Real humans, simulated attacks

Usability testing with attack scenarios

Lorrie Faith Cranor
lorrie.cranor.org
@lorrietweet

Let's talk about humans

"Humans are incapable of securely storing high-quality cryptographic keys, and they have unacceptable speed and accuracy when performing cryptographic operations… But they are sufficiently pervasive that we must design our protocols around their limitations."

— C. Kaufman, R. Perlman, and M. Speciner.
Network Security: PRIVATE Communication in a PUBLIC World. 2002.

The human threat

- Malicious humans
- Clueless humans
- Unmotivated humans
- Humans constrained by human limitations

User studies can help us better understand the human threat and design systems that meet user needs

Reasons to conduct user studies

Assess needs

What should we build?

Examine tradeoffs

Which features/approaches best fit particular needs?

Evaluate

Are requirements met?
What should be improved?

Find root causes

What underlying problems need to be fixed?

Excuses for not doing usability studies

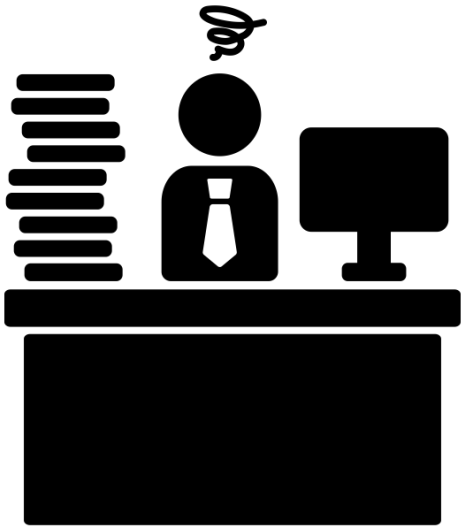
- If people weren't so lazy or stupid or careless it would work fine
- I already know what people want
- No time, no money
- I find the system easy to use
- It's so easy my kids can use it
- I'm not a usability expert

How are security user studies different from other user studies?

Security user studies usually involve the presence of an **adversary**

Need to make sure systems are usable and remain secure when…

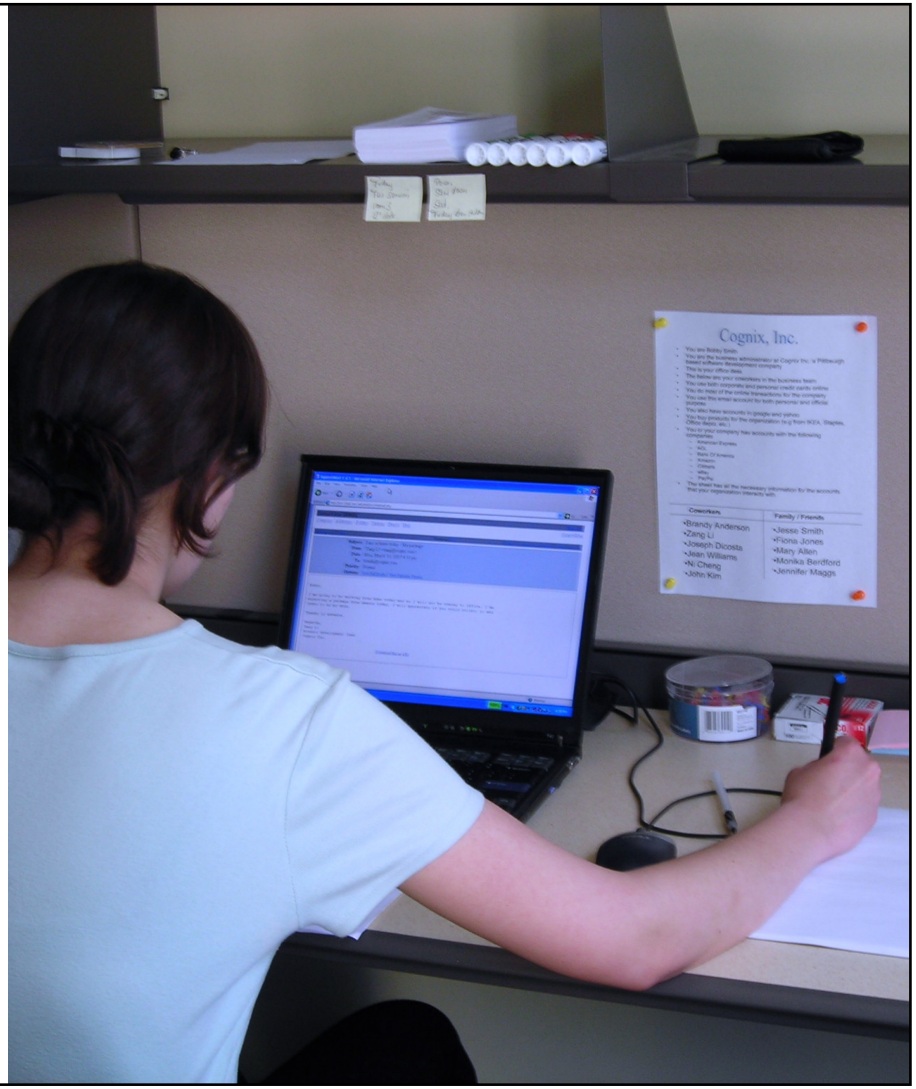
- Attackers (try to) fool users
- Users behave in predictable ways
- Users are unmotivated, careless, stressed, or busy




Security is a
secondary task

Usable security study challenges

- Keeping it real (ecological validity)
- Observing infrequent events and small differences
- Legal, ethical, and practical issues



How can we design a (legal and ethical) study that allows us to observe users in a realistic scenario being exposed to risk?

Observe real world activity

observation of real-world activity

naturally-occurring risk

Many data collection challenges

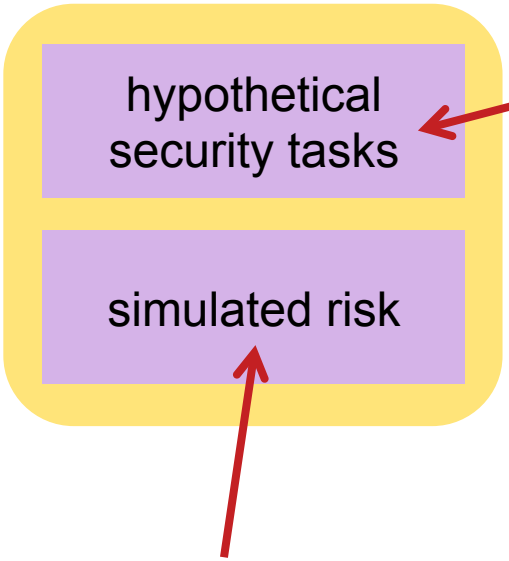
Usually not conducive to a controlled experiment

Events of interest may be infrequent

Observe hypothetical security tasks

Not ethical to harm study participants

Observe hypothetical security tasks



Users may be more alert to security issues than is natural

May use deception + debrief

Observe non-security tasks

real non-security tasks

simulated risk

But users still doing tasks they have been told to do as part of a study

3 approaches to observing users + risk

observation of real-world activity	hypothetical security tasks	real non-security tasks
naturally-occurring risk	simulated risk	simulated risk

See **cups.cs.cmu.edu** for papers!

observation of real-world activity

naturally-occurring risk

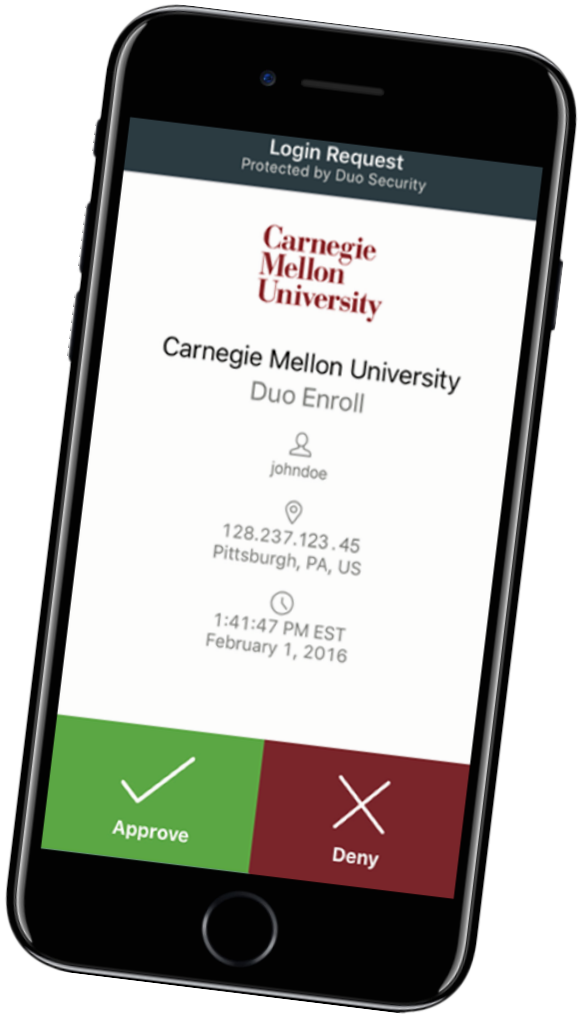
Observing 2fa rollout

observation of real-world activity

naturally-occurring risk

Observing 2fa rollout

- Spring 2017: University began requiring 2fa for employees
- Surveys of students, faculty, and staff as 2fa was being adopted
- Collecting data on problems, help desk tickets, security issues, etc.
- Data collection still underway – preliminary results



Top reasons for adoption/non-adoption

1. Beliefs about need (or lack of need) for security
2. Concerns about extra time and effort
3. Concerns about not having 2nd factor available when they need it
4. Knowledge of other users' (bad) experience

Usability + unintended consequences

- People don’t always have their phones with them
- Students getting locked out of dorm rooms
- Accidental token button pushes cause sync problem

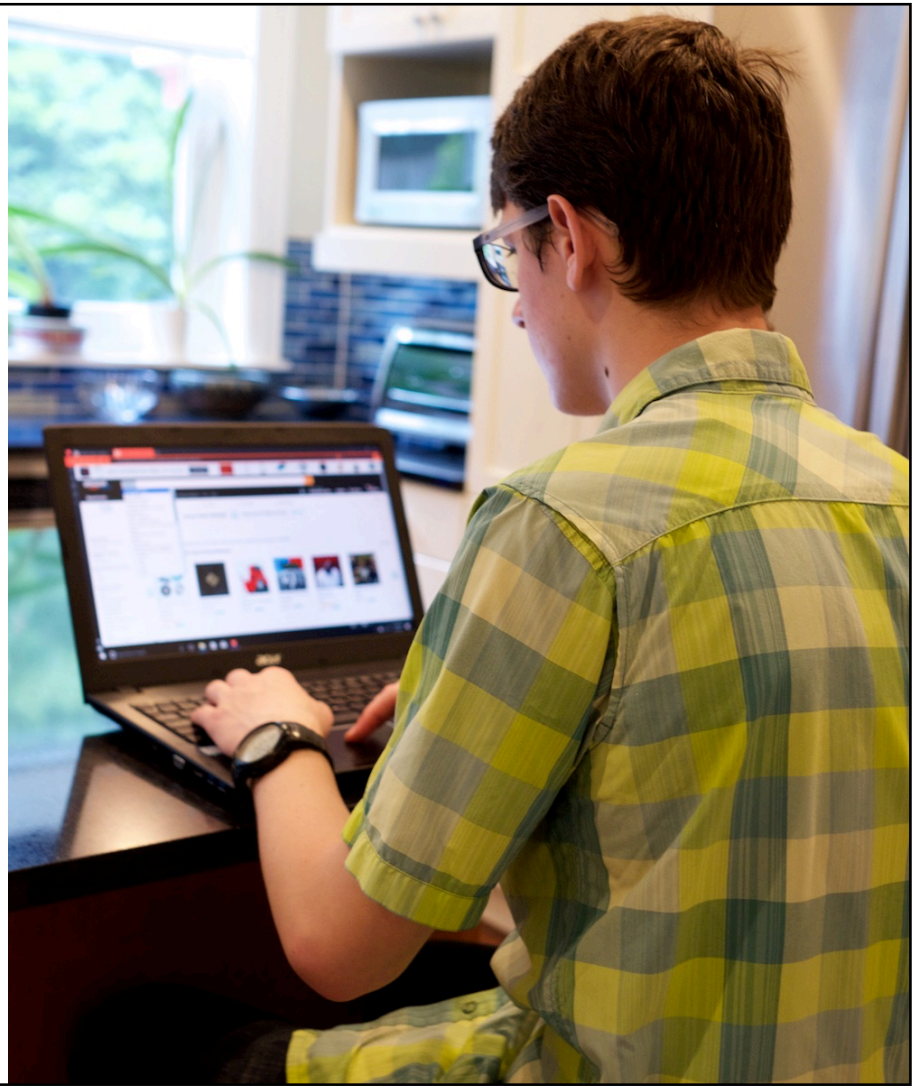
Observing home computer users in their natural habitat

observation of real-world activity

naturally-occurring risk

Security Behavior Observatory (SBO)

