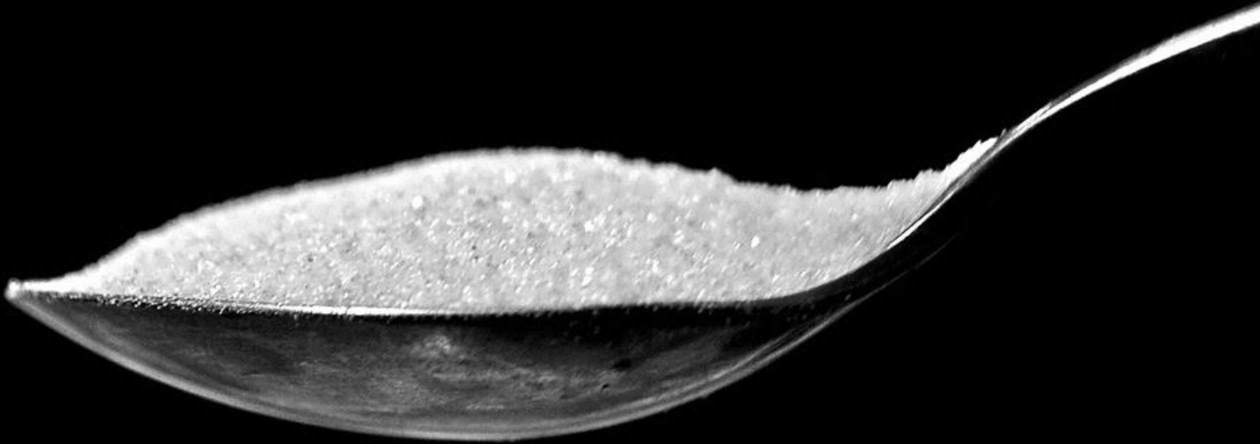
A Spoonful of Sugar?

The Impact of Guidance and Feedback on Password-Creation Behavior

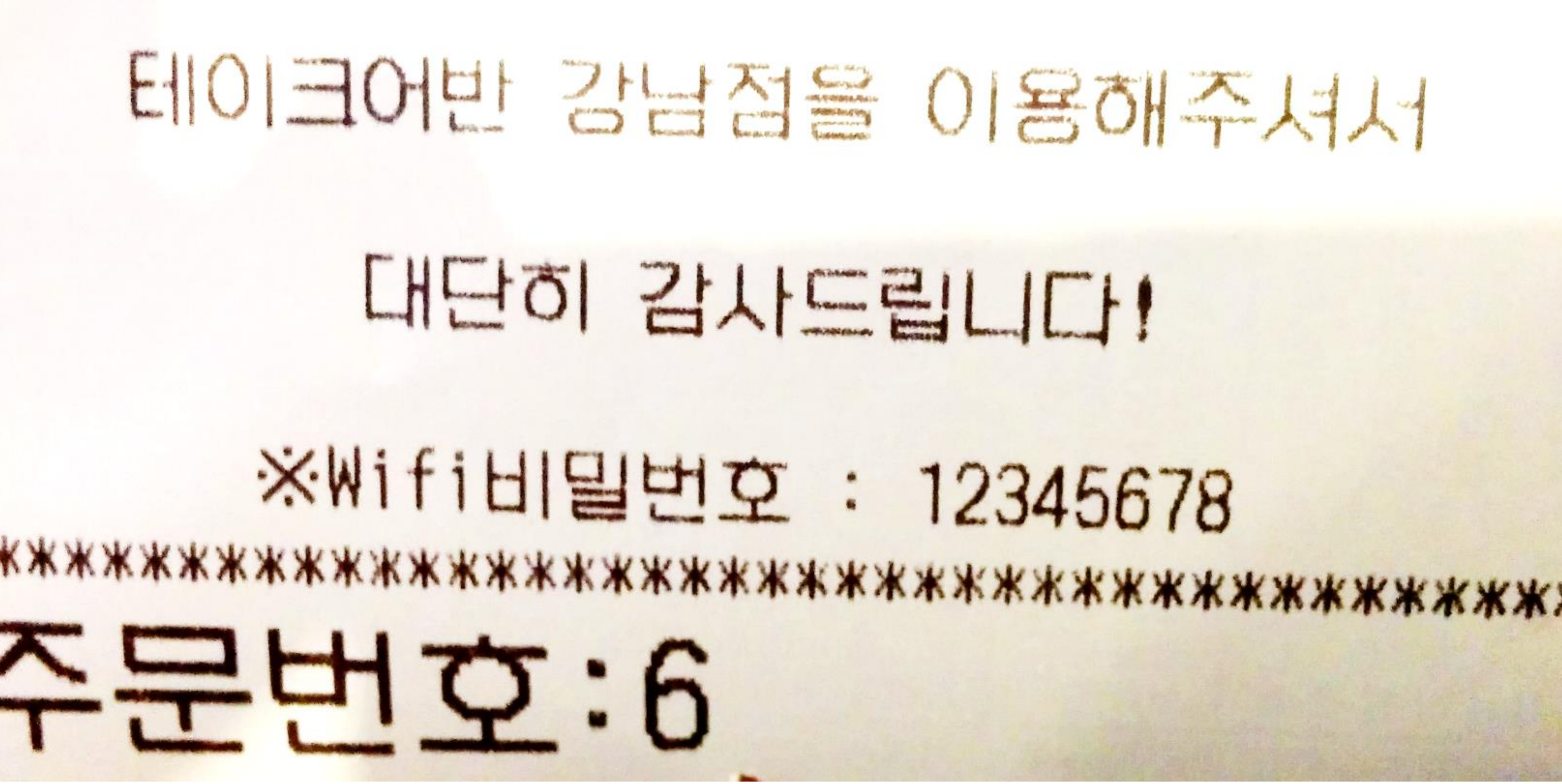


Richard Shay, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, Alain Forget, Saranga Komanduri, Michelle L. Mazurek, William Melicher, Sean M. Segreti, Blase Ur

Carnegie Mellon

UNIVERSITY OF MARYLAND

123456789
12345
password
iloveyou
princess
rockyou
1234567
12345678
abc123
nicole
daniel
babygirl
monkey
jessica
654321
michael
ashley
qwerty
111111
000000
michelle
tigger
sunshine
chocolate
password1
soccer
anthony
friends
butterfly
purple
angel
jordan
liverpool
justin
loveme
football
lovely
basketball
superman
secret
pretty
playboy
666666
jasmine
angels
andrew
tinkerbell
andrea
carlos
flower
eminem
spongebob
jennifer
iloveu
lovers
bubbles
hottie
cookie
dragon
hannah
samantha
shadow
melissa
summer
christian
matthew
elizabeth
brandon
chelsea
barbie
danielle
hello
softball
poohbear
mickey
jesus
joseph
computer
whatever
thomas
robert
forever
beautiful
charlie
america
rangers
monica
teamo
orange
rainbow
william
tweety
sakura
princess1
123123
7777777
121212
naruto
snoopy
carolina
destiny
rachel
junior
patrick
taylor
chicken
sweety
heaven
pokemon
batman
miguel



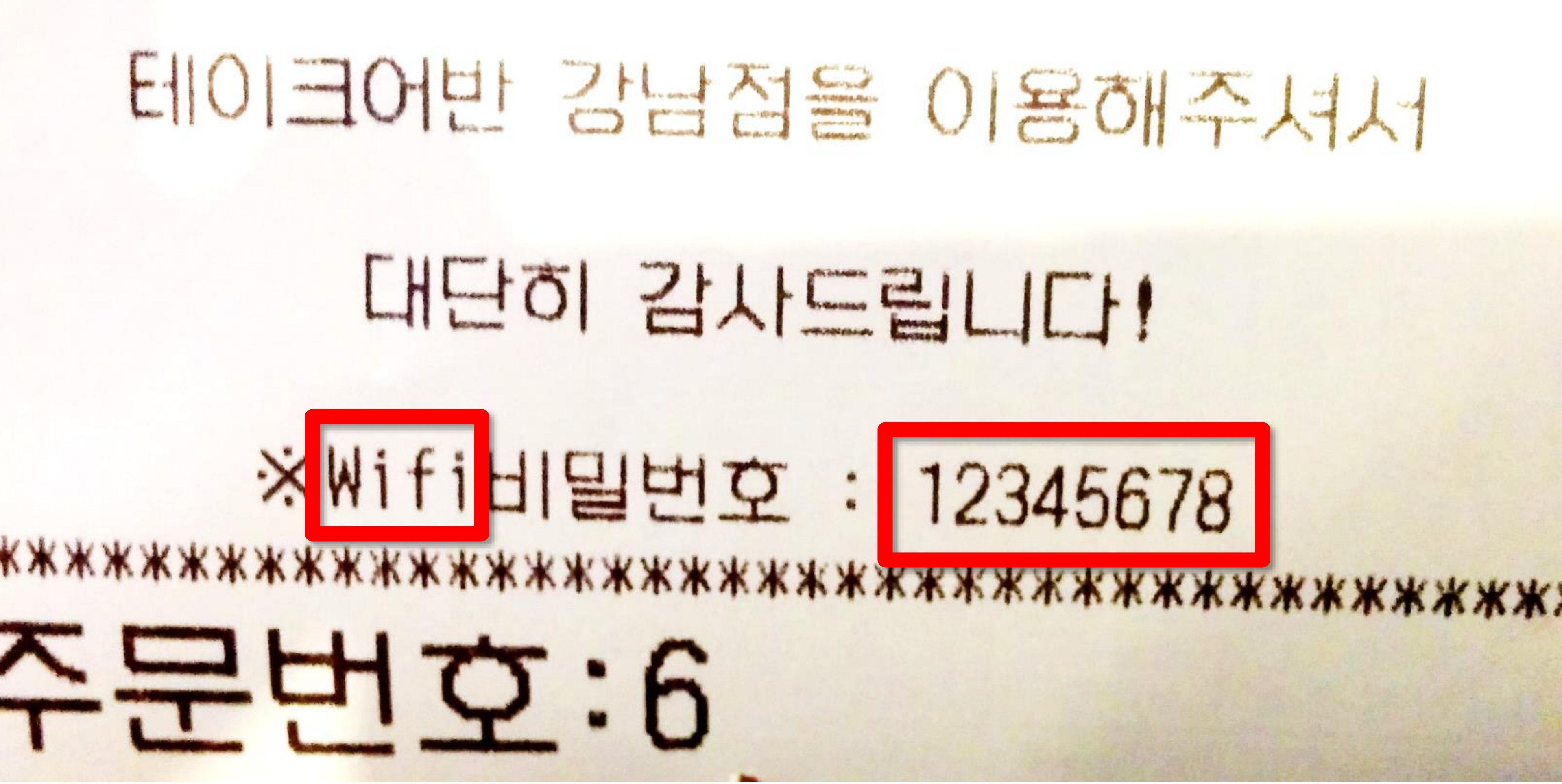
테이크어반 강남점을 이용해주셔서

대단히 감사드립니다!

※Wifi비밀번호 : 12345678

****

주문번호:6



테이크어반 강남점을 이용해주셔서

대단히 감사드립니다!

