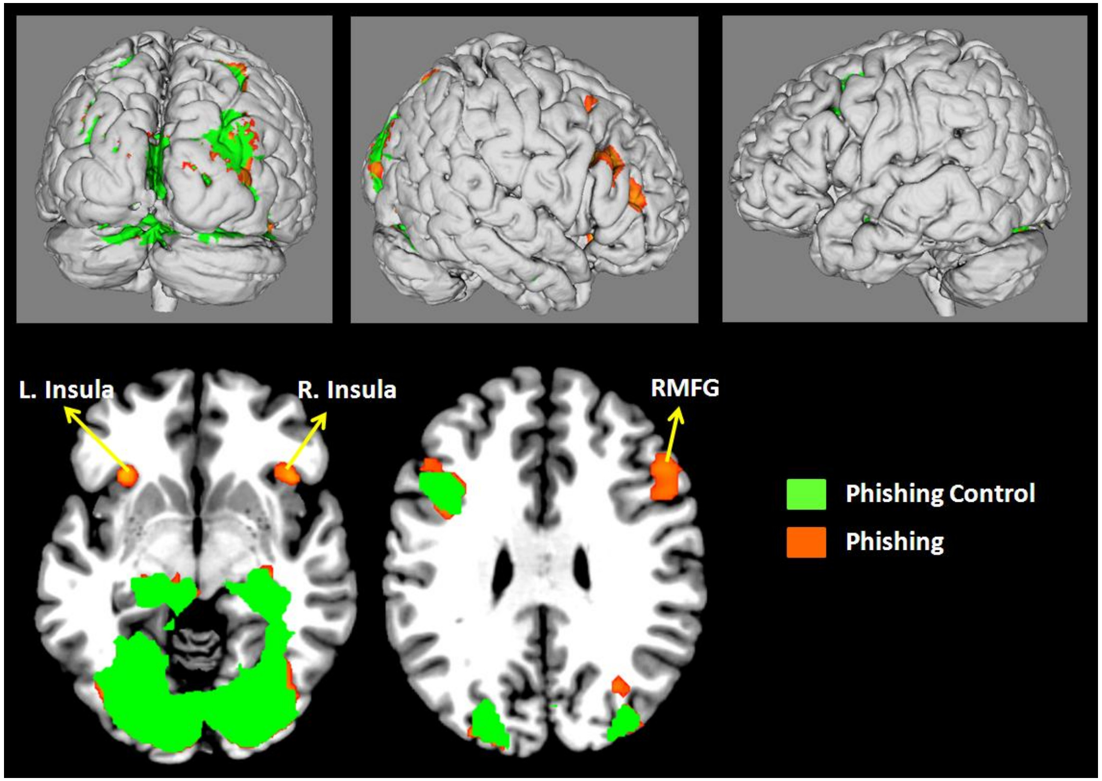
Activity in **Parietal, Frontal, and Occipital brain areas.**

What does it Mean?

- Areas associated with searching, attention shift, problem solving and decision-making
- Participants were undergoing significant effort in making important judgments about the legitimacy of the website

Data Analysis And Results (2/10).