- Network of instrumented home Windows computers
- ~200 active participants
- Natural observation + surveys and interviews



Impact of security engagement

- Matched observed security state of computer with self reports about engagement with computer security and maintenance
- Found more security engagement did not always lead to more secure computers

---

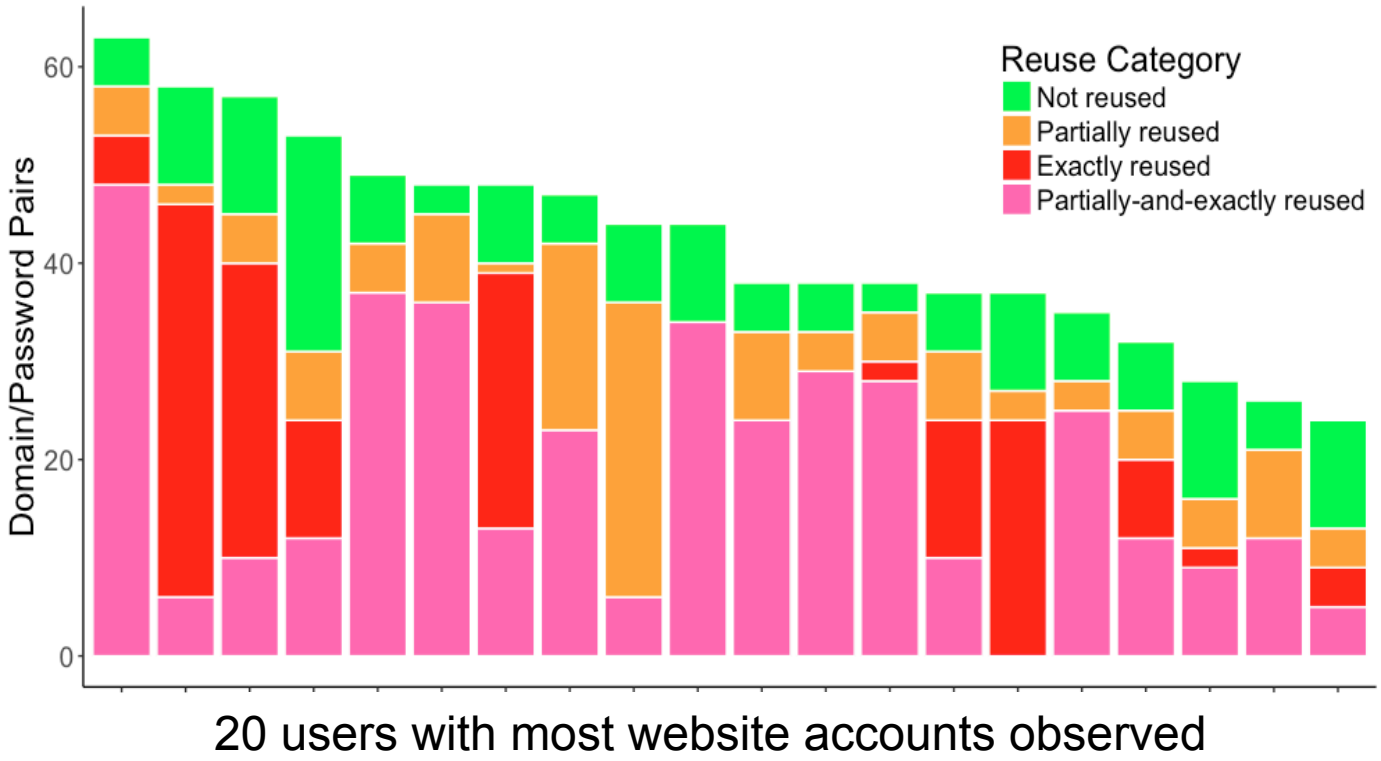
A. Forget, S. Pearman, J. Thomas, A. Acquisti, N. Christin, L. Cranor, S. Egelman, M. Harbach, and R. Telang. Do or Do Not, There Is No Try: User Engagement May Not Improve Security Outcomes. SOUPS 2016.

SBO data related to passwords

- Hashes of passwords and 4+ character substrings
- Length, strength, characters in each class (upper/lowercase, digits, special characters)

How users manage many passwords

Most users reuse passwords exactly and partially



hypothetical security tasks

simulated risk

Comparing usability and secure of password policies

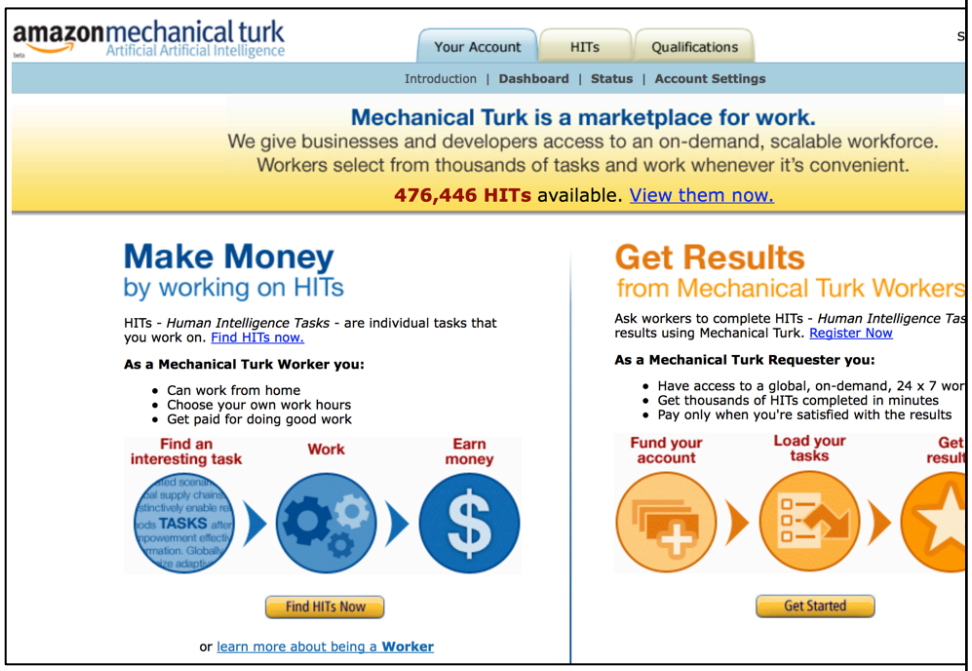
hypothetical security tasks

simulated risk

How can we help users pick passwords that are easy to remember, but hard for an attacker to guess?

Large-scale online experiments

- Amazon Mturk for easy recruitment and payment
- Email participants without collecting personally identifiable information
- 50,000+ participants



See **http://cups.cs.cmu.edu/passwords/** for papers

Participant tasks

- Create password under a randomly assigned condition
- Take a survey
- Recall password
- Return 2 days later to recall password and take survey

Hypothetical security scenario + risk

Imagine that your main email service provider has been attacked, and your account became compromised. You need to create a new password for your email account, since your old password may be known by the attackers. Because of the attack, your email service provider is also changing its password rules.

Password creation task

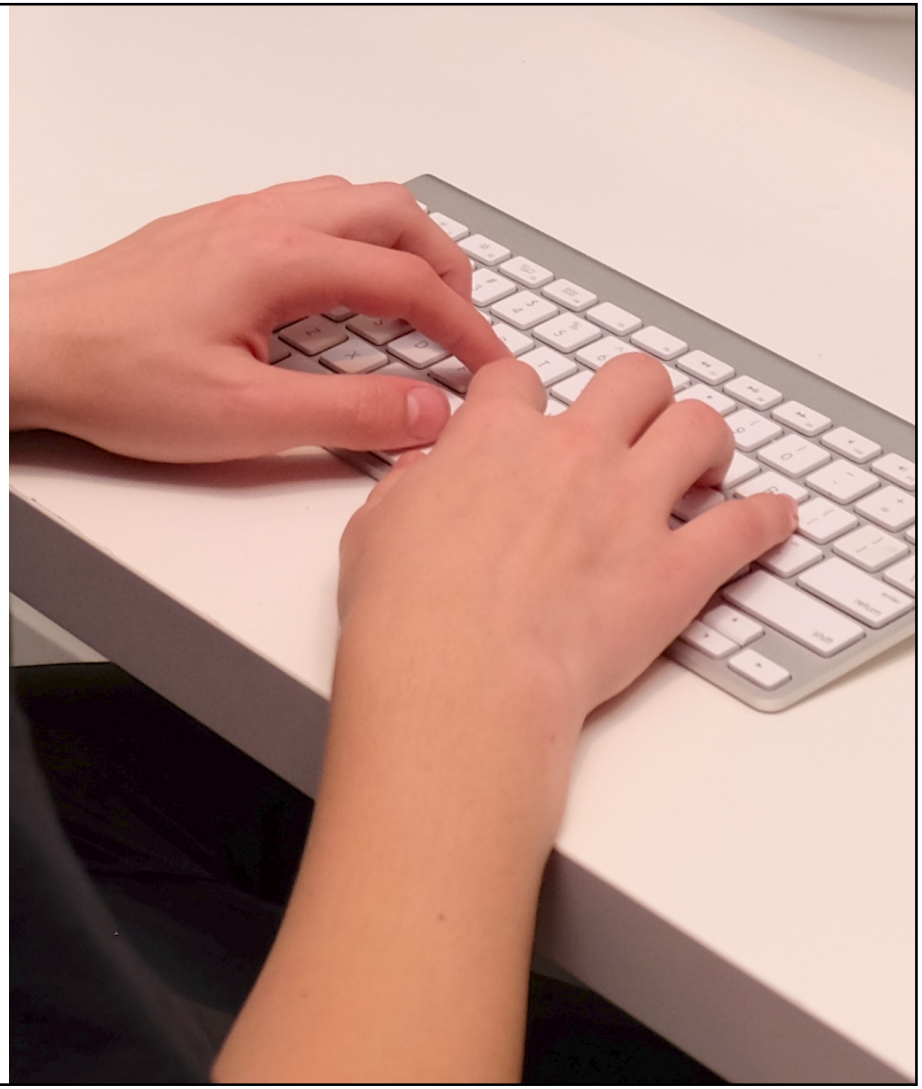
Please follow the instructions below to create a new password for your email account. We will ask you to use this password in a few days to log in again so it is important that you remember your new password.

Request to behave normally

Please take the steps you would normally take to remember your email password and protect this password as you normally would protect the password for your email account. Please behave as you would if this were your real password!