※Wifi 비밀번호 : 12345678

****

주문번호:6

Password Breaches Remain A Threat

Password Breaches Remain A Threat

Password-Composition Policies

Password-Composition Policies

- Your password must…
  - …contain 12 or more characters
  - …contain at least 3 of the following character classes: {lowercase letters, uppercase letters, digits, symbols}

Shay et al. “Can Long Passwords be Secure and Usable?” In Proc. CHI 2014
Komanduri et al. “Of Passwords and People…” In Proc. CHI 2011

Password-Composition Policies

- Your password must…
  - …contain 12 or more characters
  - …contain at least 3 of the following character classes: {lowercase letters, uppercase letters, digits, symbols}
  - …not contain a **blacklisted** substring (e.g., "1234")

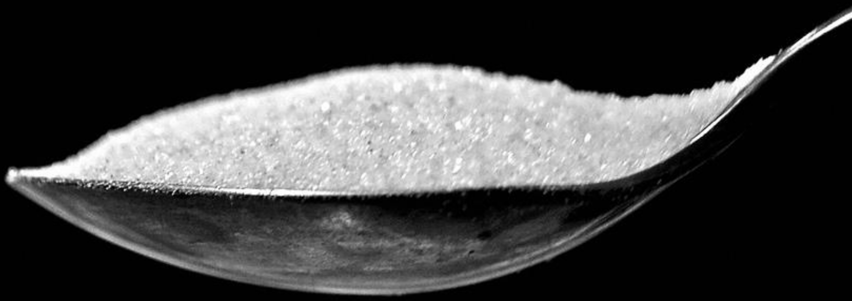
Shay et al. "Can Long Passwords be Secure and Usable?" In Proc. CHI 2014
Komanduri et al. "Of Passwords and People…" In Proc. CHI 2011

Password-Composition Policies

- Your password must…
  - …contain 12 or more characters
  - …contain at least 3 of the following character classes: {lowercase letters, uppercase letters, digits, symbols}
  - …not contain a **blacklisted** substring (e.g., "1234")
  - …start and end with a lowercase letter (**pattern**)

Shay et al. "Can Long Passwords be Secure and Usable?" In Proc. CHI 2014
Komanduri et al. "Of Passwords and People…" In Proc. CHI 2011

Images Creative Commons on Flickr by *carowallis* and *woodythrower*

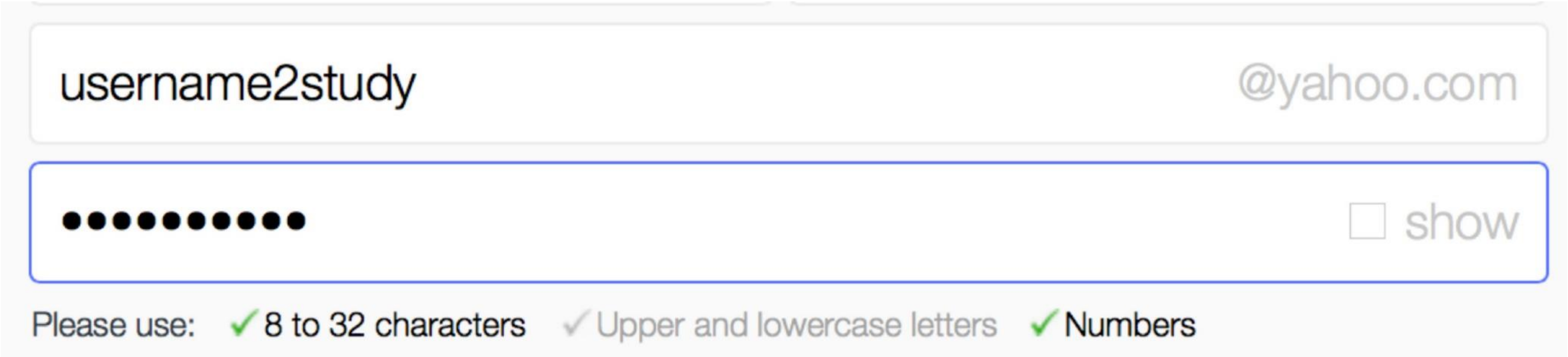


Images Creative Commons on Flickr by *carowallis* and *woodythrower*

Can we make the creation of secure passwords *more usable*?

Requirements Feedback

Requirements Feedback

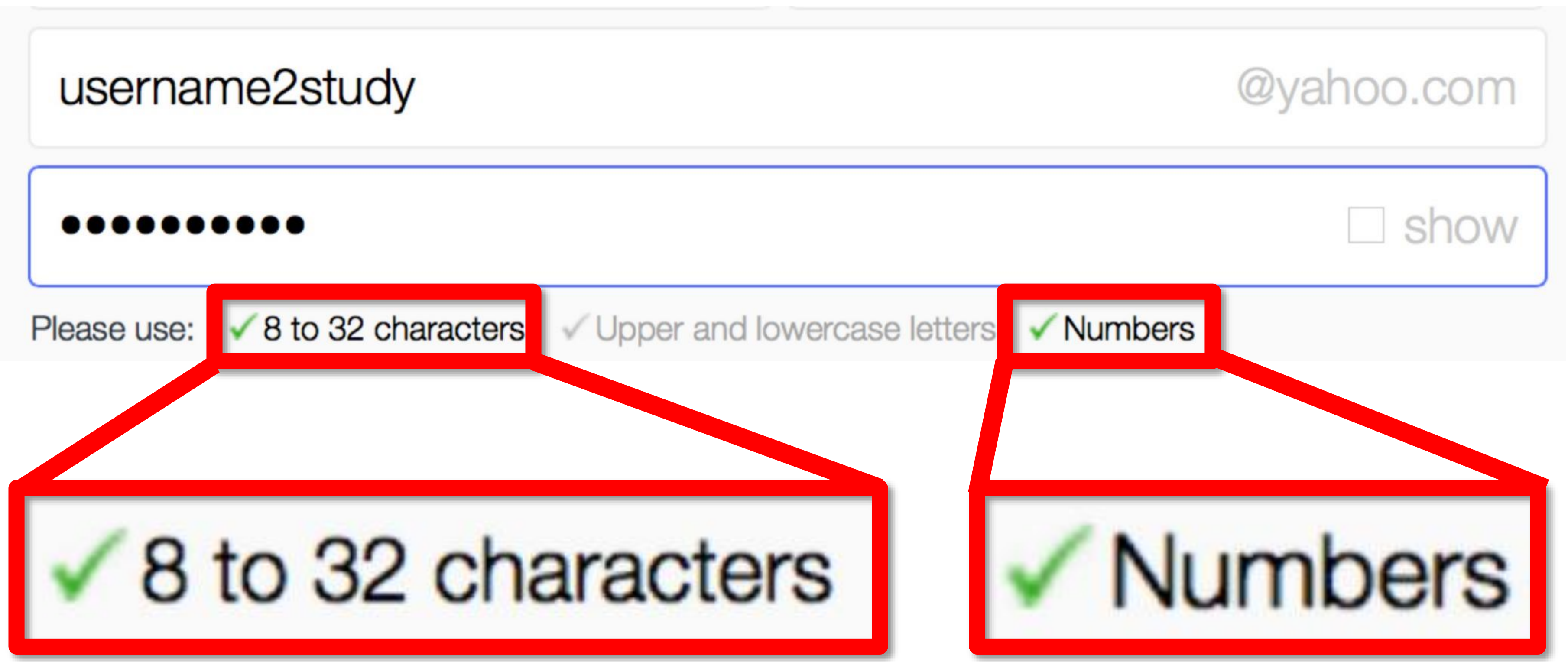


username2study @yahoo.com

●●●●●●●●●●● □ show

Please use: ✓ 8 to 32 characters ✓ Upper and lowercase letters ✓ Numbers

Requirements Feedback

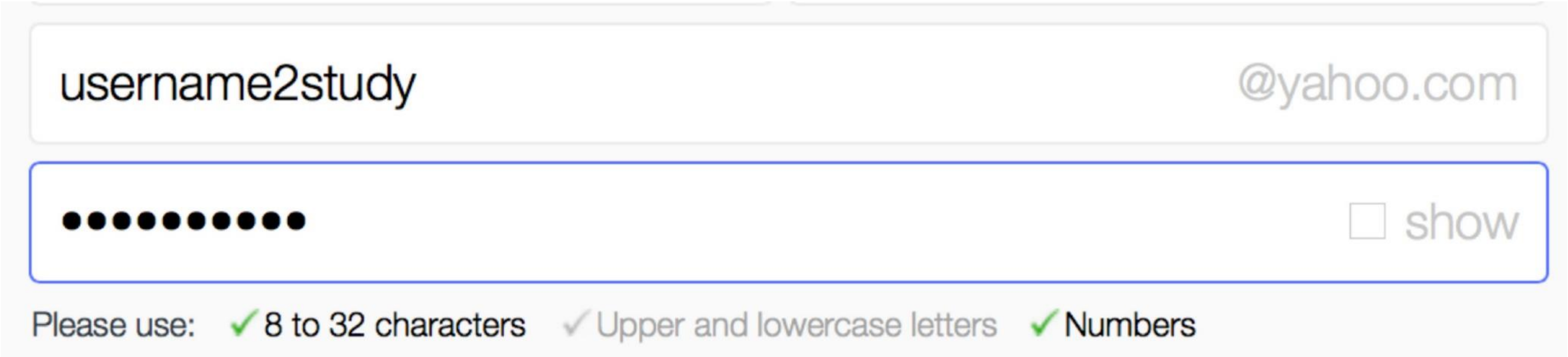


username2study @yahoo.com

●●●●●●●●●●● □ show

Please use: ✓ 8 to 32 characters ✓ Upper and lowercase letters ✓ Numbers

Requirements Feedback

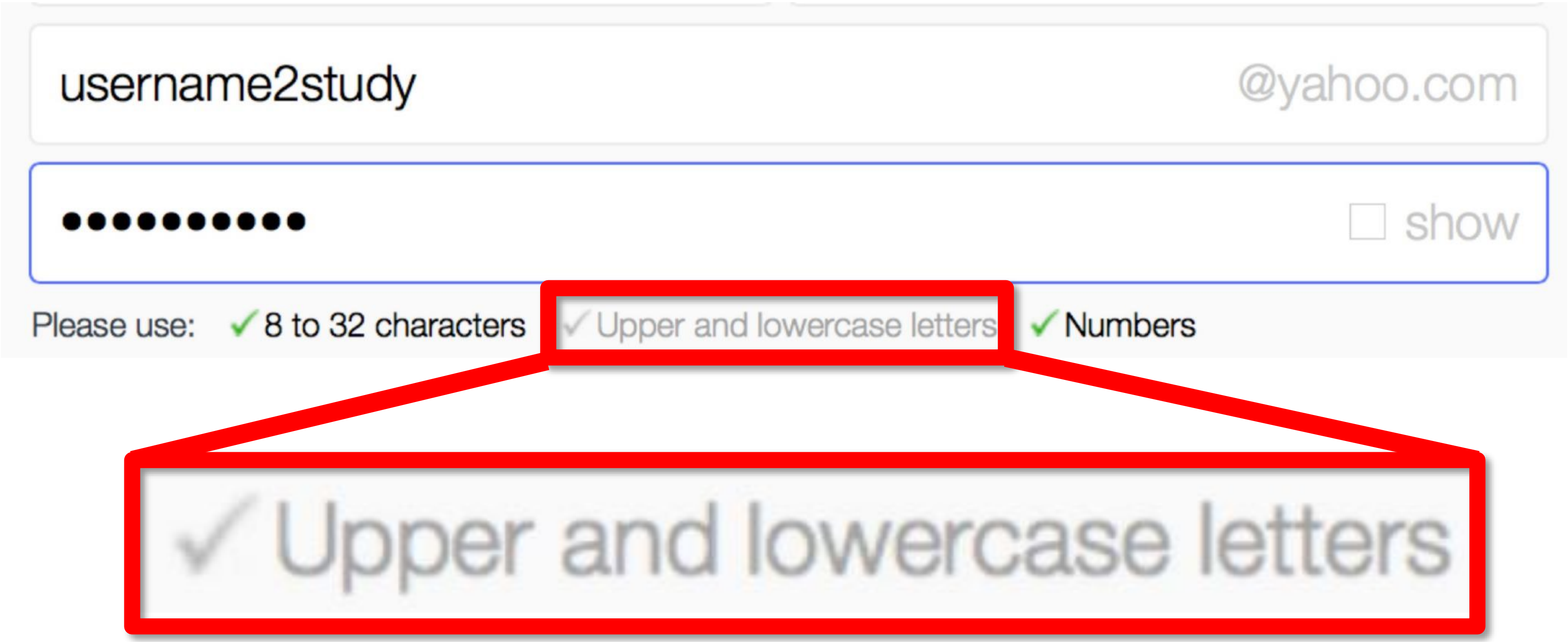


username2study @yahoo.com

●●●●●●●●●●● □ show

Please use: ✓ 8 to 32 characters ✓ Upper and lowercase letters ✓ Numbers

Requirements Feedback



username2study @yahoo.com

•••••••••• □ show

Please use: ✓ 8 to 32 characters ✓ Upper and lowercase letters ✓ Numbers

✓ Upper and lowercase letters

Multi-Step Password Guidance

Multi-Step Password Guidance

pass12word

Multi-Step Password Guidance

pass12word

Multi-Step Password Guidance

`pa$ss12wo!rd`

Primary Research Questions

Primary Research Questions

1. How do *blacklist* and *pattern* requirements impact password security and usability?

Primary Research Questions

1. How do *blacklist* and *pattern* requirements impact password security and usability?

2. Does *real-time requirements feedback* improve the usability of creating strong passwords?

Primary Research Questions

1. How do *blacklist* and *pattern* requirements impact password security and usability?

2. Does *real-time requirements feedback* improve the usability of creating strong passwords?

3. Does a *multi-step guidance process* improve the usability of creating strong passwords?

Methodology

- 6,435-participant online study
  - Recruited U.S. users of Amazon’s Mechanical Turk

Methodology

- 6,435-participant online study
  - Recruited U.S. users of Amazon’s Mechanical Turk
- Between-subjects design with 9 conditions

Methodology

- 6,435-participant online study
  - Recruited U.S. users of Amazon’s Mechanical Turk
- Between-subjects design with 9 conditions
- 2-part study, 2+ days apart
  - Compensated $0.55 and $0.70, respectively
  - Mazurek et al. CCS ‘13 and Fahl et al. SOUPS ‘13