- Phishing Experiment
  - Contrast Phishing vs Phishing-Control



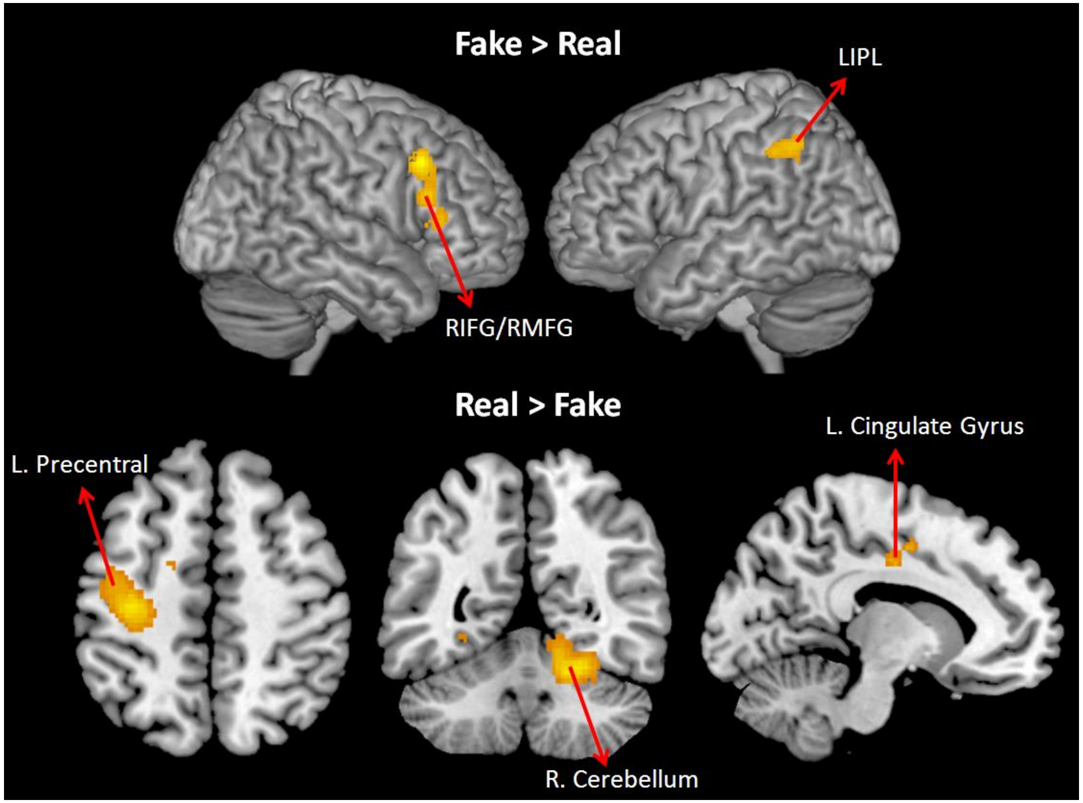
Activity in middle frontal, bilateral insula for Phishing expt.

What does it mean?

- Areas associated with cognitive judgments
- Participants were conscientiously making an effort as to differentiate real/ fake websites

Data Analysis And Results (3/10)

- Phishing Experiment
  - Fake contrasted with Real & Real contrasted with Fake



Fake > Real (increased activity in middle frontal and inferior parietal areas)

- Inline with Mengfei, Frontiers in Human Neuroscience, Vol,5, 2011

What does it Mean?

- Areas implicated with memory, decision-making
- Suggests more strategic and controlled approach, identifying fake websites

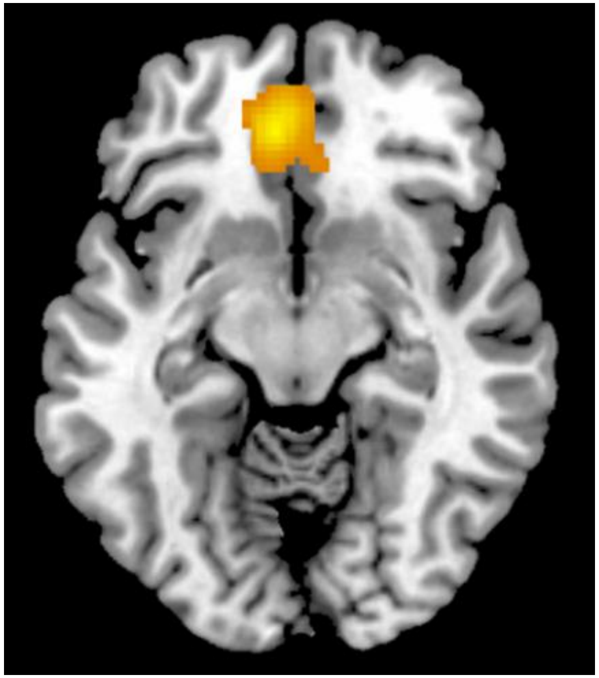
Real > Fake (Activity in cerebellum, precentral)

What does it mean?

- Attention, Decision-Making, Visual Processing

Data Analysis And Results (4/10)

- Phishing Experiment
  - **Relationship between Trait Impulsivity and Phishing task**
    - Negative relationship in medial prefrontal cortex (MPFC), more impulsive individuals had less activity in MPFC



Less Activity in Pre-Frontal Cortex

What does it mean?