Usability metrics

- Creation attempts and time
- Recall attempts
- Reported sentiment
- Write-down rate
- Study drop-out rate



Password strength metric

Guessability

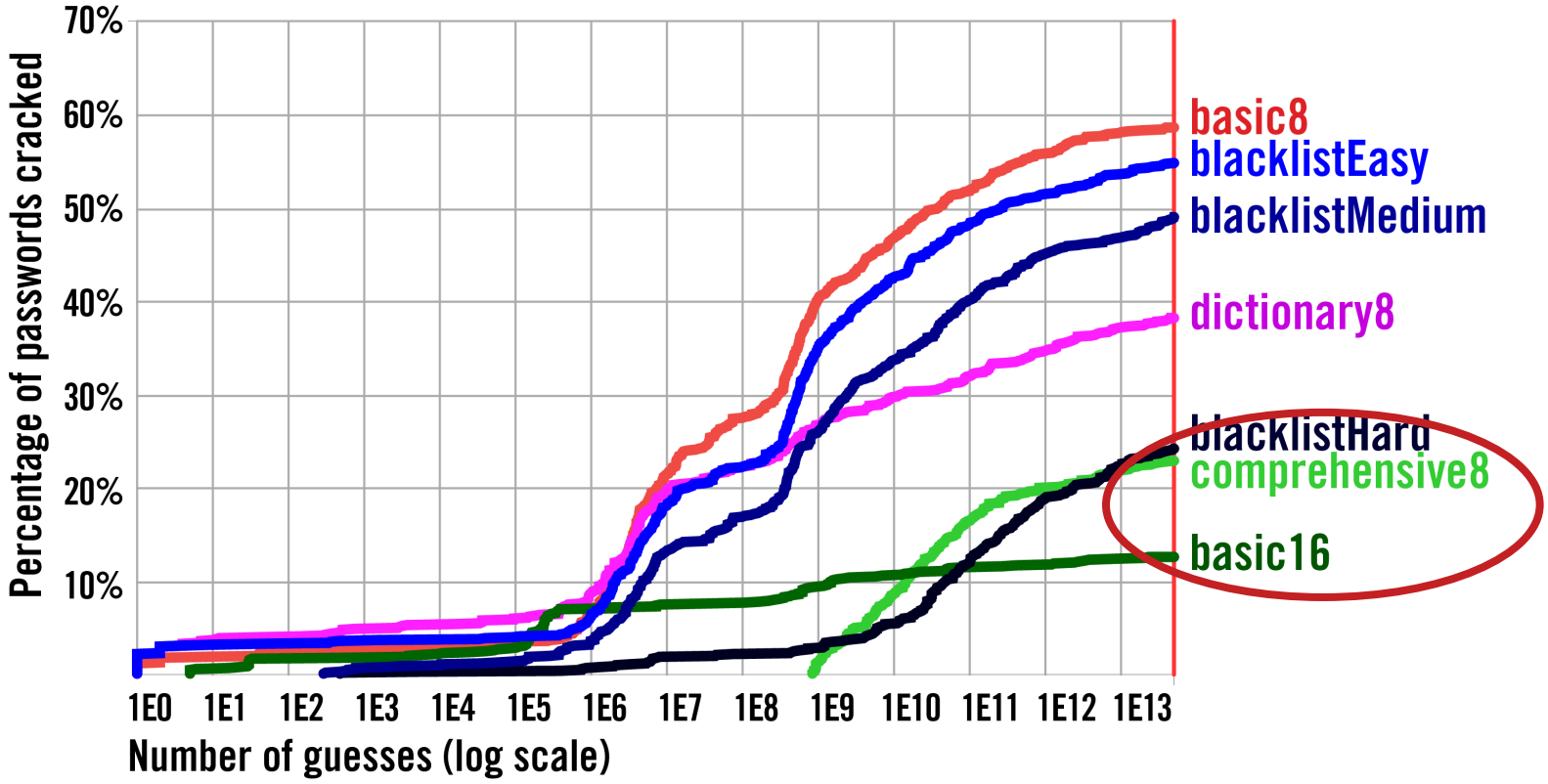
Estimate of how many guesses a sophisticated attacker will need to guess a password

Password policies

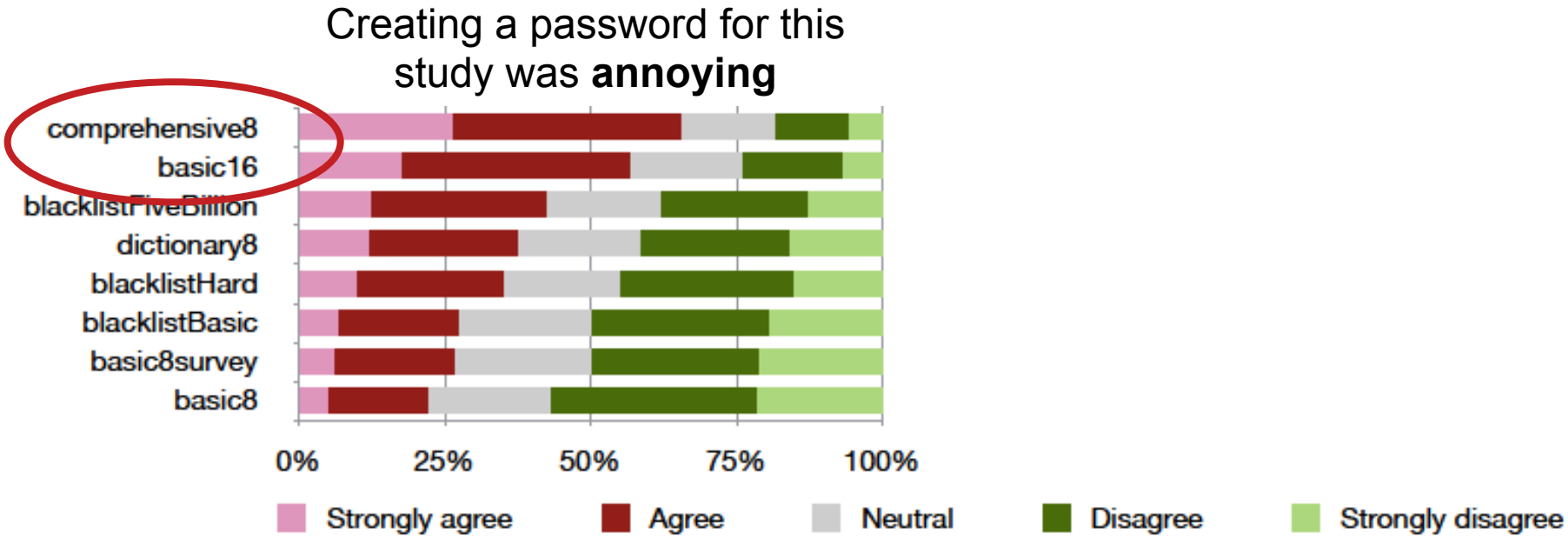
Password policies

Policy	Example password
Basic8	**password**
Dictionary8	**sapsword**
Comprehensive8	**Sapsword1!**
Basic16	**passwordpassword**

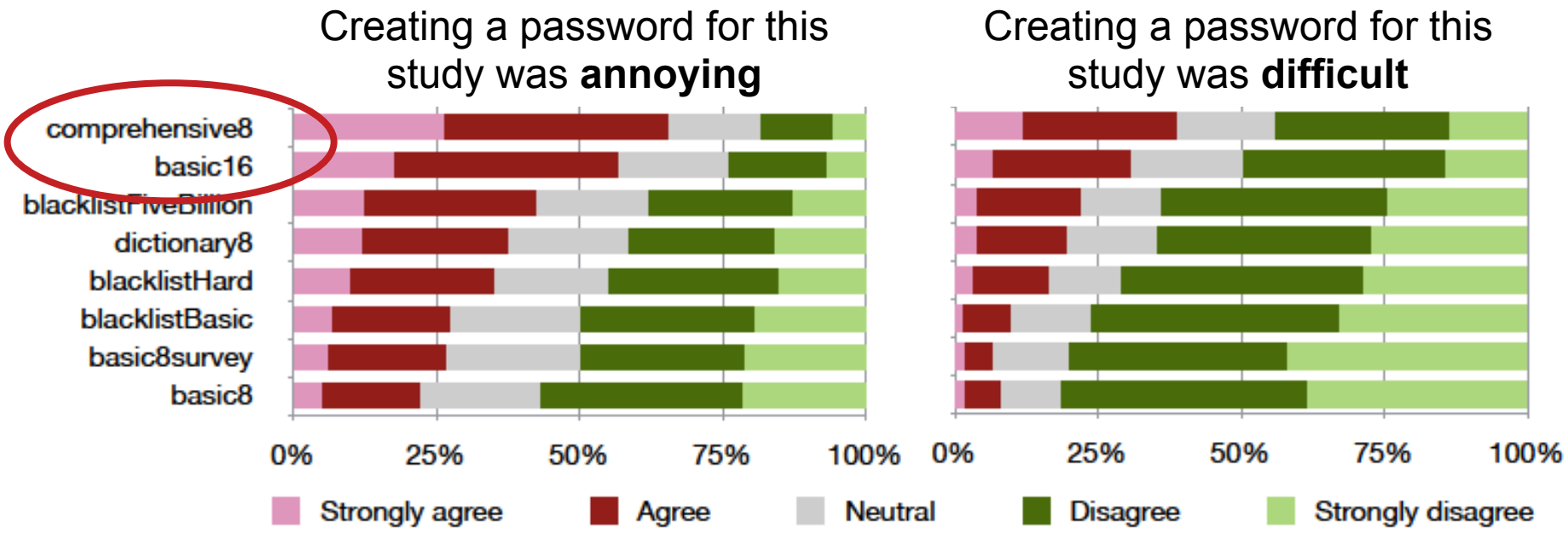
Comparing password policy strength



Comparing password policy usability



Comparing password policy usability



Benefits of this experimental approach

- Learn relative strength and usability of different password policies
  - Change policy with everything else constant
  - Observe all keystrokes while user creates and enters password
- While scenario is hypothetical, passwords are similar to passwords for real accounts

---

M.L. Mazurek, S. Komanduri, T. Vidas, L. Bauer, N. Christin, L.F. Cranor, P.G. Kelley, R. Shay, and B. Ur. Measuring Password Guessability for an Entire University. ACM CCS 2013.

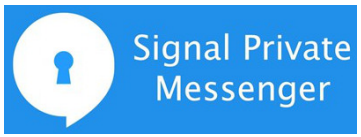
Users' accuracy when comparing crypto key fingerprints

hypothetical security tasks

simulated risk

Secure messaging

- Private communications tools
- Sender needs to reliably obtain recipient's public key to send an encrypted message
- Important to check to make sure you have correct key