Methodology

- Part 1: Create password & take survey
  - Scenario: Email provider requires password change

Methodology

- Part 1: Create password & take survey
    - Scenario: Email provider requires password change
- Part 2: Return, re-enter password, & take survey

Security Metric: Guessability

- Guessability – how many guesses to crack?
  - Threat model: offline attack

Image Creative Commons on Flickr by *ilahi*

Security Metric: Guessability

- Guessability – how many guesses to crack?
  - Threat model: offline attack
    - Naïve first guesses: *aaaaaaaaaaaa, aaaaaaaaaaab*

Image Creative Commons on Flickr by *ilahi*

Security Metric: Guessability

- Guessability – how many guesses to crack?
  - Threat model: offline attack
    - Naïve first guesses: *aaaaaaaaaaaa, aaaaaaaaaaab*
    - Better first guesses: *123456781234, password1234*

Image Creative Commons on Flickr by *ilahi*

Security Metric: Guessability

- Guessability – how many guesses to crack?
  - Threat model: offline attack
    - Naïve first guesses: *aaaaaaaaaaaa, aaaaaaaaaaab*
    - Better first guesses: *123456781234, password1234*

- 20 trillion guesses per condition

Image Creative Commons on Flickr by *ilahi*

Usability Metrics

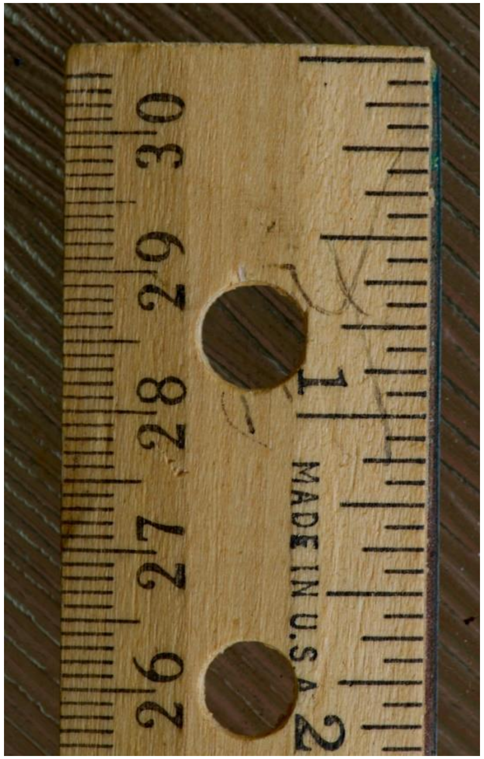


Image Creative Commons on Flickr by *ilahi*

Usability Metrics

- Password creation
  - Time
  - # failed attempts

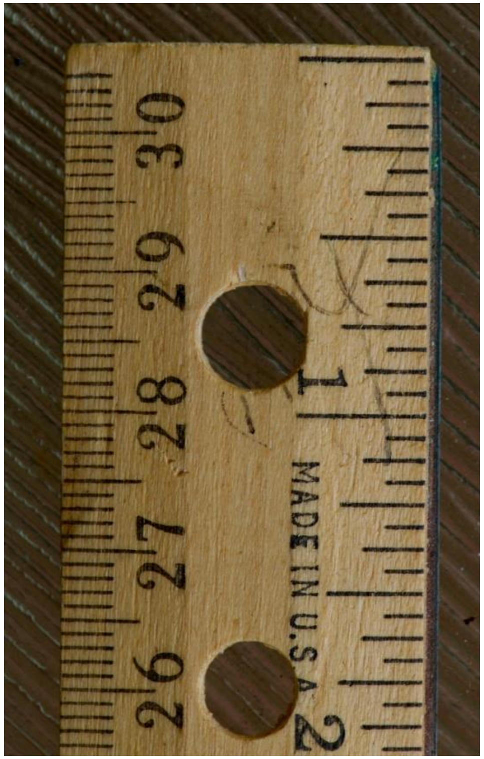


Image Creative Commons on Flickr by *ilahi*

Usability Metrics

- Password creation
  - Time
  - # failed attempts
- Participant sentiment
  - Self-reported
  - Study drop-out

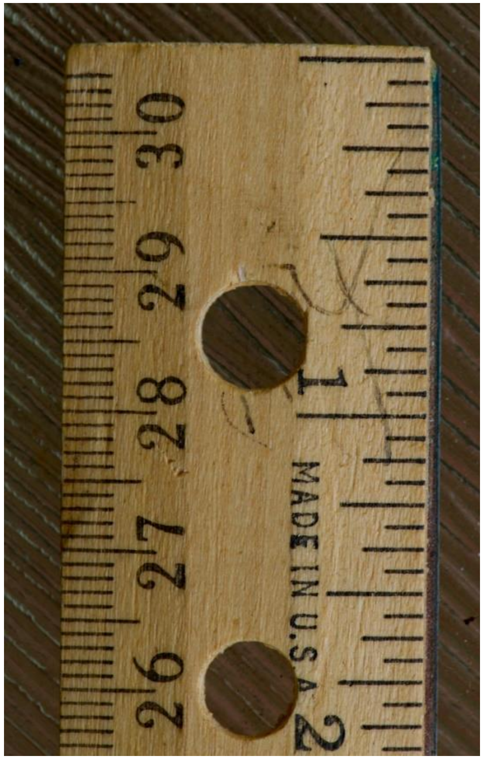


Image Creative Commons on Flickr by *ilahi*

Usability Metrics

- Password creation
  - Time
  - # failed attempts
- Participant sentiment
  - Self-reported
  - Study drop-out
- Memorability
  - ~5 minutes after creation
  - 2-5 days after creation

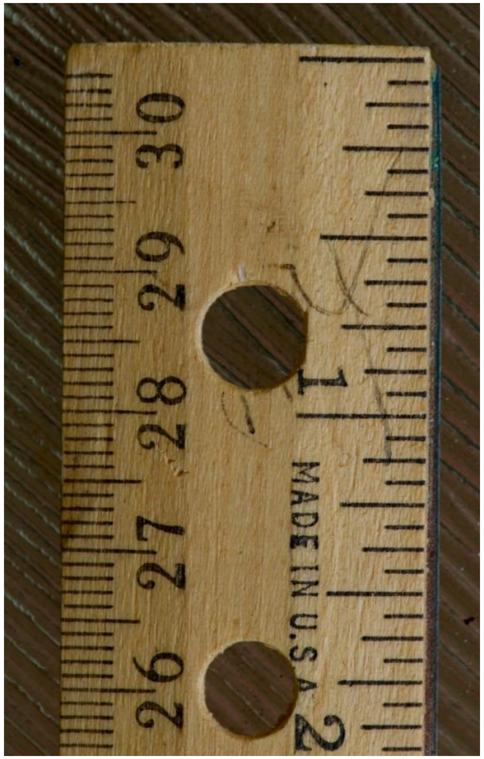
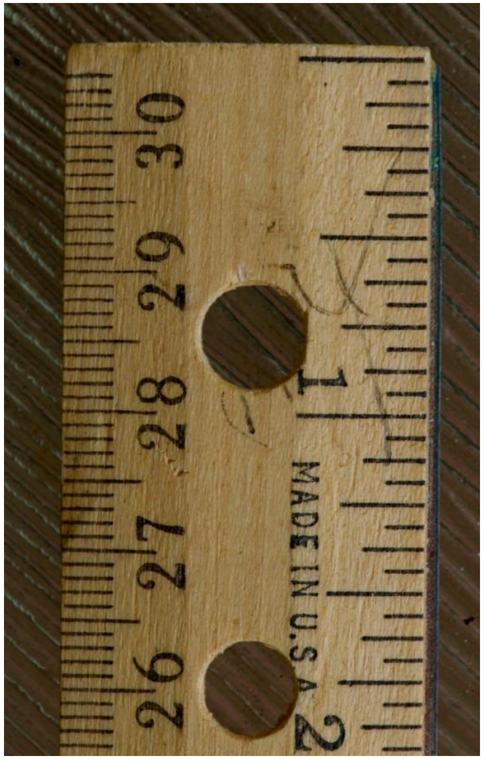


Image Creative Commons on Flickr by *ilahi*

Usability Metrics



- Password creation
  - Time
  - # failed attempts
- Participant sentiment
  - Self-reported
  - Study drop-out
- Memorability
  - ~5 minutes after creation
  - 2-5 days after creation
- Writing down/storing password

Image Creative Commons on Flickr by *ilahi*

Participants

- 6,435 participants
- 47% male, 53% female
- Median age 28

Primary Research Questions

1. How do *blacklist* and *pattern* requirements impact password security and usability?

2. Does *real-time requirements feedback* improve the usability of creating strong passwords?

3. Does a *multi-step guidance process* improve the usability of creating strong passwords?

RQ1 Conditions

- **Base**: 12+ characters from 3+ classes

RQ1 Conditions

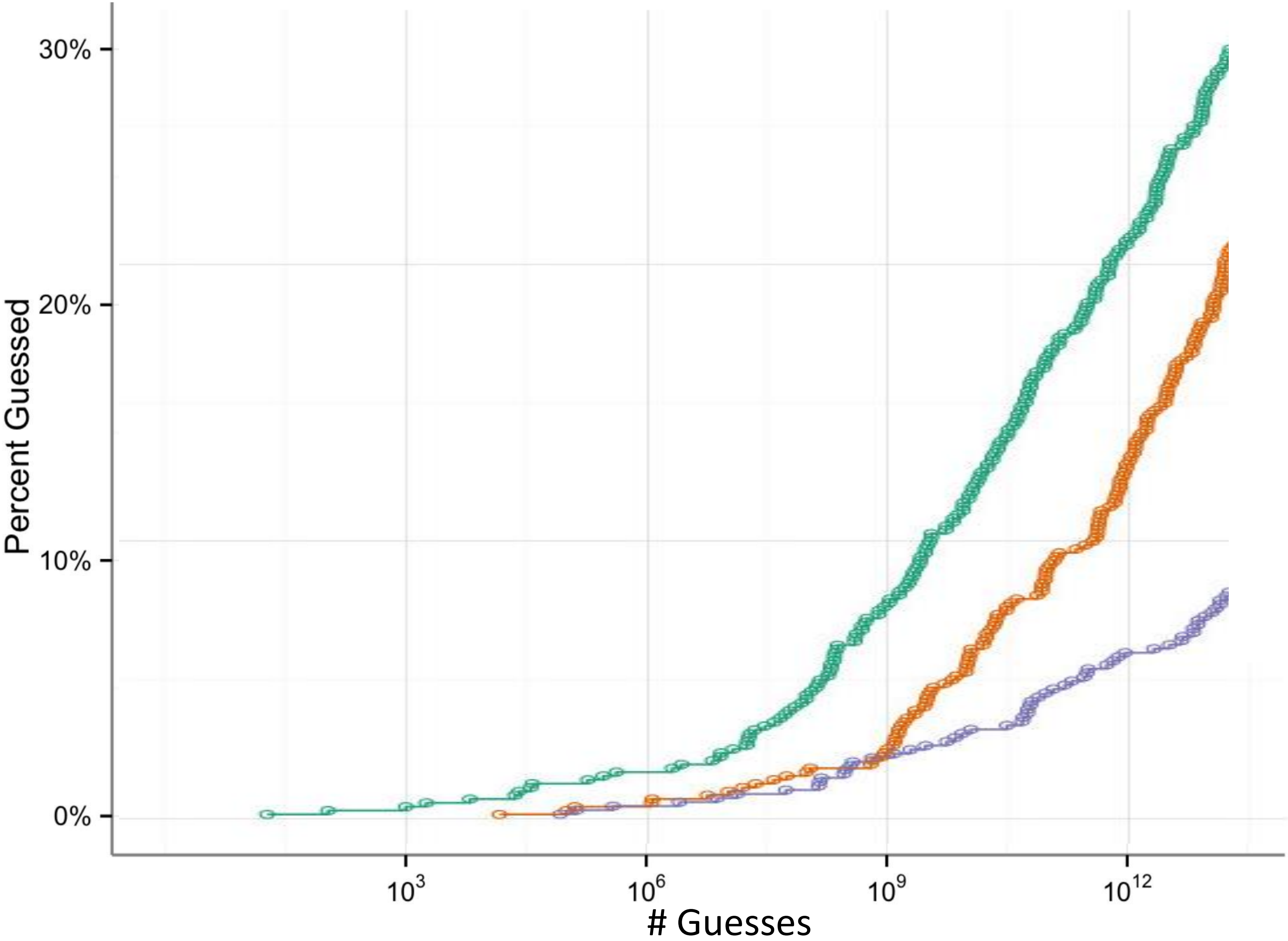
- **Base**: 12+ characters from 3+ classes