- Area associated with decision-making and Problem-Solving
- Conflict and difficulty involved in making real or fake decisions

Data Analysis And Results (5/10)

- Phishing Experiment
  - Behavioral Data Analysis

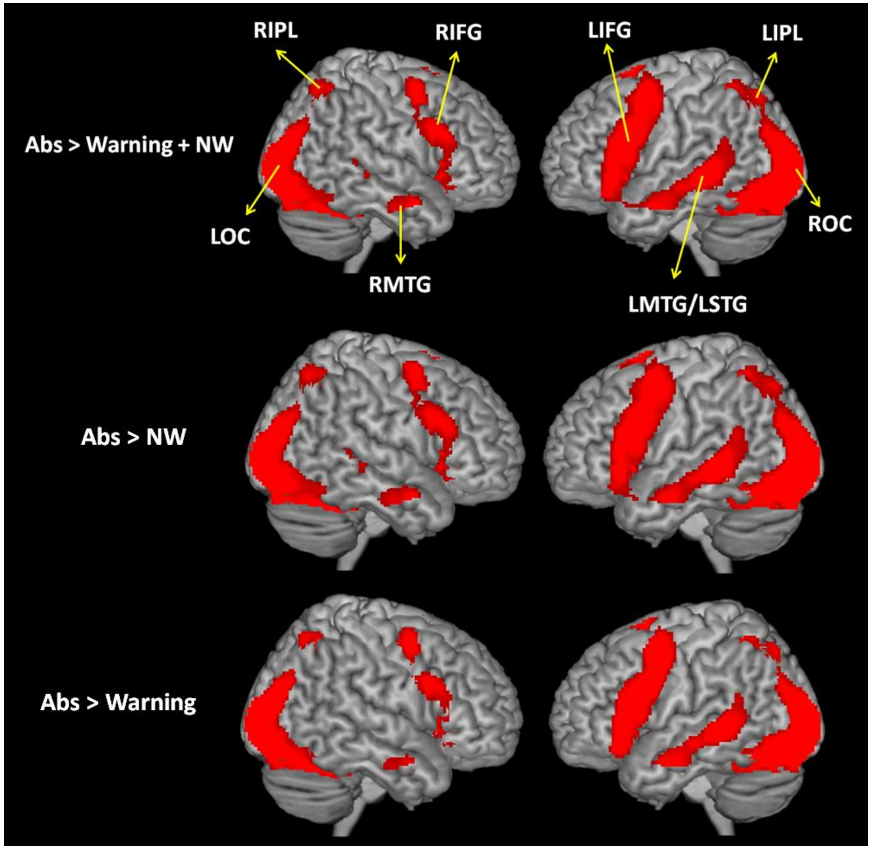
STATISTICS FOR ACCURACY AND RESPONSE TIME – PHISHING EXPERIMENT

Trials	μ_{acc} (σ_{acc})	μ_{time} (σ_{time})
Real	76.68% (18.84%)	3323 ms (1066 ms)
Fake	46.48% (20.58%)	3276 ms (584 ms)
Easy Fake	56.57% (23.29%)	3077 ms (625 ms)
Difficult Fake	33.98% (23.61%)	3538 ms (645 ms)
All	60.42% (13.99%)	3347 ms (654 ms)

- Accuracy of identifying fake websites low.
  - Inline with Dhamija’s study, CHI’06

Data Analysis And Results (6/10)

- Malware Experiment
  - **Contrasts Abstract vs. Warning/ Abstract vs. Non-Warning/Abstract vs. both**



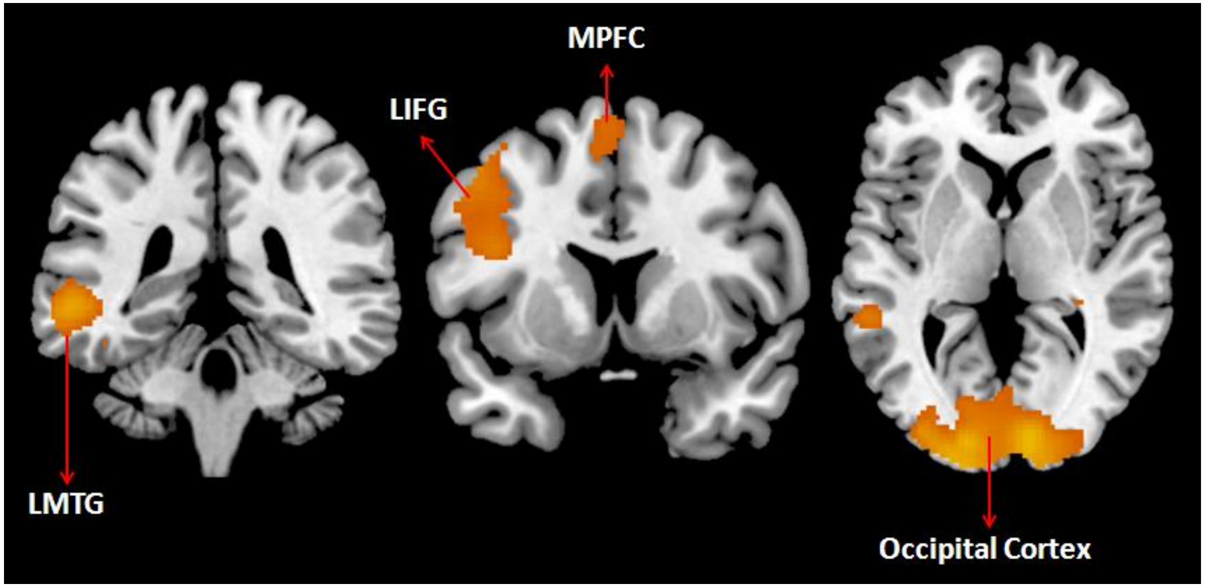
<u>Frontal, temporal and occipital activity</u>