Public key → fingerprint

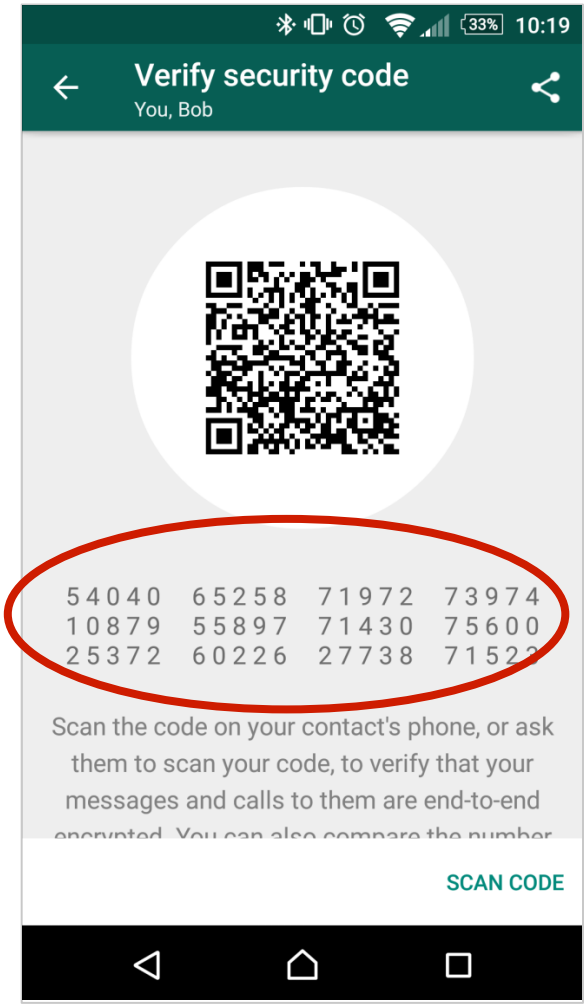
```
-----BEGIN PGP PUBLIC KEY BLOCK-----
Version: SKS 1.1.5
Comment: Hostname: pgp.mit.edu

mQINBFLsrT0BEADI72WmFPt4Q8+3zhtXfxg7MtIilamR0XLk0CSy5jEJk38rLb6Sxr7TCHD1
sD/W/Iy8atV3UA5MUwTZ12iU08MAGW49qmEp9atY7alFtL2p1mGBV0nd8gx0nuLFstGaFIUv
WRVlmeRxiU5zneH2Slt+dgjDsUWMN4nFNnP+87FMI98Q82OdwDai7hXtGKaxLYpzIo9gfFGy
W2x47FXvMxQTC4pUyavkKsv4Q9qfx4cS/Bxv5eezNn/O76b47L/xwJOgCUJILt4udig7RYyI
y8Y0wO5cBwVIfd/XzIig7q0vzEgVCLFnhghyJsguLMjRXa/pCuCAiNkeiqHHwdT3GRHSbGh+
SsUJ6JUcj5nzh5ODpExEGDv1wlncE7DIpwpxM+ct4muVMYqhe6moP6rsOa/aTi+3Jw+Hg80n
FsKlpizCUsAtTFft94tOFZw+uplu+AGPZ8qD1J490V5GZo+7RkUFYxNq/Zt0GAcB+KaW4MTZ
CpDBUJRAnWm/k/n0OYbdjQsTR/Si7cnkLFhQMRN3yaETLsE0WKUYBBmJPug7bhkDEWkF15MJ
dF1N5EQ7Hb1tlFi39zYBhZYMkYEaVviRYAPlVQLOCzVSsS4xUyivRsDRmSX7DLmaW8tY1NwE
8QvJ6mjNQy+V/DdSQf9cMdVu7NMnk8Cb5HOuEgjl9wywm4wWgQARAQABtB5Kb3NodWEgVGFu
IDxqdGFuMTg5QGdtYWlsLmNvbT6JAj0EEwEKACcFAlLsrT0CGwMFCQHhM4AFCwkIBwMFFQoJ
CAsFFgIDAQACHgECF4AACgkQiZDZY75OwYzPaA//aH6+4lN6d1egxPG+NDzcaCPv73gbIxtZ
u19fi9WtVAnLBqGykOHL1Yw+hCH9jFWYfRq8vmiRaRuVQn/7Wf+JcsQway2M7XICeOEg2bPv
uR3eQ50jYyvqEkxSgzoBRp46aSm/9S1wHvwp62C5Hu3Cnjlvb/vFQgWB4tfuyVVjqcpn//Qv
0Jas5SZ6TUid6yLpkFq8U1AQo24Wl2Ns8pfXJoUAfeL0fUoDoQ++0t1V7Zsog7sOIxVXfEyk
…
```

C6C2 78B5 6F92 2B8F 5A07
5B17 69F5 2C6E F103 4425

Key → Fingerprint

Alice wants to verify Bob's fingerprint

- WhatsApp provides numeric fingerprints
- Alice can compare this with fingerprint on Bob's business card or other source



What type of fingerprint is best?

```
8174 5886 6247 7685 4281 4047
0930 1306 7201 2113 8177 9827
```

```
+--[ECDSA  256]---+
|         o o.    |
|          = o    |
|         + . .   |
|          o .    |
|         S .     |
|          o E .  |
|           + o +.|
|          . o * +o|
|           o.++*o.|
+-----------------+
```

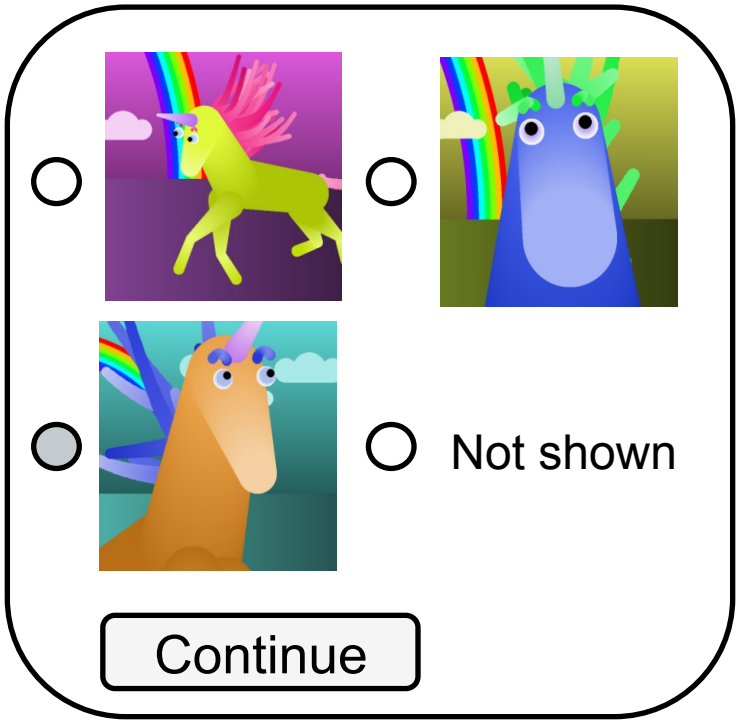
tin yellow blood short attention tax danger bulb wood the normal healthy up false nut bright

buri padi luya kilo yise rada deyu sipi hofe hage xata rite

Joshua Tan, Lujo Bauer, Joseph Bonneau, Lorrie Faith Cranor, Jeremy Thomas, Blase Ur. Can Unicorns Help Users Compare Crypto Key Fingerprints? CHI 2017

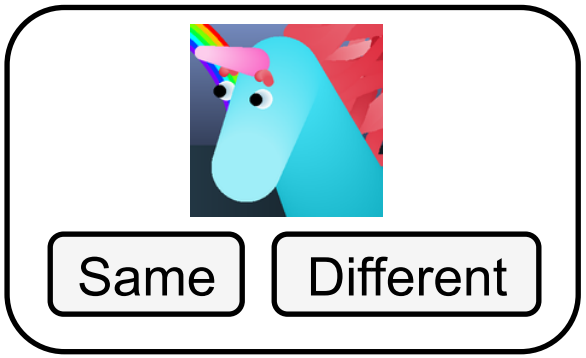
Comparison modes

Compare-and-confirm



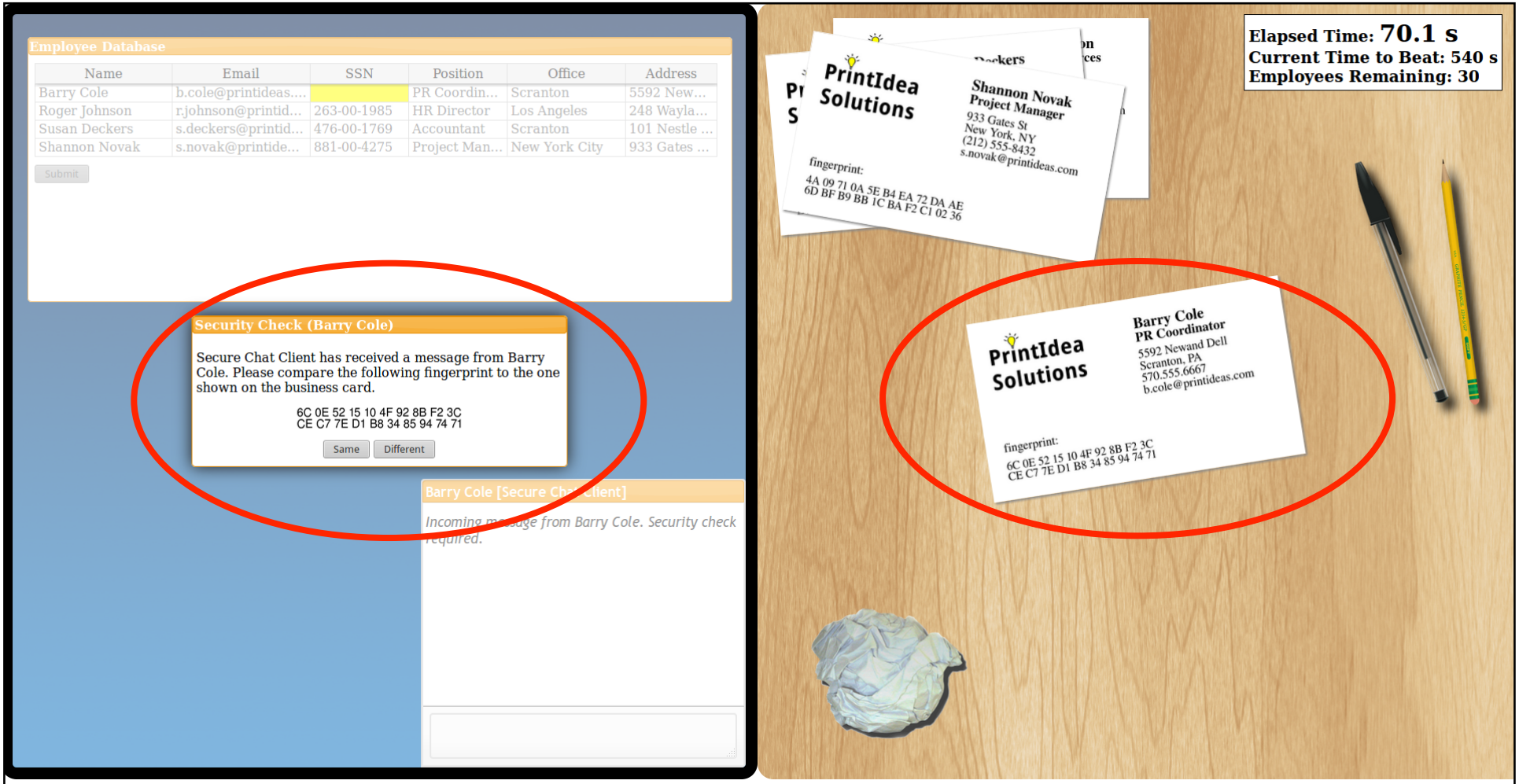
Compare-and-select

Do certain representations and comparison modes lead to more accurate comparisons?



661-participant Mturk experiment

- Participants role-played accountant tasked with updating employee SSNs in database
- For each of 30 employees, required security check involving fingerprint comparison
- Each participant saw 30 fingerprints of same format, **including 1 attack**
- Tested 5 textual formats, 3 graphical formats


Employee Database
Name
Email
SSN
Position
Office
Address
Barry Cole
b.cole@printideas....
PR Coordin...
Scranton
5592 New...
Roger Johnson
r.johnson@printid...
263-00-1985
HR Director
Los Angeles
248 Wayla...
Susan Deckers
s.deckers@printid...
476-00-1769
Accountant
Scranton
101 Nestle ...
Shannon Novak
s.novak@printide...
881-00-4275
Project Man...
New York City
933 Gates ...
Submit
Security Check (Barry Cole)
Secure Chat Client has received a message from Barry Cole. Please compare the following fingerprint to the one shown on the business card.
6C 0E 52 15 10 4F 92 8B F2 3C
CE C7 7E D1 B8 34 85 94 74 71
Same
Different
Barry Cole [Secure Chat Client]
Incoming message from Barry Cole. Security check required.
PrintIdea Solutions
Shannon Novak
Project Manager
933 Gates St
New York, NY
(212) 555-8432
s.novak@printideas.com
fingerprint:
4A 09 71 0A 5E B4 EA 72 DA AE
6D BF B9 BB 1C BA F2 C1 02 36
PrintIdea Solutions
Barry Cole
PR Coordinator
5592 Newand Dell
Scranton, PA
570.555.6667
b.cole@printideas.com
fingerprint:
6C 0E 52 15 10 4F 92 8B F2 3C
CE C7 7E D1 B8 34 85 94 74 71
Elapsed Time: 70.1 s
Current Time to Beat: 540 s
Employees Remaining: 30

Results: people aren't good at this!