- **Blacklist**: Base + disallowed 41,329 substrings (e.g., “1234”, years, “abcd”)

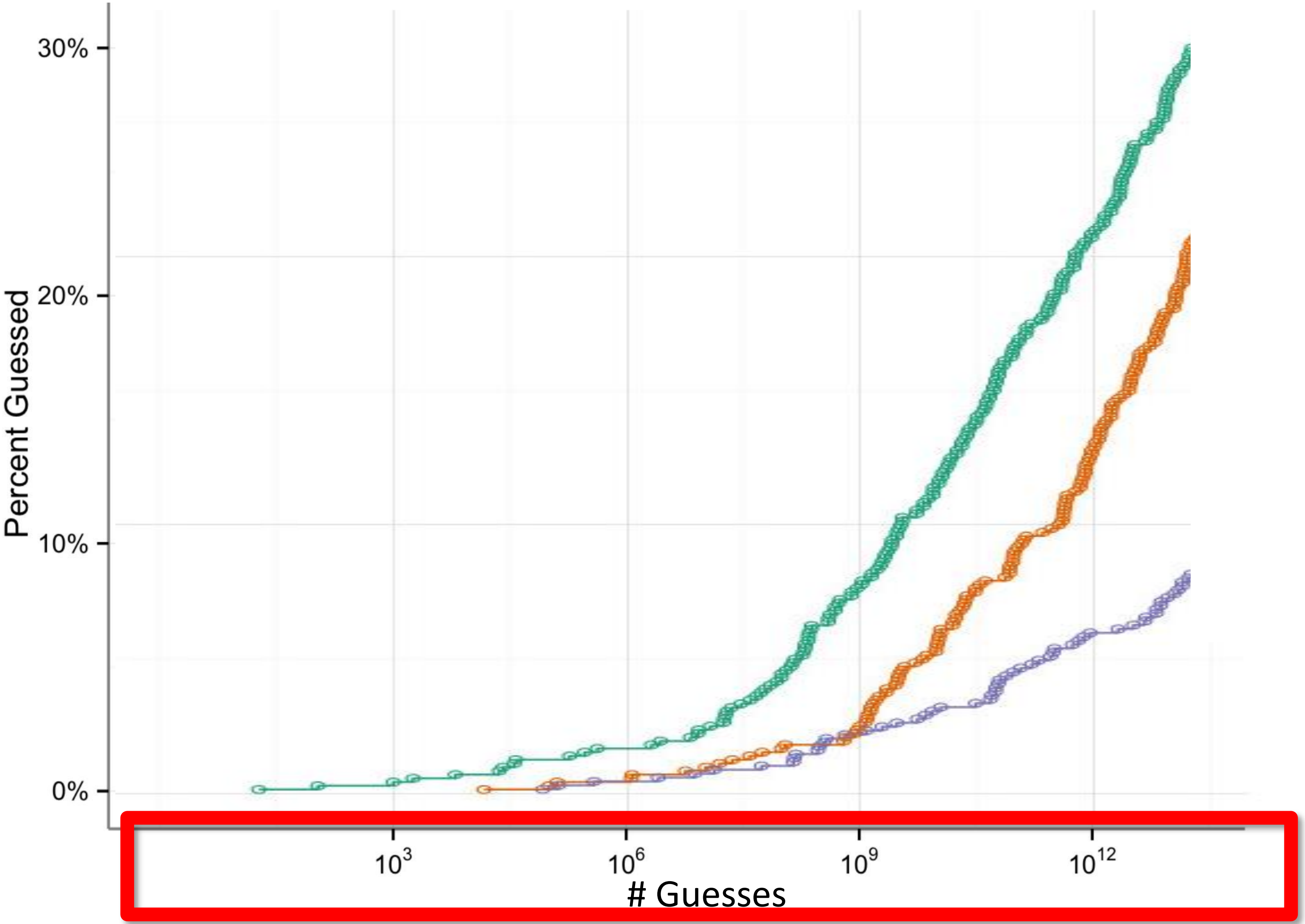
RQ1 Conditions

- **Base**: 12+ characters from 3+ classes
- **Blacklist**: Base + disallowed 41,329 substrings (e.g., “1234”, years, “abcd”)
- **Pattern**: Base + start and end w/ lowercase letter