What does it mean?

- Areas associated with Language Comprehension, Visual Processing, Reading
- Participants were potentially reading warnings/non-warnings

Data Analysis And Results (7/10)

- Malware Experiments
  - **Contrasts Warning vs. Non-Warning**



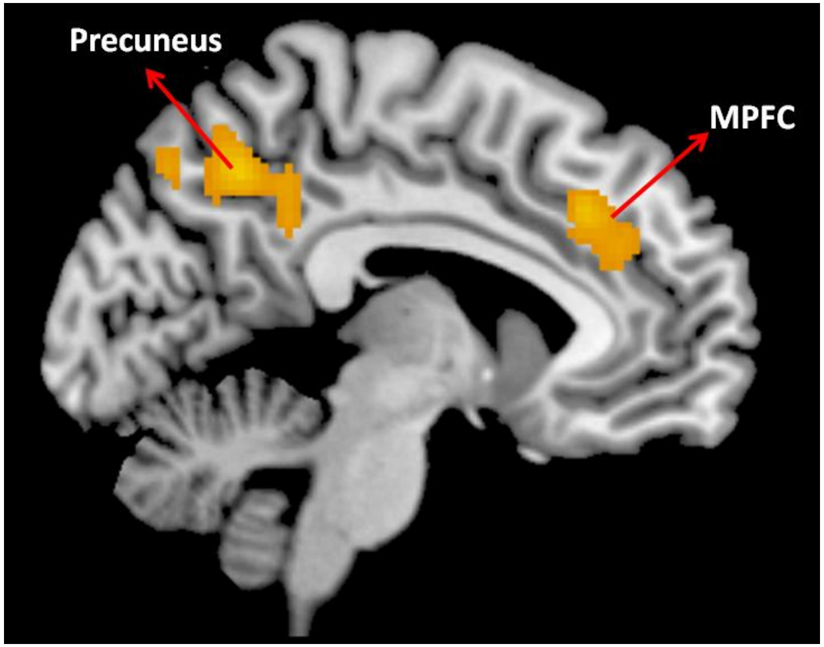
More Activity in Inferior Frontal Middle Temporal and Occipital areas

What does it mean?

- Areas associated with Language Comprehension, Visual Processing, Reading
- Participants were potentially reading through Warnings carefully

Data Analysis And Results (8/10)

- Malware Experiments
  - **Impulsivity as Covariate:**
    - Negative relationship in medial prefrontal cortex (MPFC), more impulsive individuals had less activity in MPFC



Less Activity in Prefrontal Cortex and Precuneus

What does it mean?

- Area associated with decision-making and Problem-Solving
- Tendency to react quickly

Data Analysis And Results (9/10)