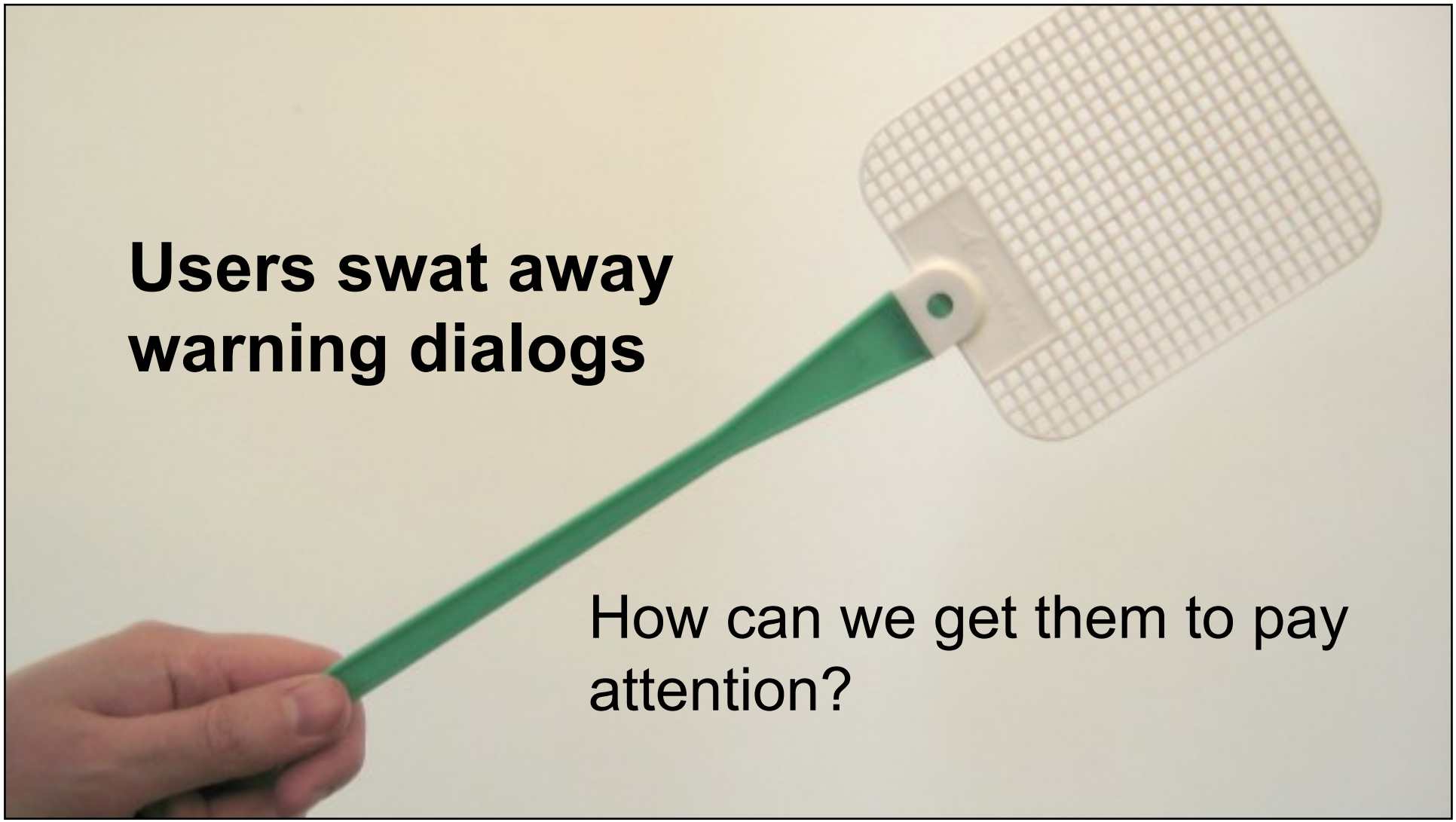
- Compare-and-select caused more mistakes than compare-and-confirm
- Textual formats all had similar missed attack rates
- Graphical formats more varied in attack rates, faster to compare
- Most attacks missed in unicorn condition
- No fingerprints performed very well

real non-security tasks

simulated risk


Security Error: Domain Name Mismatch
You have attempted to establish a connection with "www.whitehouse.gov". However, the security certificate presented belongs to "a248.e.akamai.net". It is possible, though unlikely, that someone may be trying to intercept your communication with this web site.
If you suspect the certificate shown does not belong to "www.whitehouse.gov", please cancel the connection and notify the site administrator.
View Certificate
Cancel
OK

Image courtesy of Johnathan Nightingale


Users swat away warning dialogs
How can we get them to pay attention?

Study design challenges

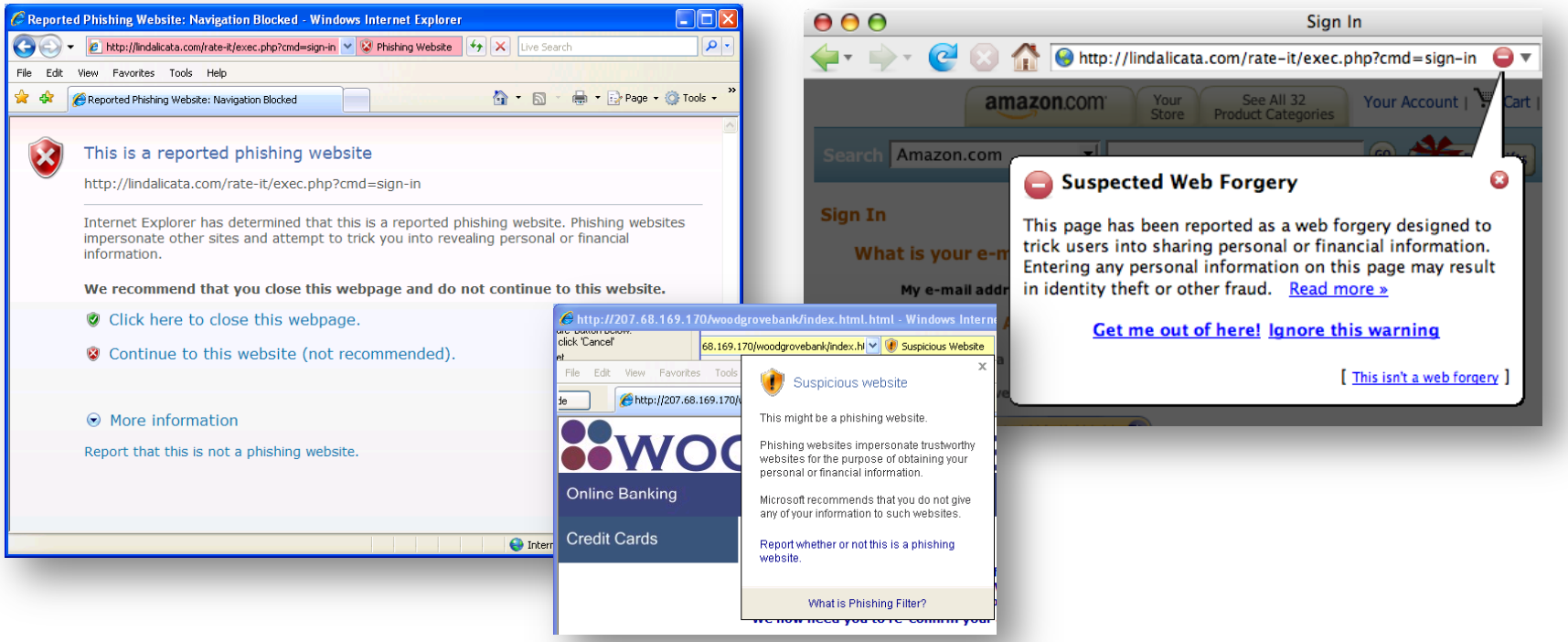
- Observe users interacting with warnings without them knowing we're interested in warnings
- Make users feel like they are experiencing an attack without actually putting them at risk

Evaluating phishing warnings

real non-security tasks

simulated risk

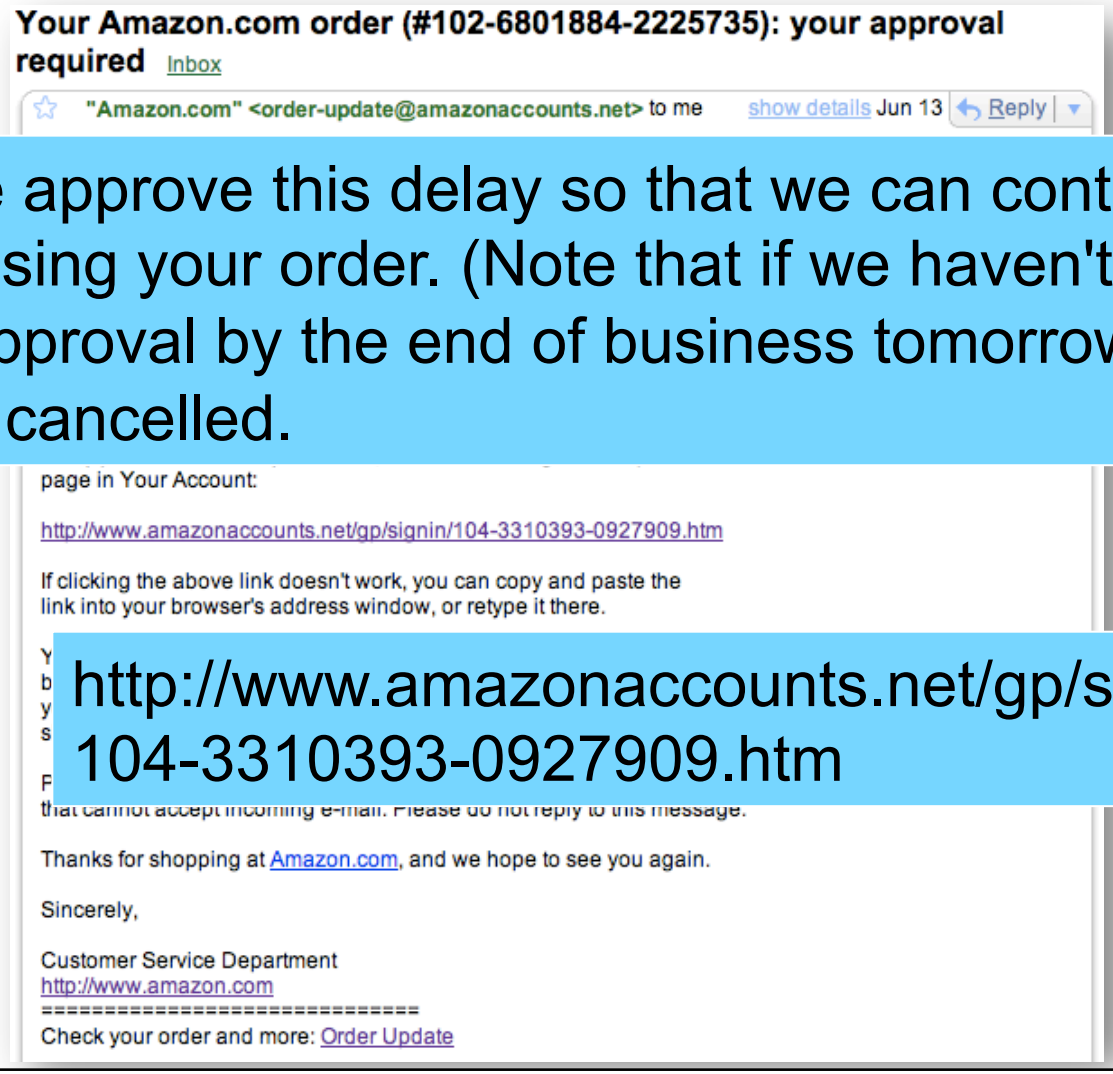
Browser phishing warning study



S. Egelman, L. Cranor, and J. Hong. You've Been Warned: An Empirical Study of the Effectiveness of Web Browser Phishing Warnings. CHI 2008.