RQ1 Results – Security

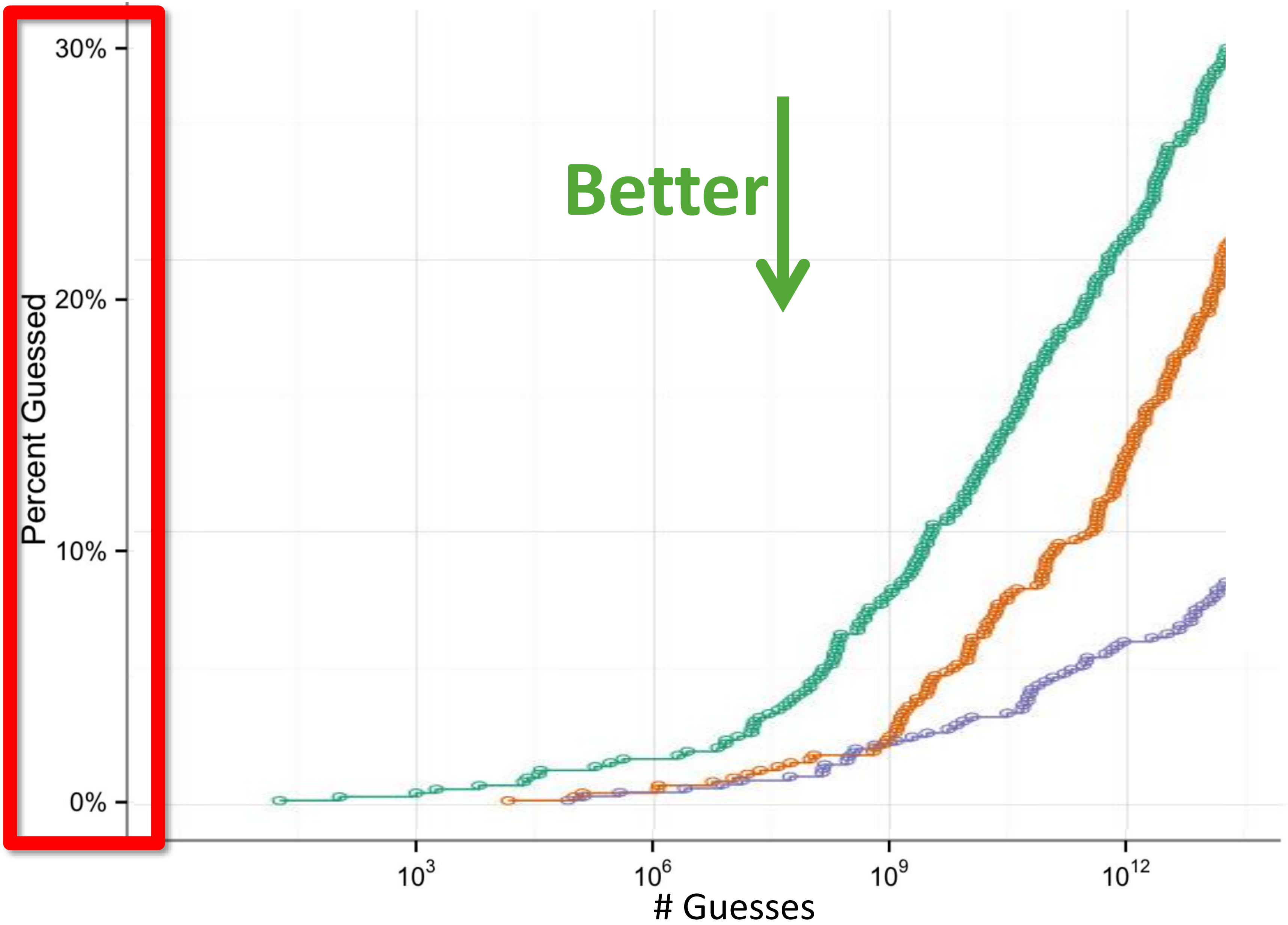


RQ1 Results – Security



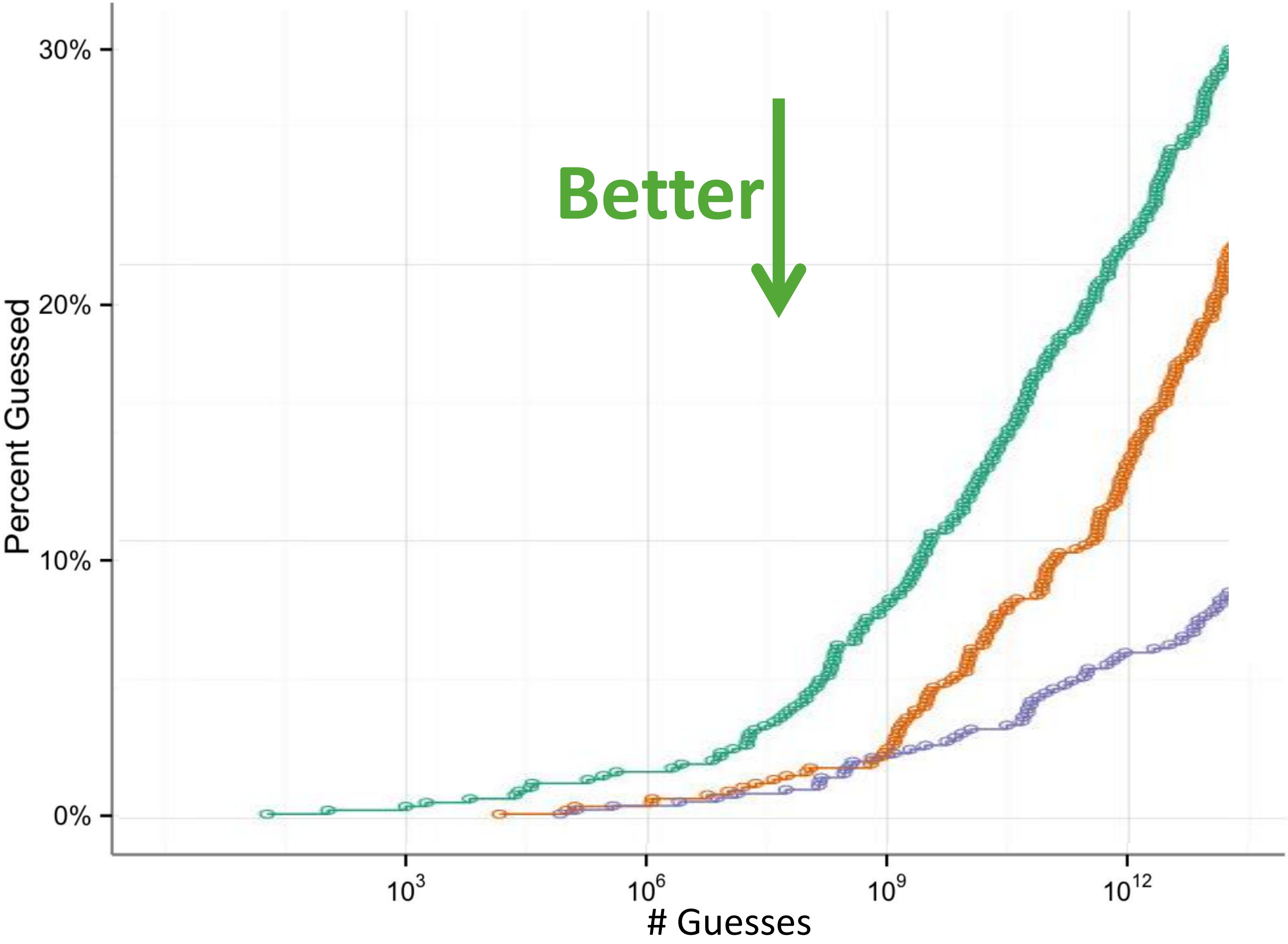
# Guesses

RQ1 Results – Security

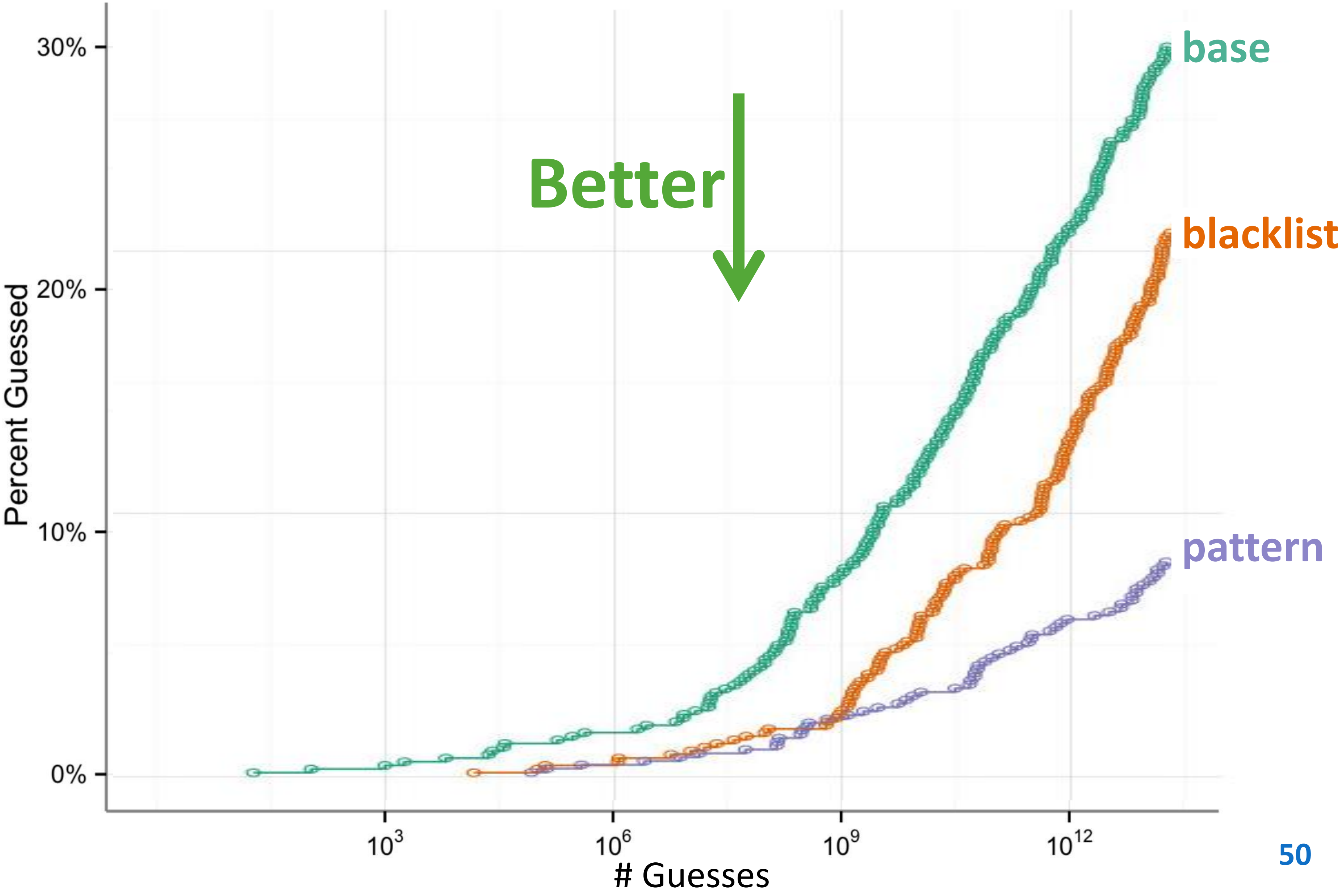


Better

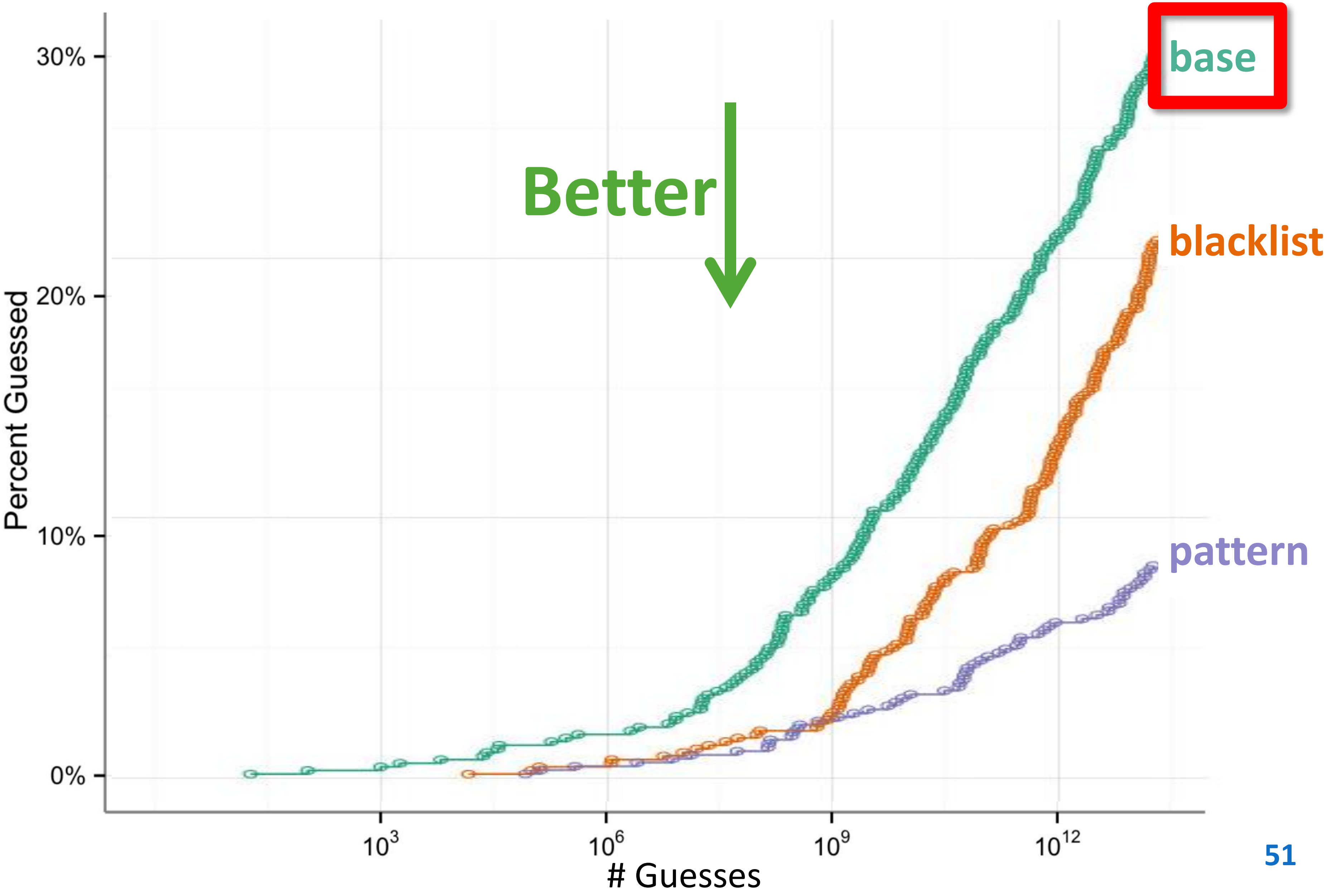
RQ1 Results – Security



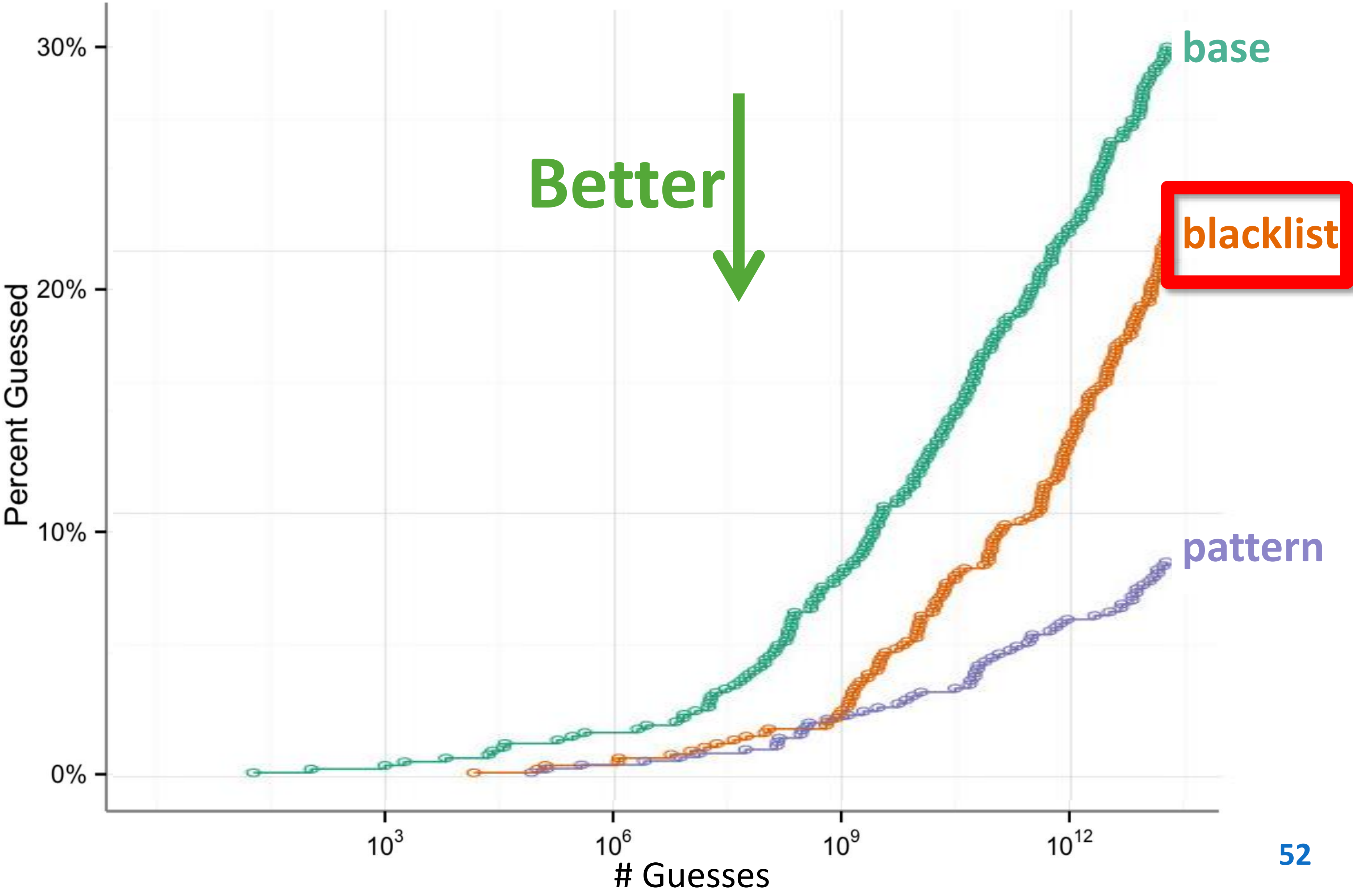
RQ1 Results – Security



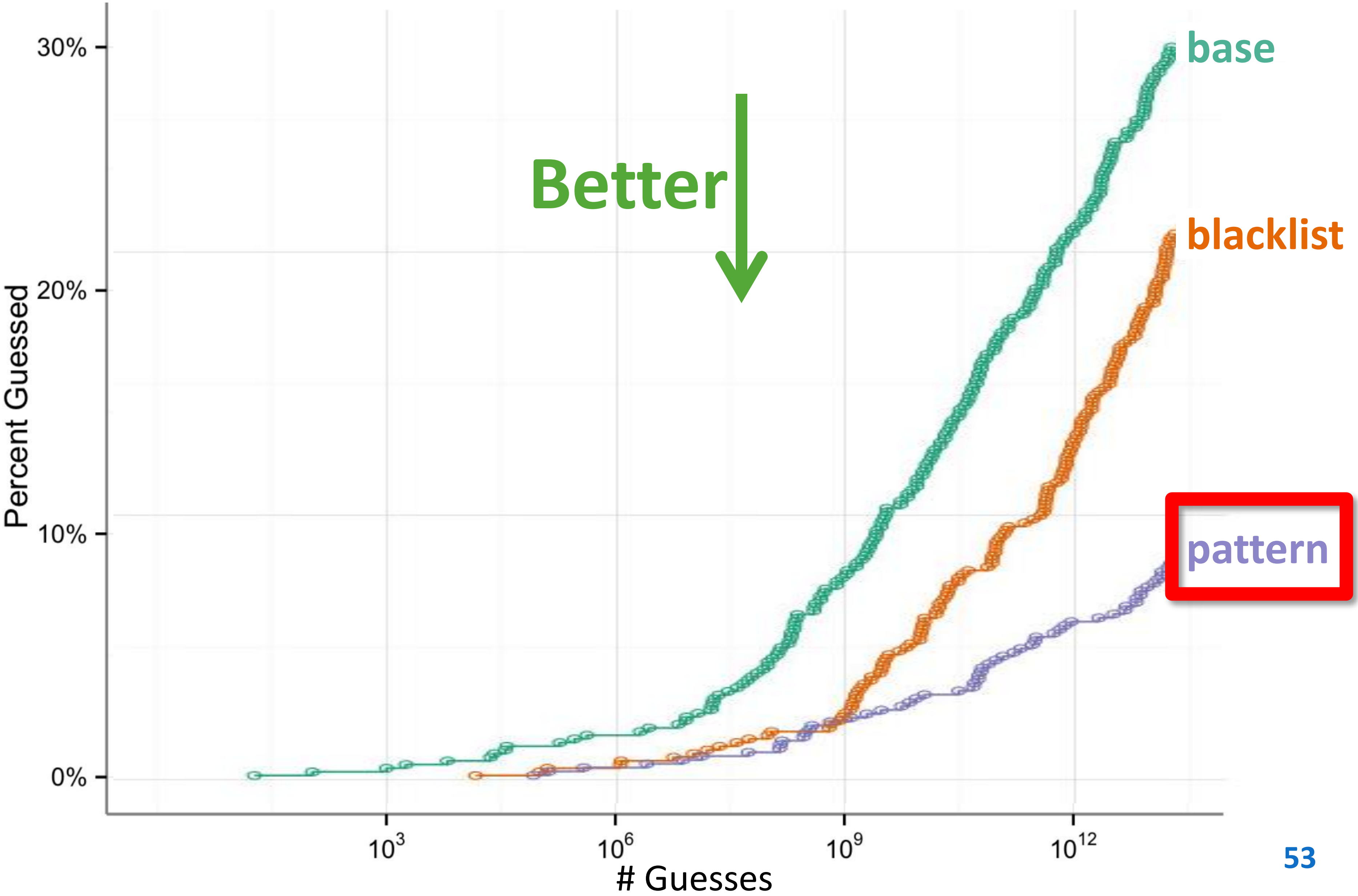
RQ1 Results – Security



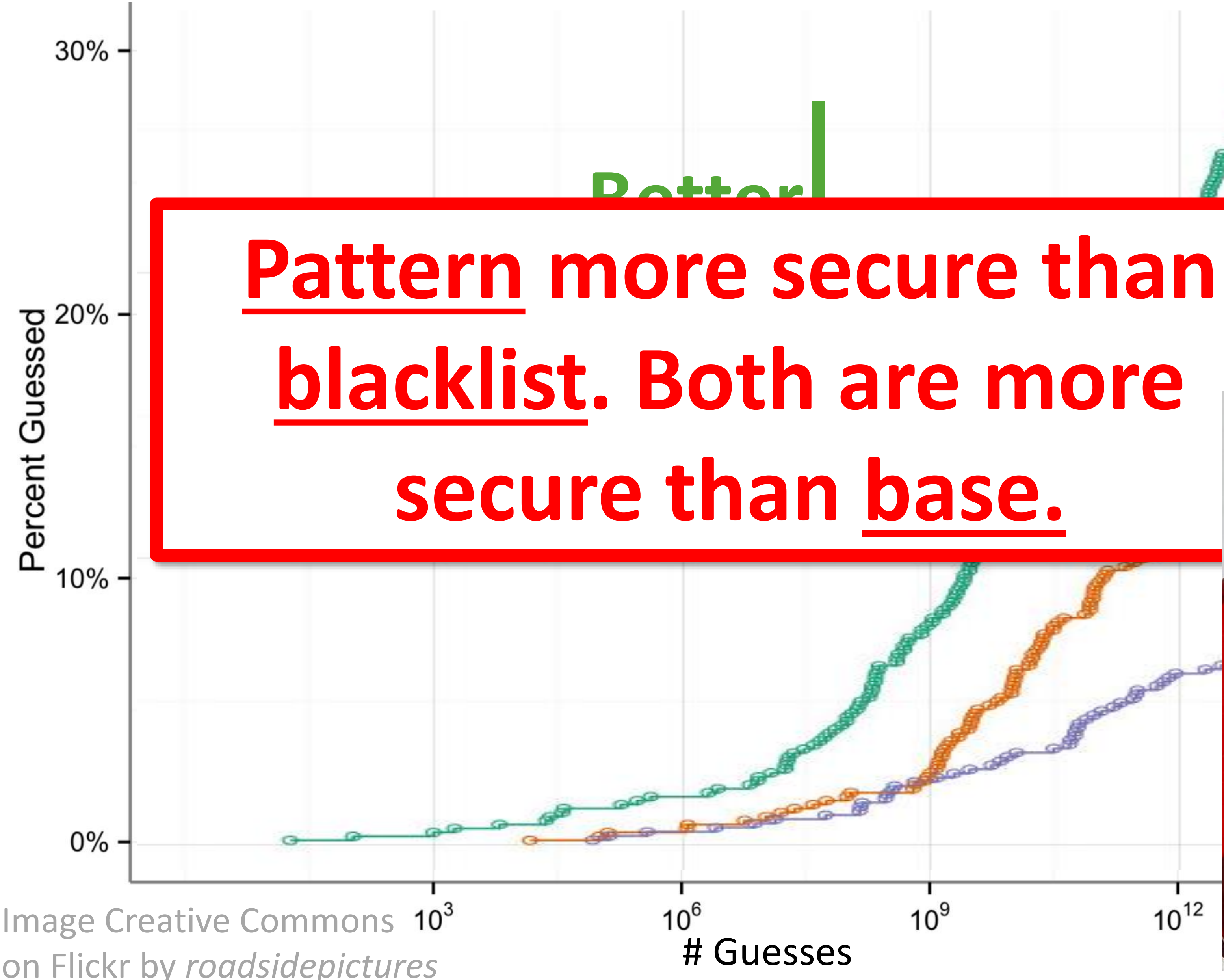
RQ1 Results – Security



RQ1 Results – Security



RQ1 Results – Security



Pattern more secure than blacklist. Both are more secure than base.

Image Creative Commons on Flickr by *roadsidepictures*

RQ1 Results – Usability

RQ1 Results – Usability

- *Pattern* took longer to create than *blacklist*; *blacklist* longer than *base*

RQ1 Results – Usability

- *Pattern* took longer to create than *blacklist*; *blacklist* longer than *base*

- *Pattern* more difficult to create than *base/blacklist*

RQ1 Results – Usability

- *Pattern* took longer to create than *blacklist*; *blacklist* longer than *base*

- *Pattern* more difficult to create than *base/blacklist*

- *Pattern* stored or written down more than *base/blacklist*

Primary Research Questions

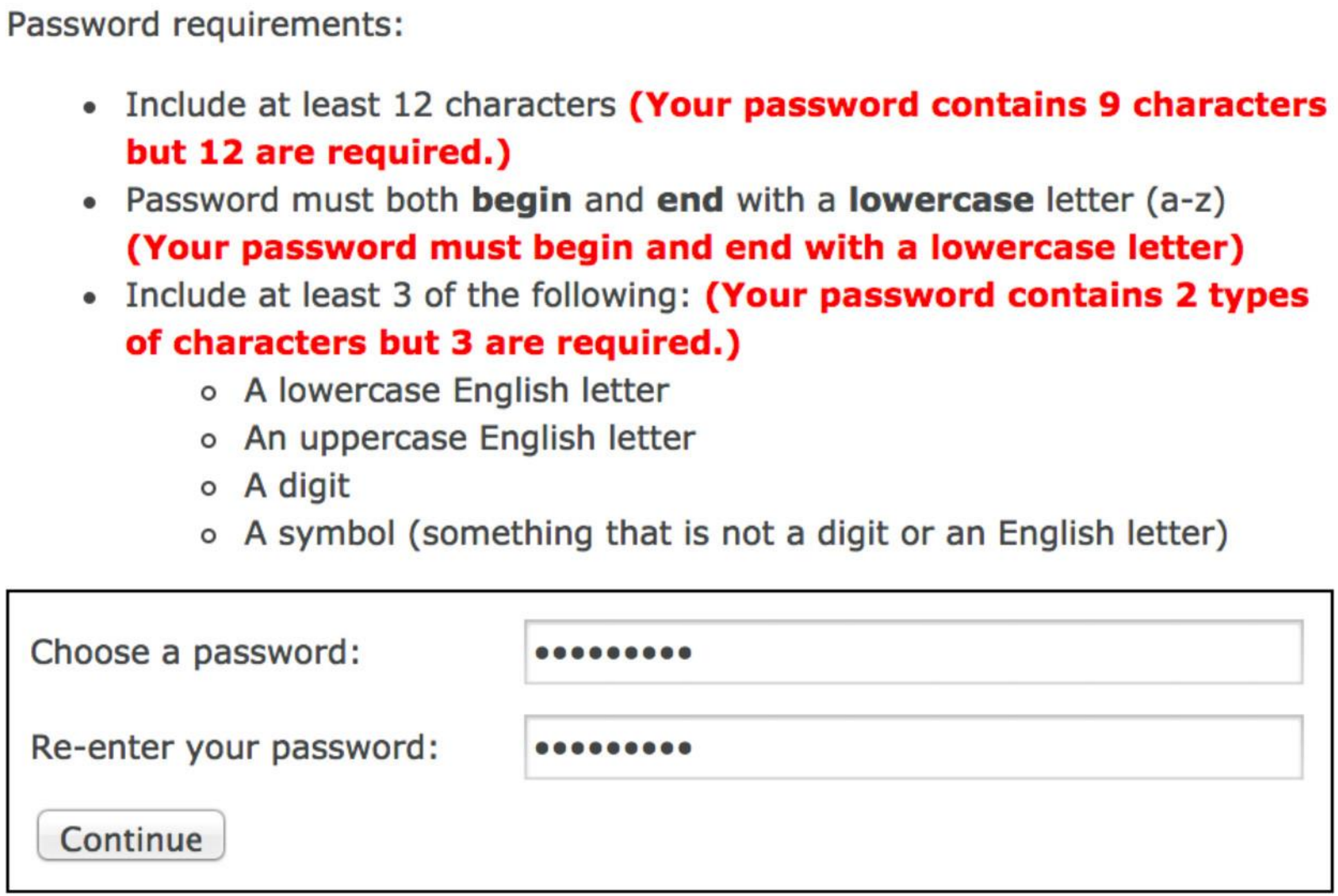
1. How do *blacklist* and *pattern* requirements impact password security and usability?