- Malware Experiment
  - Behavioral Data Analysis

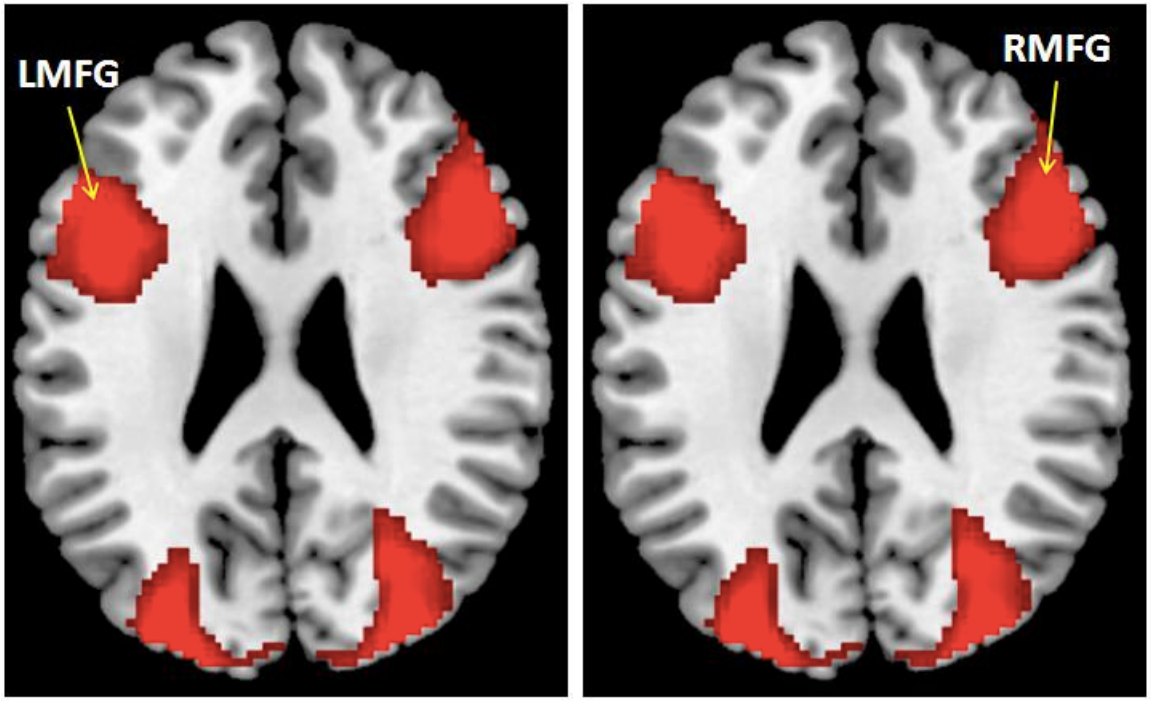
STATISTICS FOR ACCURACY AND RESPONSE TIME – MALWARE EXPERIMENT

Conditions	μ_{acc} (σ_{acc})	μ_{time} (σ_{time})
Non-Warnings	67.49 % (26.57%)	4228 ms (664 ms)
Warnings	88.71% (28.62%)	3715 ms (1141 ms)
All	81.05% (19.59%)	4022 ms (588 ms)

- Participant heeded warnings almost 89% times.
  - Inline with Akhawe-Felt field study
- Response Time for warnings shorter than Non-Warnings

Data Analysis And Results (10/10)

- Cross-Experimental Analysis
  - Phishing vs. Malware



Activity in middle frontal

What does it mean?

- Areas associated with decision making
- One’s ability to heed malware warnings may be associated with his decisions about the legitimacy of a website and vice versa.

Conclusions and Implications

- **Users engage actively in security tasks.**
  - Suggested by activation in different brain areas
  - **Phishing -** behavioral performance was poor despite significant activation in brain regions correlated with higher order cognitive procession
    - Lack of User's Knowledge
  - **Malware Warnings** - the level of brain activation matched with user's good task performance reflected by behavioral data
- **Personality traits impact security**
  - Impulsive individuals showed lower brain activation and may eventually have poor task performance
- **Behavior in one task may potentially be related to another**
  - High degree of correlation in brain activity (with respect to decision making areas) across phishing and malware tasks

Questions ?

Demographics

Table: Participant Demographics Summary

N=25	
Gender	14 Male; 11 female
Age Range	19-32 years
Handedness	24 right-handed; 1 left- handed
Race	13 Caucasian; 5 Hispanic; 6 Asian; 1 African American
Non-Native English Speakers	7
Education Programs	Biology, Music, Athletics, Psychology, Communicational Studies, Physical Education, Biomedical Engineering, pathology , Physical Therapy, Mathematics, Medicine and CS

Pre-Scanning Phase

- Barrat’s Impulsivity Questionnaire
  - Paper and Pencil Test
  - http://www.impulsivity.org/pdf/BIS11English.pdf
  - http://www.impulsivity.org/measurement/bis11
- Edinburgh Handedness form
  - The purpose of the handedness may relate to the lateralization of hemispheric activity in the participants