Required a little deception

- Lab study on online shopping
- Purchase paper clips from Amazon
- Answer questions about shopping (for another study)
- That's when we phished them
- Check email to get your receipt
- That's when they fell for it

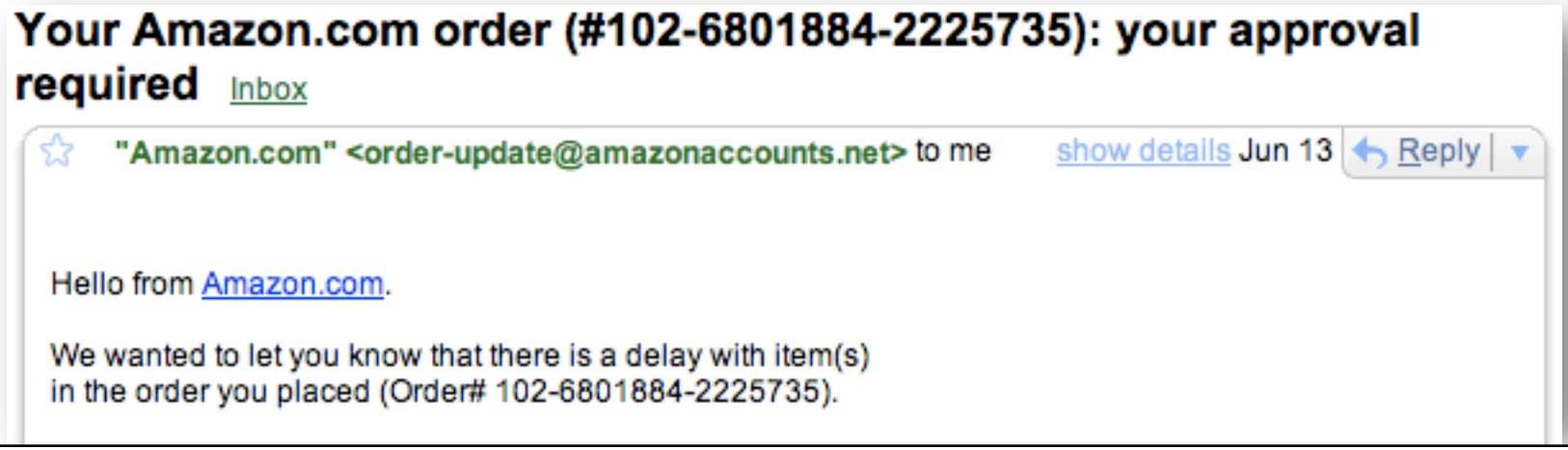

Your Amazon.com order (#102-6801884-2225735): your approval required Inbox
"Amazon.com" <order-update@amazonaccounts.net> to me show details Jun 13 Reply
Please approve this delay so that we can continue processing your order. (Note that if we haven't received your approval by the end of business tomorrow, the item will be cancelled.
page in Your Account:
http://www.amazonaccounts.net/gp/signin/104-3310393-0927909.htm
If clicking the above link doesn't work, you can copy and paste the link into your browser's address window, or retype it there.
http://www.amazonaccounts.net/gp/signin/104-3310393-0927909.htm
Thanks for shopping at Amazon.com, and we hope to see you again.
Sincerely,
Customer Service Department
http://www.amazon.com
==========
Check your order and more: Order Update

Success!

- Most participants got phished
- Significant differences between conditions
- Observed interesting user behavior that helped us understand root cause of failures

Confused by domain names

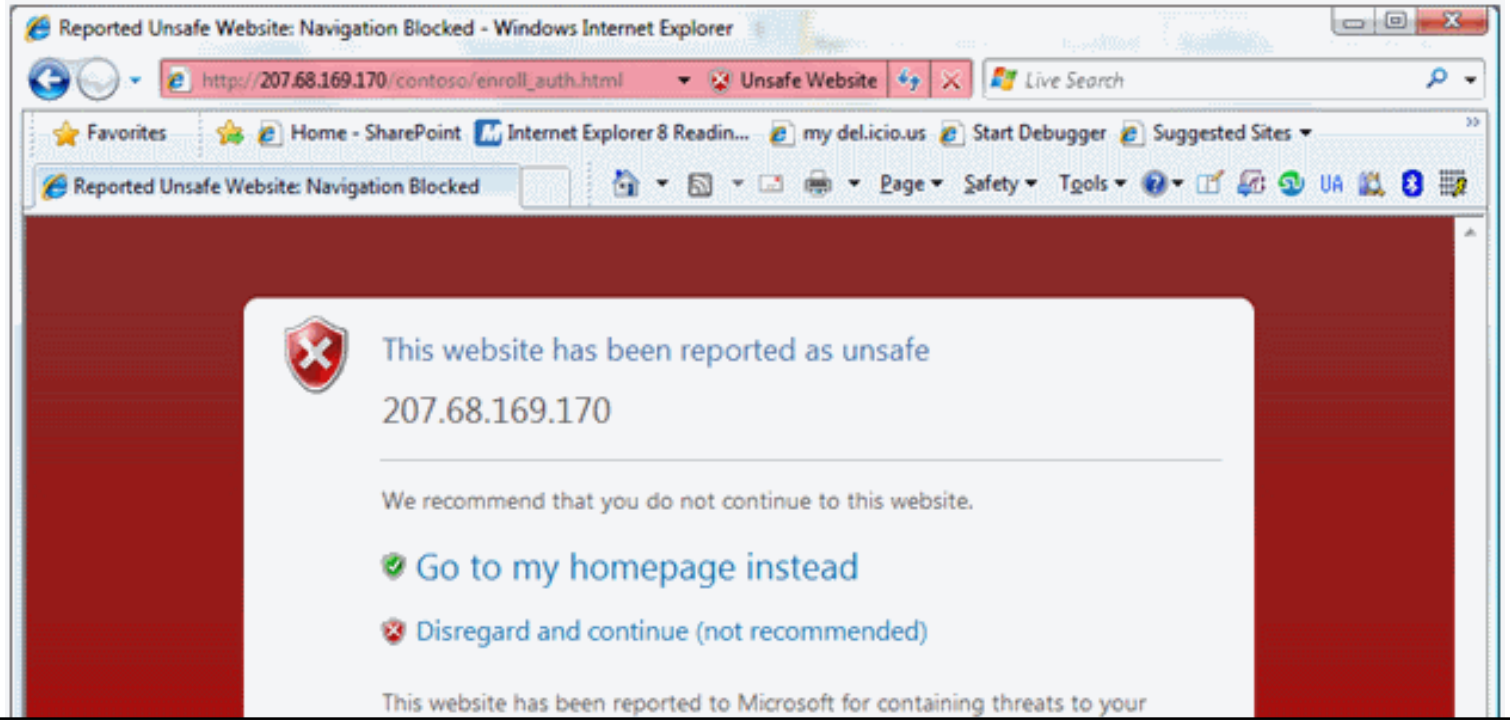
“The address in the browser was of amazonaccounts.net which is a genuine address”



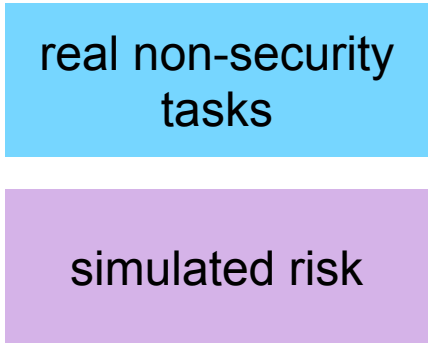
Confused mental models

Some users repeatedly closed their browser, returned to the phishing email, and clicked on the link again

Research led to better phishing warnings



Attracting attention to key information



Some hazards are ALWAYS dangerous

Some hazards are context dependent

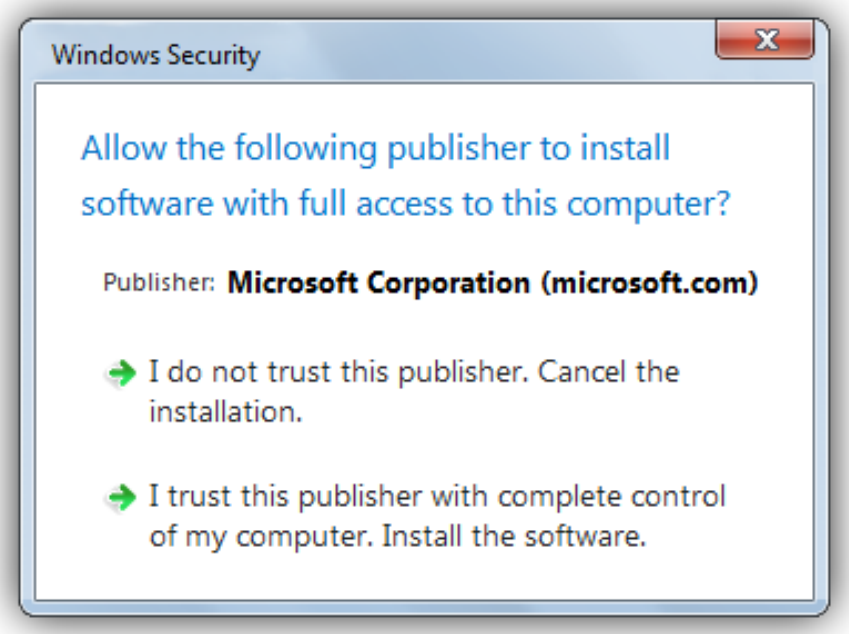
Security dialogs context dependent

- Security warning dialogs more like warnings on wine than warnings on poison
- Software developers place burden of assessing risk on users

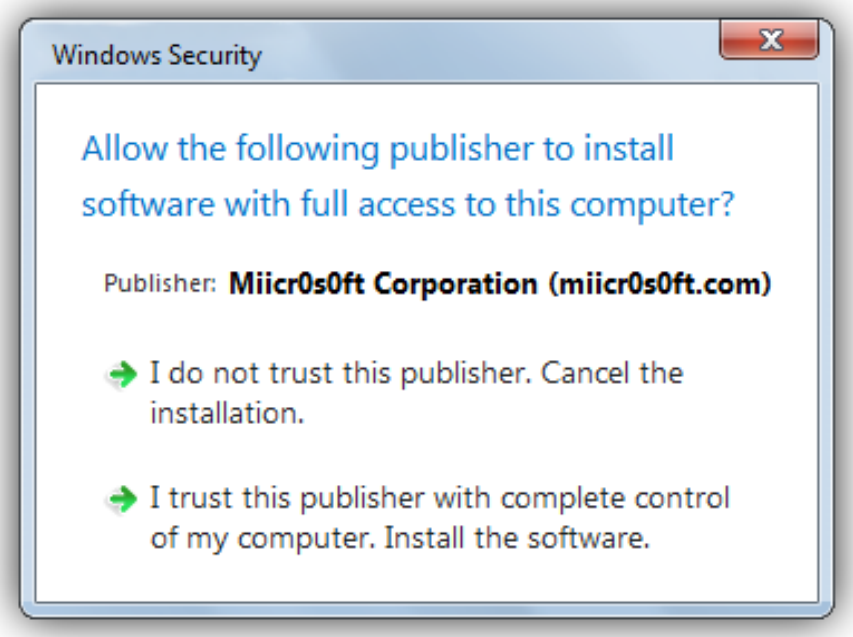
A good warning helps users determine whether they are at risk

- Stops users from doing something dangerous in risky context
- Doesn’t interfere with non-risky contexts
- Need to test warnings in both contexts

Can you spot the suspicious software?