2. Does *real-time requirements feedback* improve the usability of creating strong passwords?

3. Does a *multi-step guidance process* improve the usability of creating strong passwords?

RQ2 Conditions

- Realtime (“**rt**”) feedback



Password requirements:

- Include at least 12 characters **(Your password contains 9 characters but 12 are required.)**
- Password must both **begin** and **end** with a **lowercase** letter (a-z) **(Your password must begin and end with a lowercase letter)**
- Include at least 3 of the following: **(Your password contains 2 types of characters but 3 are required.)**
    - A lowercase English letter
    - An uppercase English letter
    - A digit
    - A symbol (something that is not a digit or an English letter)

Choose a password: ●●●●●●●●●

Re-enter your password: ●●●●●●●●●

RQ2 Results – Security

- Requirements feedback did not significantly impact security

RQ2 Results – Security

- Requirements feedback did not significantly impact security
- Feedback → Less likely to exceed requirements

RQ2 Results – Usability

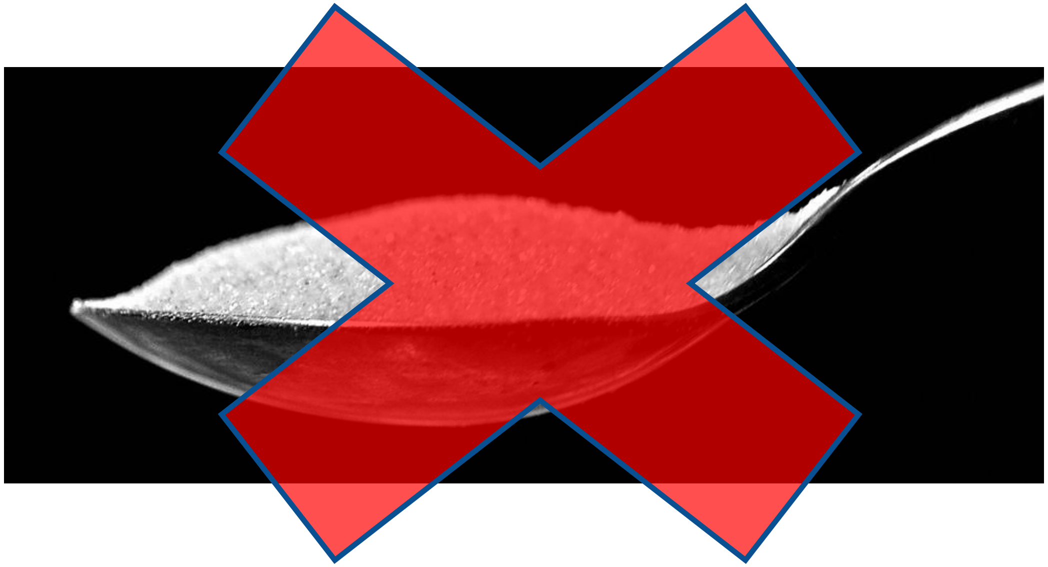
RQ2 Results – Usability

- More likely to submit compliant password with requirements feedback

- No significant impact on other usability metrics

RQ2 Results – Usability

- More likely to submit compliant password with requirements feedback
- No significant impact on other usability metrics



Primary Research Questions

1. How do *blacklist* and *pattern* requirements impact password security and usability?

2. Does *real-time requirements feedback* improve the usability of creating strong passwords?

3. Does a *multi-step guidance process* improve the usability of creating strong passwords?

RQ3 Conditions

- **Pattern-rt**: 12+ characters, 3+ character classes, start & end with lowercase letter, feedback

RQ3 Conditions

- **Pattern-rt**: 12+ characters, 3+ character classes, start & end with lowercase letter, feedback
- **Guide:** Multi-step creation process
  - Step 1: 10+ character pattern password
  - Step 2: The user adds 2 characters

RQ3 Conditions

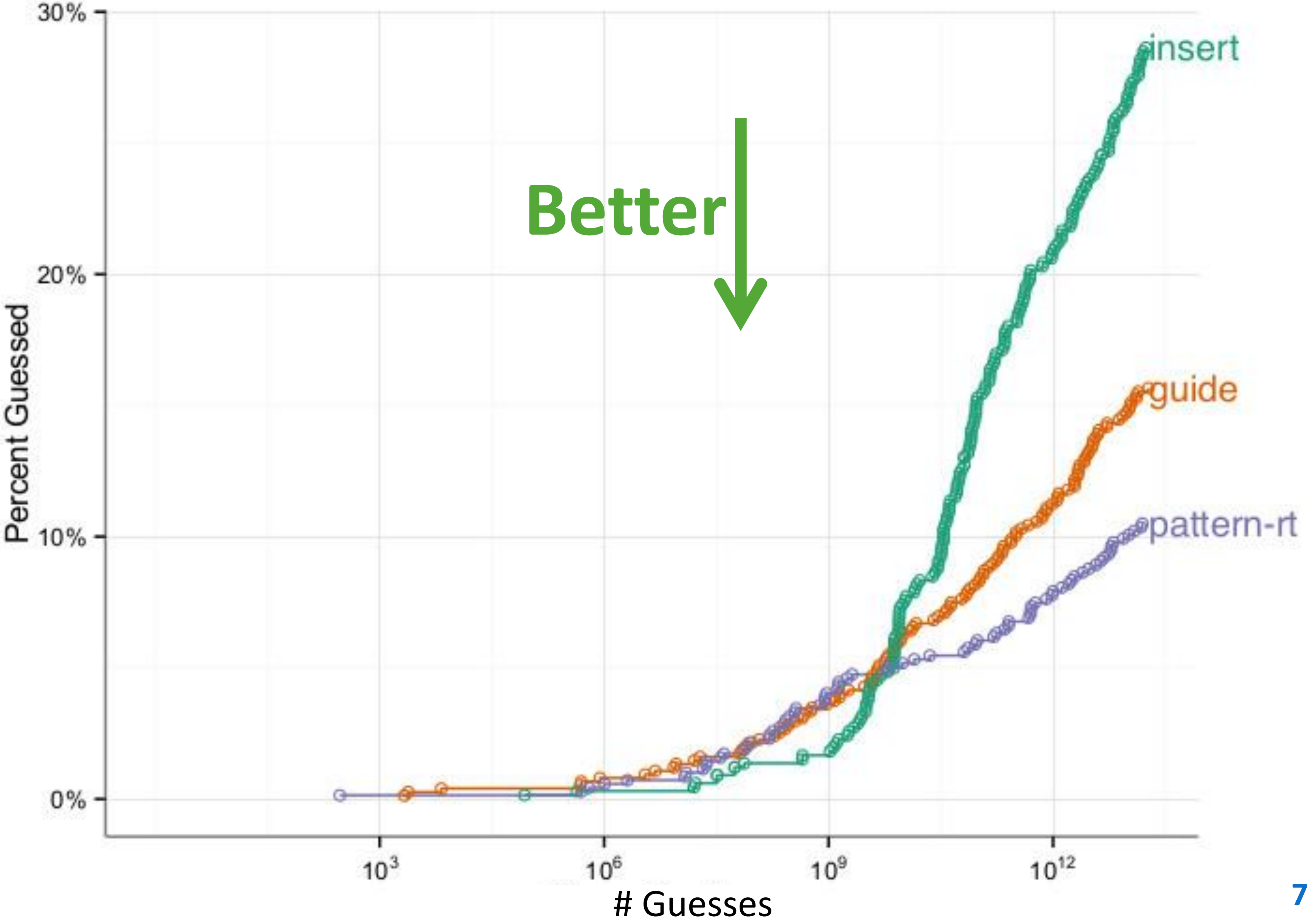
- **Pattern-rt**: 12+ characters, 3+ character classes, start & end with lowercase letter, feedback

- **Guide:** Multi-step creation process
  - Step 1: 10+ character pattern password
  - Step 2: The user adds 2 characters

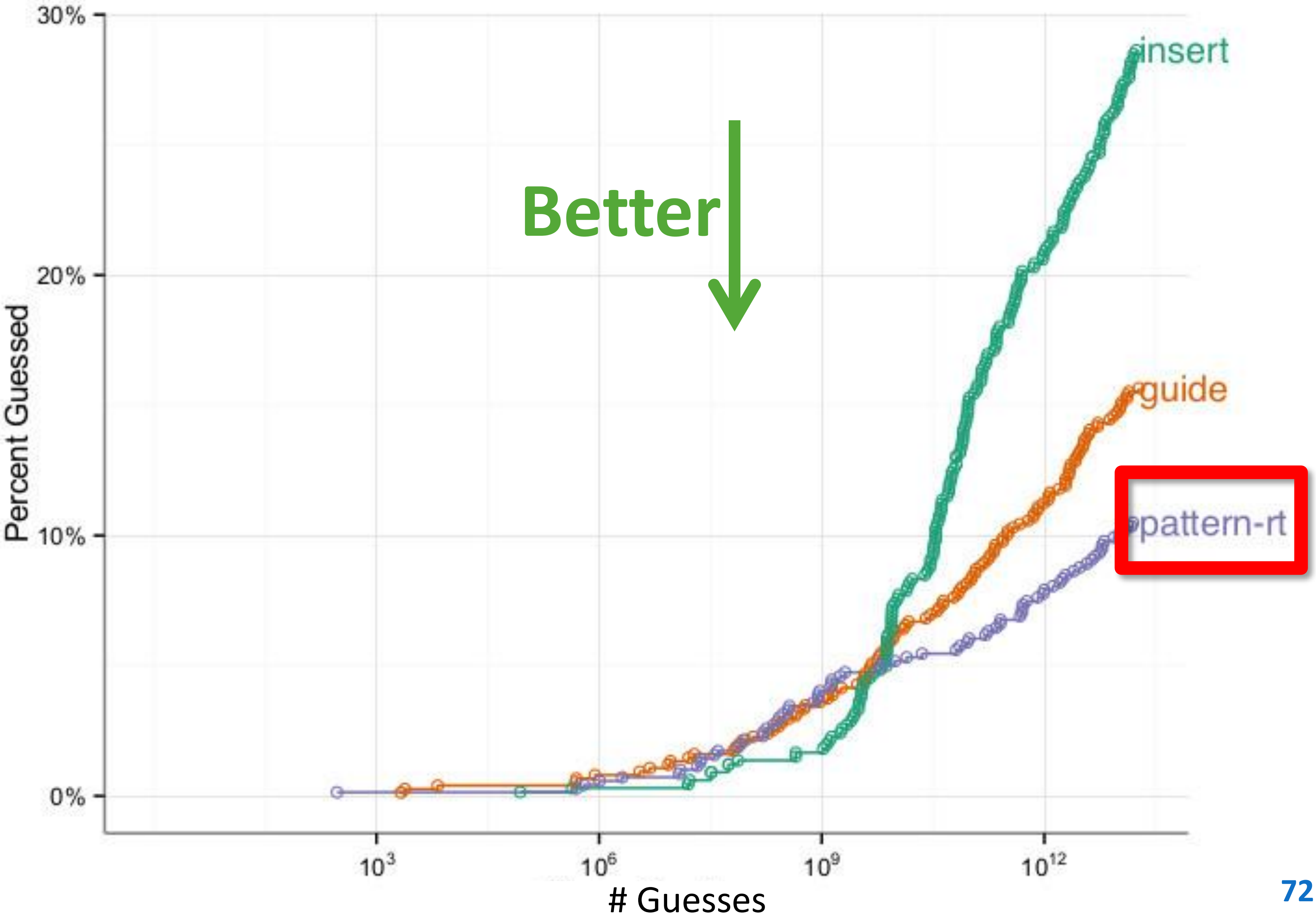
- **Insert:** Multi-step creation process
  - Step 1: 10+ character pattern password
  - Step 2: The system adds 2 random characters