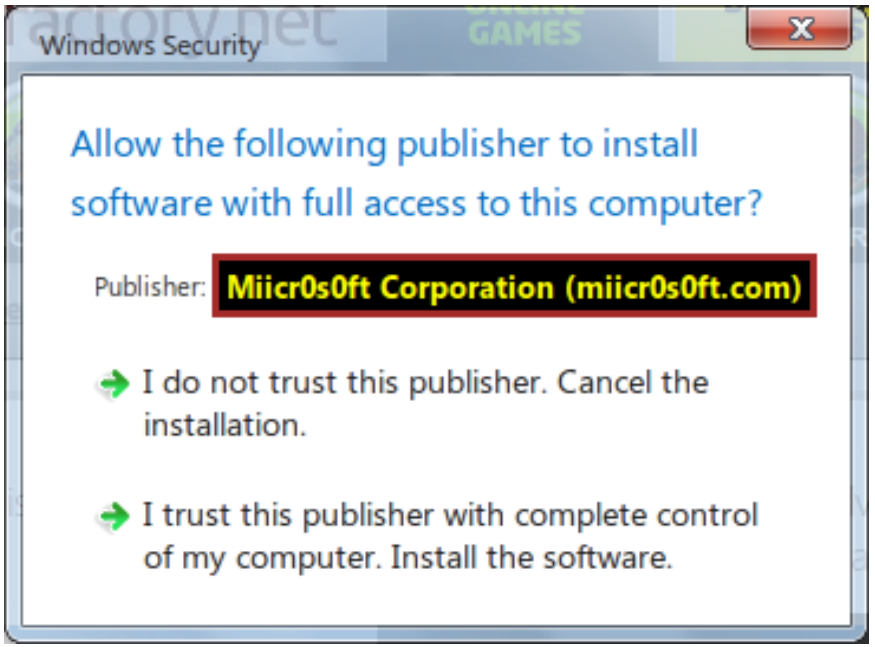
benign



suspicious

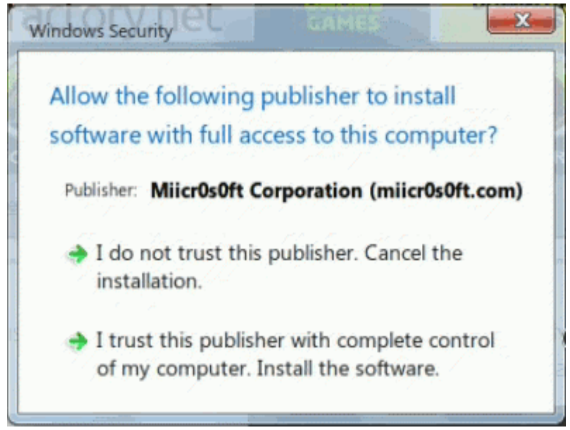
Attracting users' attention

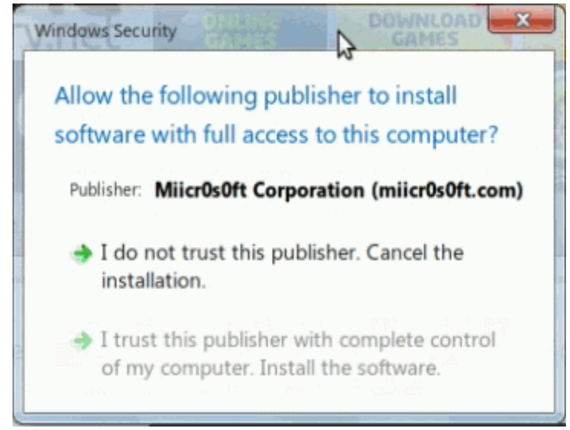
How can we focus users' attention on key information they need to make informed decisions?



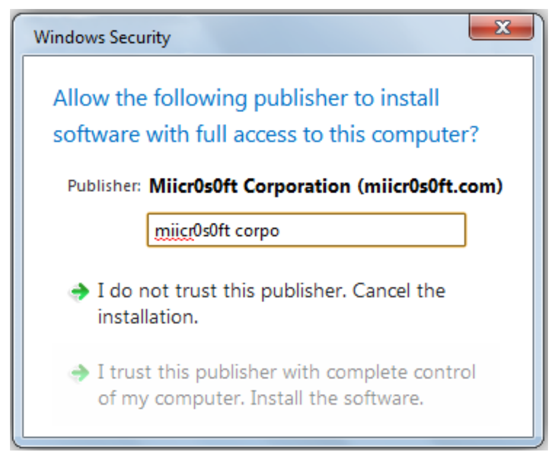
C. Bravo-Lillo, L.F. Cranor, J. Downs, S. Komanduri, R.W. Reeder, S. Schechter, and M. Sleeper. Your Attention Please: Designing security-decision UIs to make genuine risks harder to ignore. SOUPS 2013.


Windows Security
Allow the following publisher to install software with full access to this computer?
Publisher: Miicr0s0ft Corporation (miicr0s0ft.com)
I do not trust this publisher. Cancel the installation.
I trust this publisher with complete control of my computer. Install the software.


Windows Security
Allow the following publisher to install software with full access to this computer?
Publisher: Miicr0s0ft Corporation (miicr0s0ft.com)
I do not trust this publisher. Cancel the installation.
I trust this publisher with complete control of my computer. Install the software.


Windows Security
Allow the following publisher to install software with full access to this computer?
Publisher: Miicr0s0ft Corporation (miicr0s0ft.com)
I do not trust this publisher. Cancel the installation.
I trust this publisher with complete control of my computer. Install the software.


Windows Security
Allow the following publisher to install software with full access to this computer?
Publisher: Miicr0s0ft Corporation (miicr0s0ft.com)
I do not trust this publisher. Cancel the installation.
I trust this publisher with complete control of my computer. Install the software.

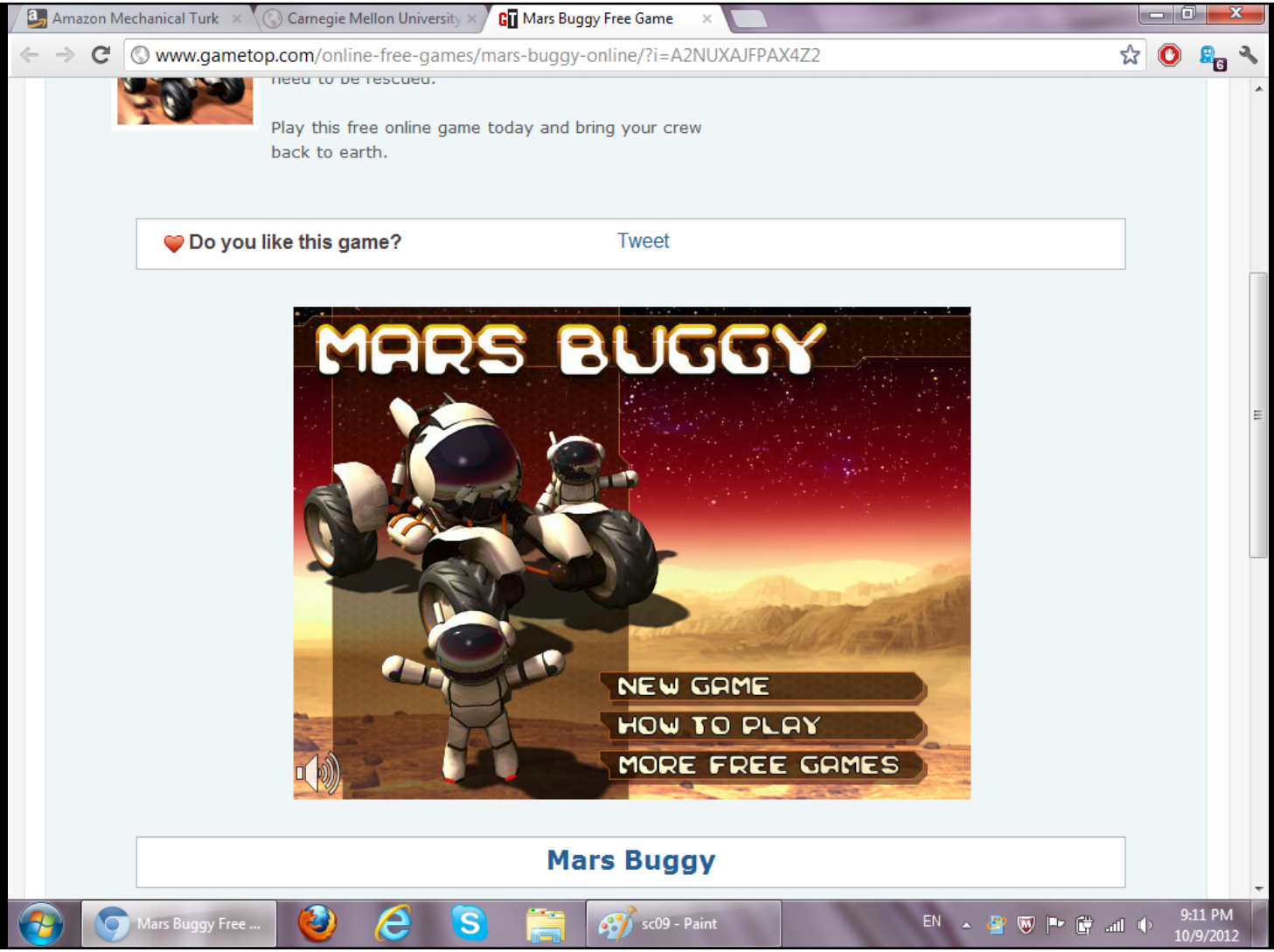

Windows Security
Allow the following publisher to install software with full access to this computer?
Publisher: Miicr0s0ft Corporation (miicr0s0ft.com)
miicr0s0ft corpo
I do not trust this publisher. Cancel the installation.
I trust this publisher with complete control of my computer. Install the software.

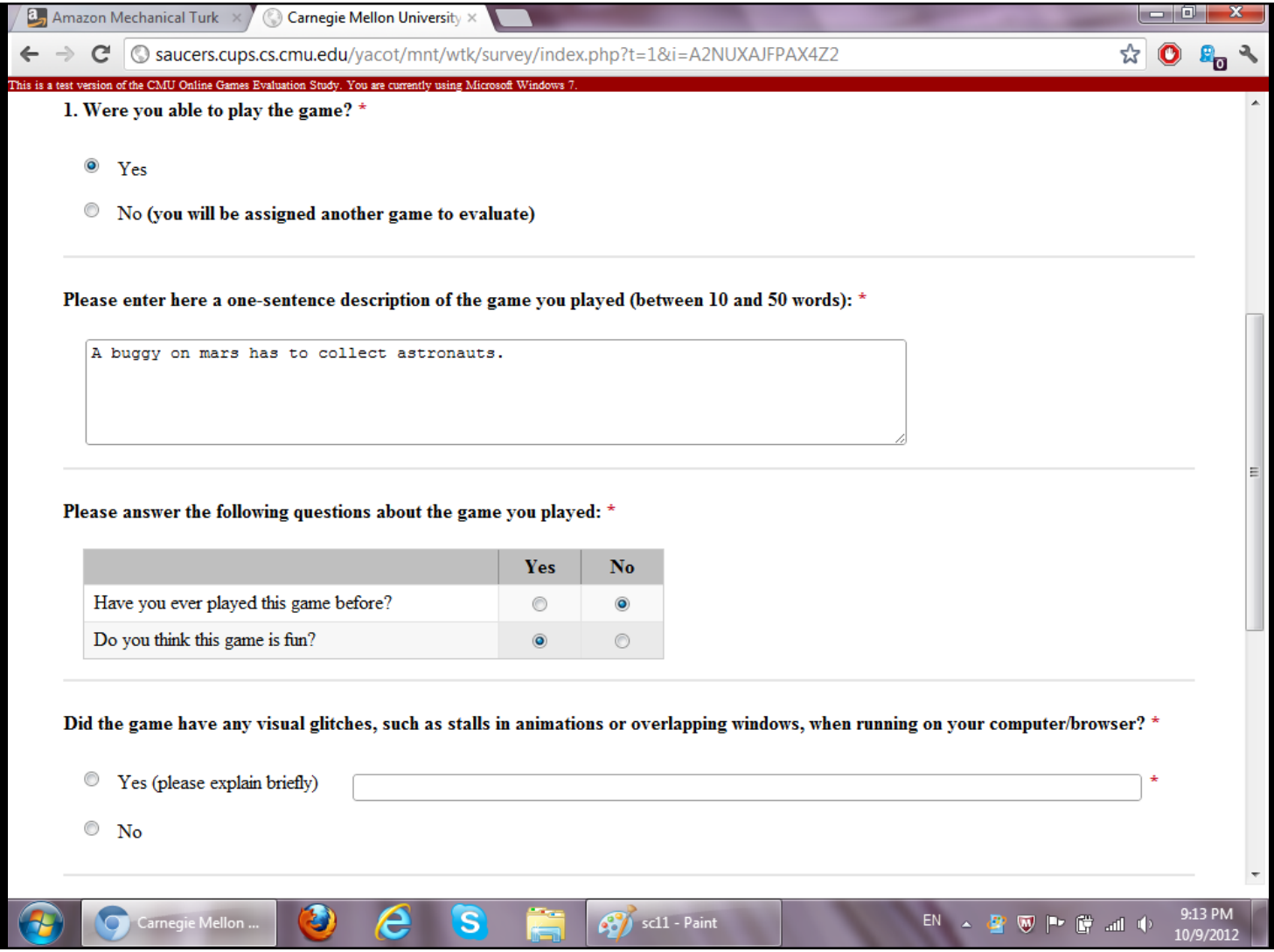
Do any of these work?