RQ3 Results – Security

RQ3 Results – Security

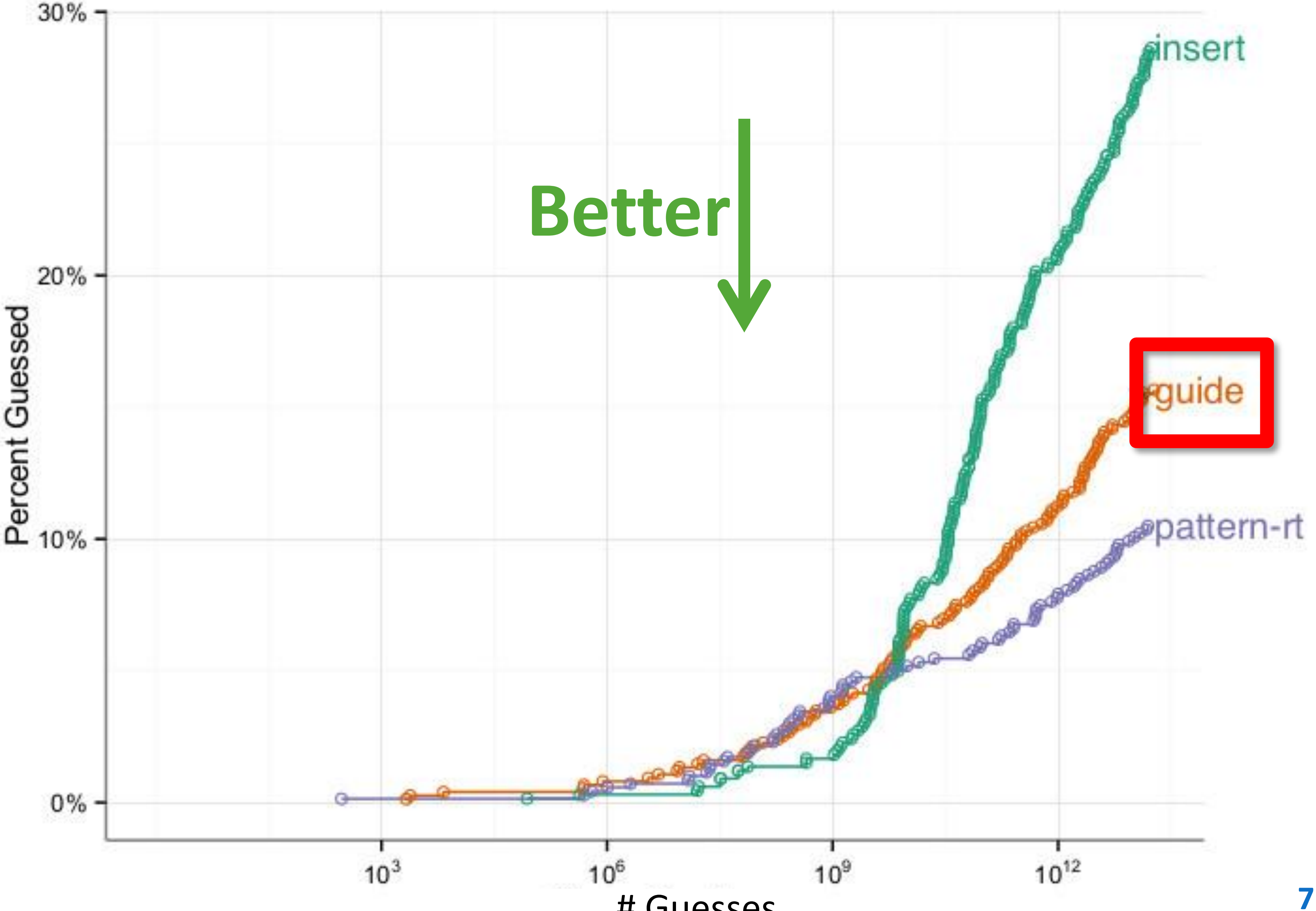


RQ3 Results – Security

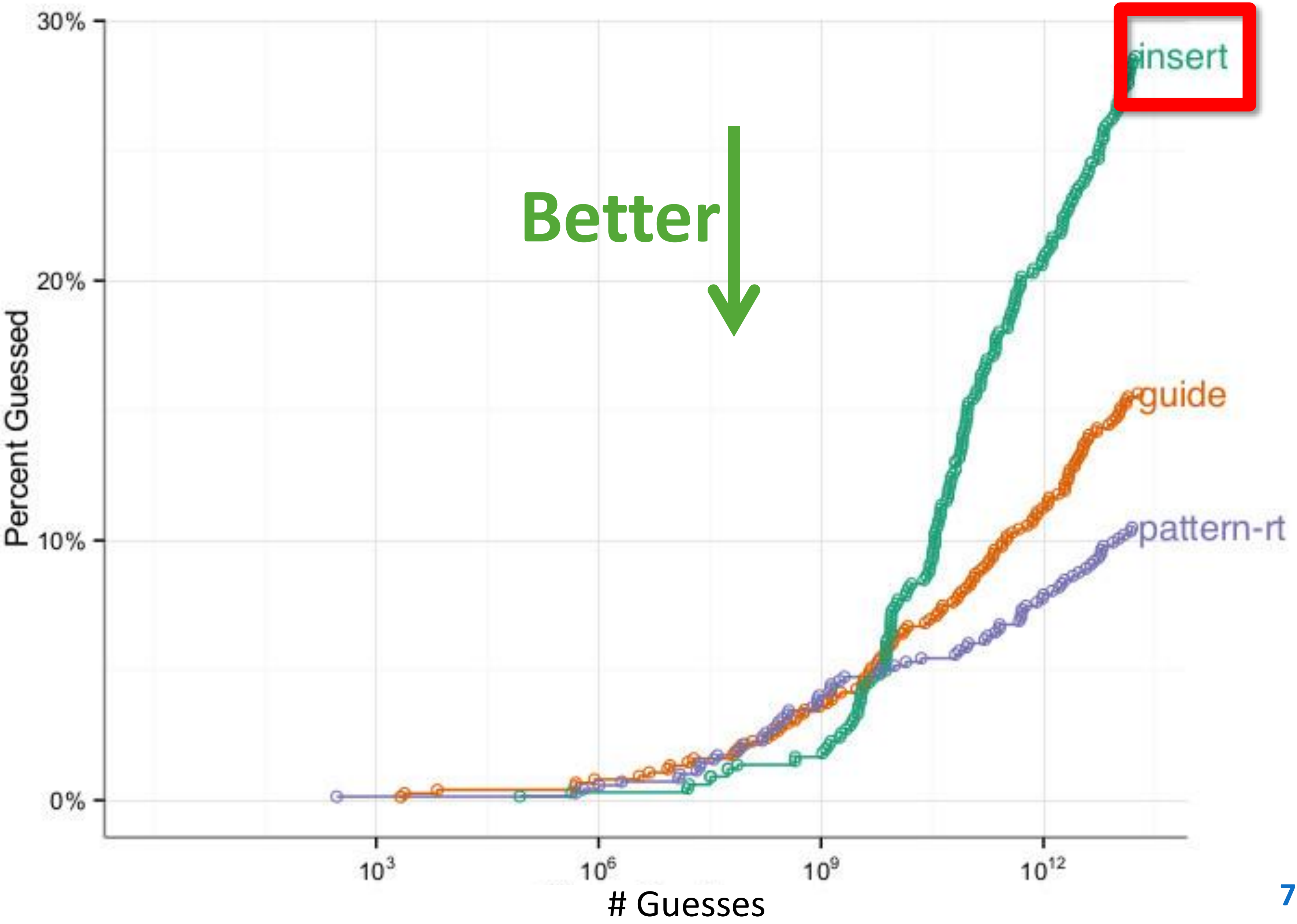


Better

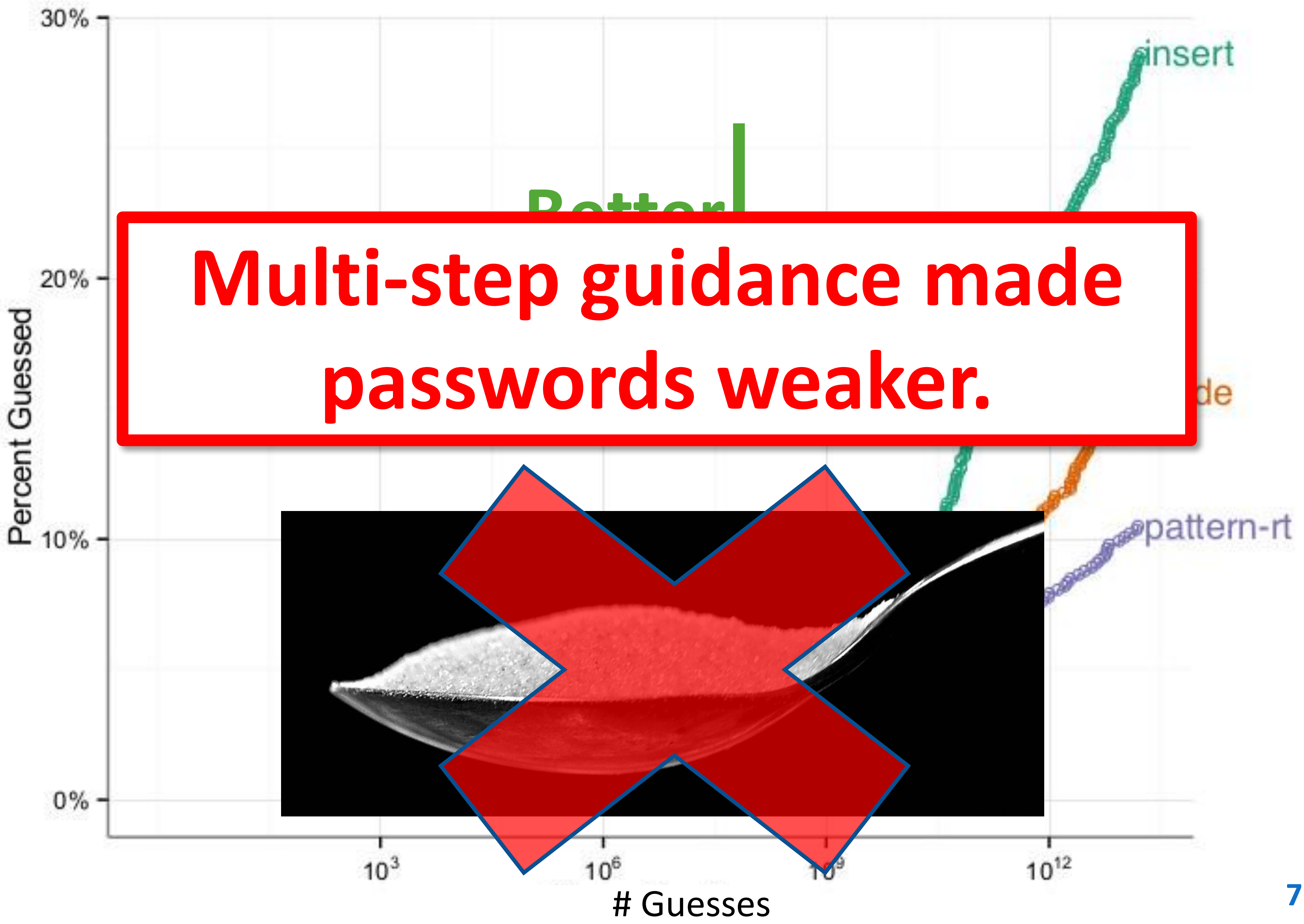
RQ3 Results – Security



RQ3 Results – Security



RQ3 Results – Security



Multi-step guidance made passwords weaker.

RQ3 Results – Usability

RQ3 Results – Usability

- *Guide* and *insert* passwords less difficult to create than *pattern-rt*

RQ3 Results – Usability

- *Guide* and *insert* passwords less difficult to create than *pattern-rt*

- *Guide* and *insert* participants less likely to drop out than *pattern-rt*

RQ3 Results – Usability

- *Guide* and *insert* passwords less difficult to create than *pattern-rt*
- *Guide* and *insert* participants less likely to drop out than *pattern-rt*
- *Insert* more likely to be written down/stored than *pattern-rt*

Limitations

- Tested recall at only two points
- Passwords created for a research study
  - Mazurek et al. CCS ‘13 and Fahl et al. SOUPS ‘13
- Did not test multiple devices

Primary Research Questions

1. How do *blacklist* and *pattern* requirements impact password security and usability?

2. Does *real-time requirements feedback* improve the usability of creating strong passwords?

3. Does a *multi-step guidance process* improve the usability of creating strong passwords?

Conclusions

1. How do *blacklist* and *pattern* requirements impact password security and usability?

2. Does *real-time requirements feedback* improve the usability of creating strong passwords?

3. Does a *multi-step guidance process* improve the usability of creating strong passwords?

Conclusions

1. ***Blacklist* and *pattern* requirements make passwords stronger, but reduce usability**

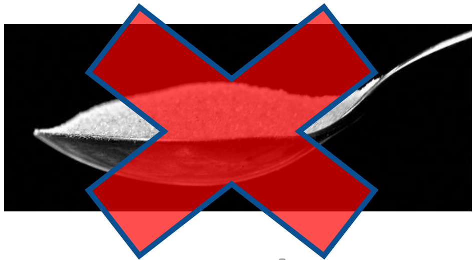
2. Does *real-time requirements feedback* improve the usability of creating strong passwords?

3. Does a *multi-step guidance process* improve the usability of creating strong passwords?

Conclusions

1. ***Blacklist*** **and *pattern* requirements make passwords stronger, but reduce usability**

2. ***Real-time requirements feedback*** **did not have a major security or usability impact**

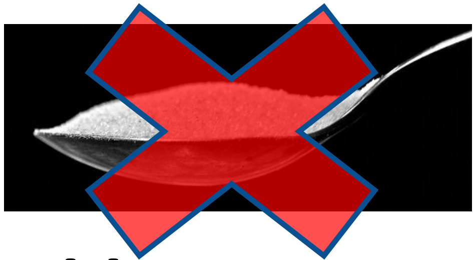


3. Does a *multi-step guidance process* improve the usability of creating strong passwords?

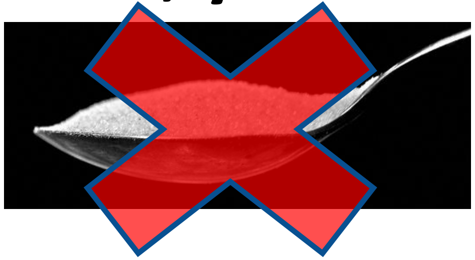
Conclusions

1. *Blacklist* and *pattern* requirements make passwords stronger, but reduce usability

2. *Real-time requirements feedback* did not have a major security or usability impact



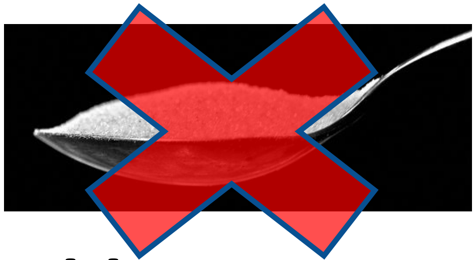
3. *Multi-step guidance process* more usable, yet leads to weaker passwords



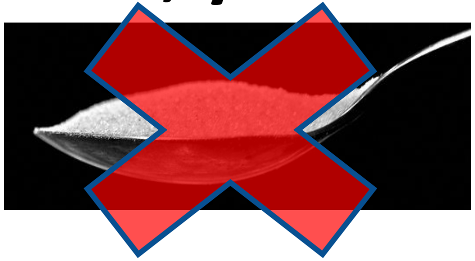
Conclusions

1. *Blacklist* and *pattern* requirements make passwords stronger, but reduce usability

2. *Real-time requirements feedback* did not have a major security or usability impact



3. *Multi-step guidance process* more usable, yet leads to weaker passwords



Richard Shay, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, Alain Forget, Saranga Komanduri, Michelle L. Mazurek, William Melicher, Sean M. Segreti, Blase Ur

blase@blaseur.com

Carnegie Mellon

UNIVERSITY OF MARYLAND