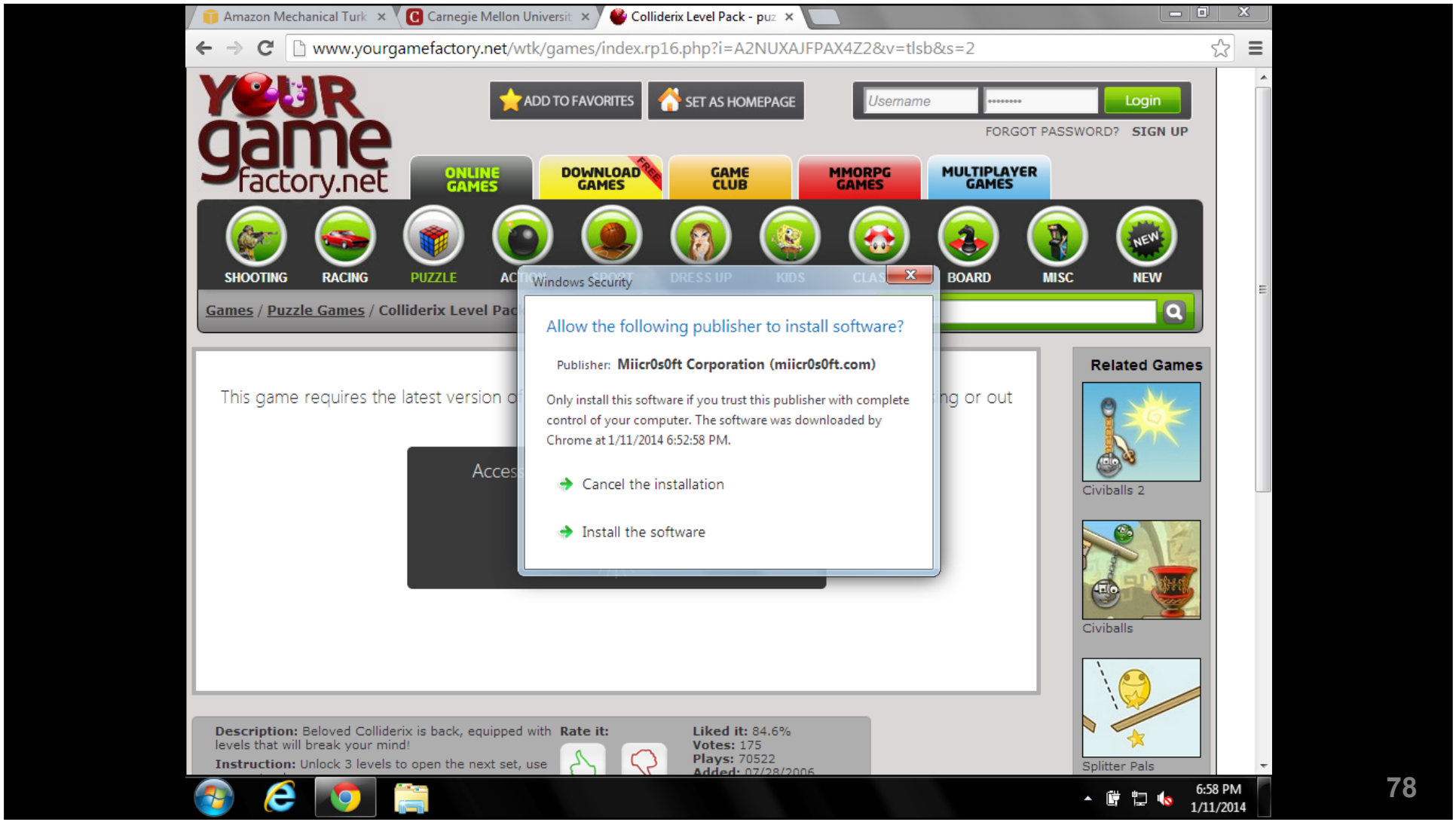
- Do attractors and other techniques prevent suspicious installs without preventing benign installs?
- How much do attractors delay benign installs?

Methodology requirements

- Massive, inexpensive, quick
- Remote observation/recording of behavior
- Participants should feel safety/risk and behave as they would in real life
- But should not actually be at increased risk through participation in experiment


Amazon Mechanical Turk
Carnegie Mellon University
Mars Buggy Free Game
www.gametop.com/online-free-games/mars-buggy-online/?i=A2NUXAJFPAX4Z2
need to be rescued.
Play this free online game today and bring your crew back to earth.
Do you like this game?
Tweet
MARS BUGGY
NEW GAME
HOW TO PLAY
MORE FREE GAMES
Mars Buggy
Mars Buggy Free ...
sc09 - Paint
EN
9:11 PM
10/9/2012

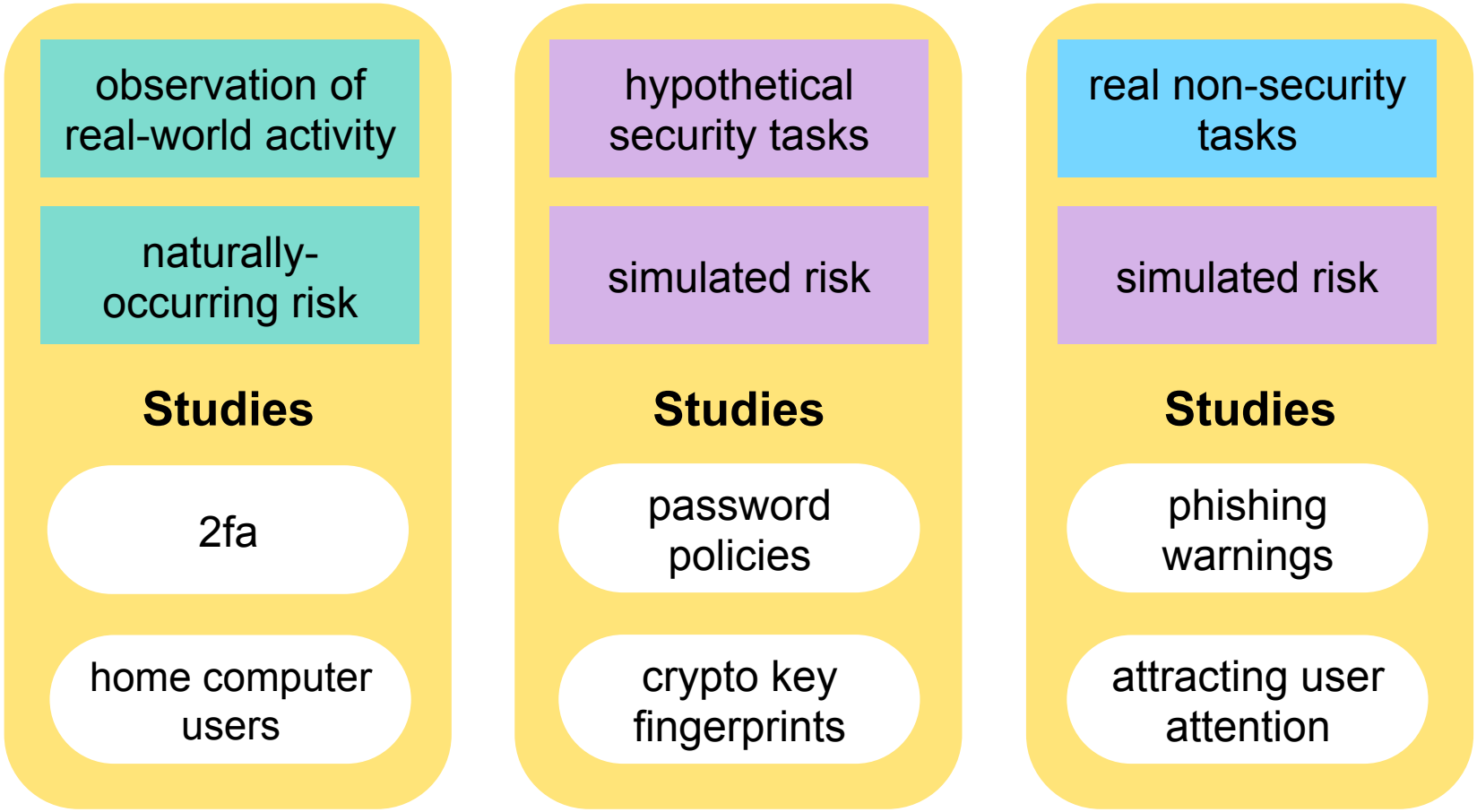

Amazon Mechanical Turk
Carnegie Mellon University
saucers.cups.cs.cmu.edu/yacot/mnt/wtk/survey/index.php?t=1&i=A2NUXAJFPAX4Z2
This is a test version of the CMU Online Games Evaluation Study. You are currently using Microsoft Windows 7.
1. Were you able to play the game? *
Yes
No (you will be assigned another game to evaluate)
Please enter here a one-sentence description of the game you played (between 10 and 50 words): *
A buggy on mars has to collect astronauts.
Please answer the following questions about the game you played: *
Yes
No
Have you ever played this game before?
Do you think this game is fun?
Did the game have any visual glitches, such as stalls in animations or overlapping windows, when running on your computer/browser? *
Yes (please explain briefly) *
No
Carnegie Mellon ...
sc11 - Paint
EN
9:13 PM
10/9/2012


Amazon Mechanical Turk
Carnegie Mellon Universit
Colliderix Level Pack - puz
www.yourgamefactory.net/wtk/games/index.rp16.php?i=A2NUXAJFPAX4Z2&v=tlsb&s=2
YOUR game factory.net
ADD TO FAVORITES
SET AS HOMEPAGE
Username
Login
FORGOT PASSWORD? SIGN UP
ONLINE GAMES
DOWNLOAD GAMES
GAME CLUB
MMORPG GAMES
MULTIPLAYER GAMES
SHOOTING
RACING
PUZZLE
BOARD
MISC
NEW
Games / Puzzle Games / Colliderix Level Pac
Windows Security
Allow the following publisher to install software?
Publisher: Miicr0s0ft Corporation (miicr0s0ft.com)
Only install this software if you trust this publisher with complete control of your computer. The software was downloaded by Chrome at 1/11/2014 6:52:58 PM.
Cancel the installation
Install the software
This game requires the latest version
Related Games
Civiballs 2
Civiballs
Splitter Pals
Description: Beloved Colliderix is back, equipped with levels that will break your mind!
Instruction: Unlock 3 levels to open the next set, use
Rate it:
Liked it: 84.6%
Votes: 175
Plays: 70522
6:58 PM
1/11/2014

Results are encouraging

- 2,227 Mturk participants encountered dialogs
- New dialogs reduced installations in suspicious scenario without preventing benign installations
- Some dialogs slowed people down
- Swipe, type, and delay were particularly effective
- Follow-up study: Swipe and type remained effective after many exposures

Review and wrap-up


observation of real-world activity
naturally-occurring risk
Studies
2fa
home computer users
hypothetical security tasks
simulated risk
Studies
password policies
crypto key fingerprints
real non-security tasks
simulated risk
Studies
phishing warnings
attracting user attention

Black hat sound bytes

- Don't assume you know how humans will behave – do a study!
- Observe real world activity if you can
- Otherwise, observe realistic scenarios under simulated risk

Real humans, simulated attacks

Usability testing with attack scenarios

Lorrie Faith Cranor
lorrie.cranor.org
@lorrietweet