Telepathwords: preventing weak passwords by reading users' minds

Saranga Komanduri, Richard Shay, Lorrie Faith Cranor, Cormac Herley, Stuart Schechter

Carnegie Mellon University
CyLab

institute for SOFTWARE RESEARCH

CyLab Usable Privacy & Security Laboratory
HTTP://CUPS.CS.CMU.EDU

Microsoft Research

Authentication ecosystem

Authentication ecosystem

System administrators

Users

Adversaries

Defending users

- System administrators set *password policies*
  - Constraints on passwords

Character requirements

- Common component of policies
  - Length
  - Uppercase, digit, symbol

Character requirements

- Common component of policies
  - Length
  - Uppercase, digit, symbol

- Can’t prevent weak passwords:

Character requirements

- Default policies often use only character requirements

- In Microsoft Active Directory (**3class8**)
  - 3 of the following: uppercase, lowercase, digit, symbol
  - 8 character minimum

These requirements don’t improve security, and they make passwords harder to type

Goal

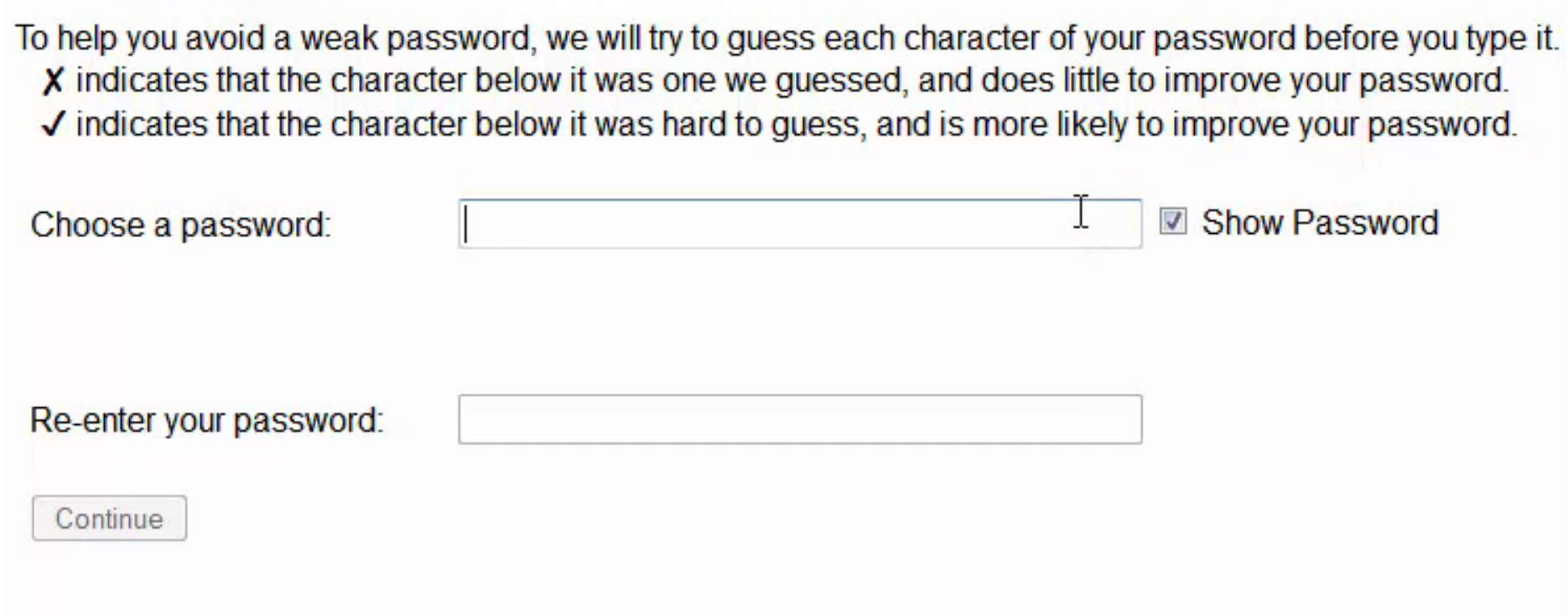
- Focus on weakest passwords
  - Threat model: online attack of an organization
  - Policies should make the weakest passwords harder to guess

Contributions

- Show that character requirements don’t prevent weak passwords
- Introduce Telepathwords
  - Detects weak passwords while providing real-time feedback
- Show that real-time feedback coupled with prevention of common patterns works well
  - Significantly better security than character requirements
  - Not significantly different in usability (after creation)

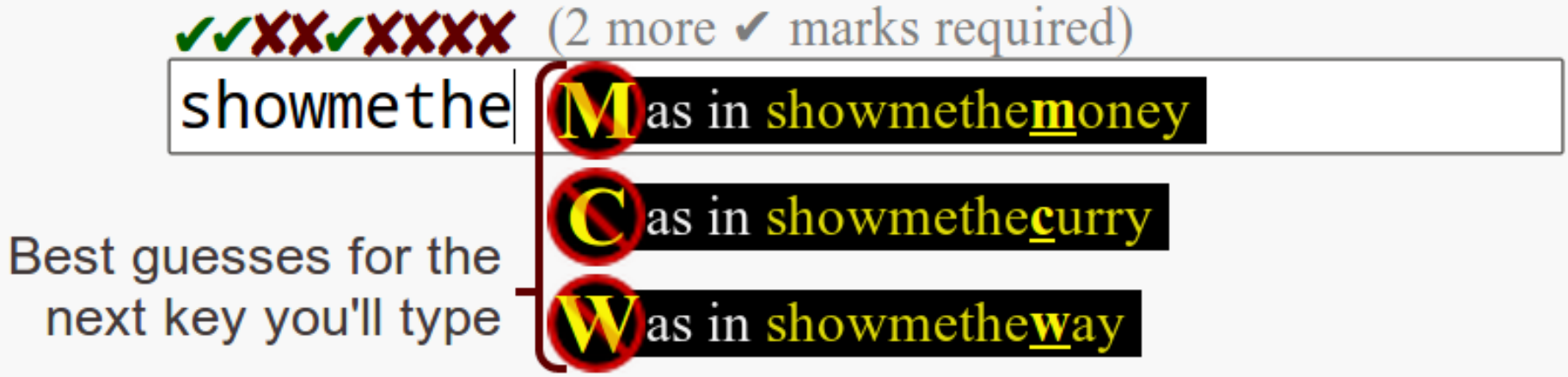
Telepathwords

- Real-time predictions



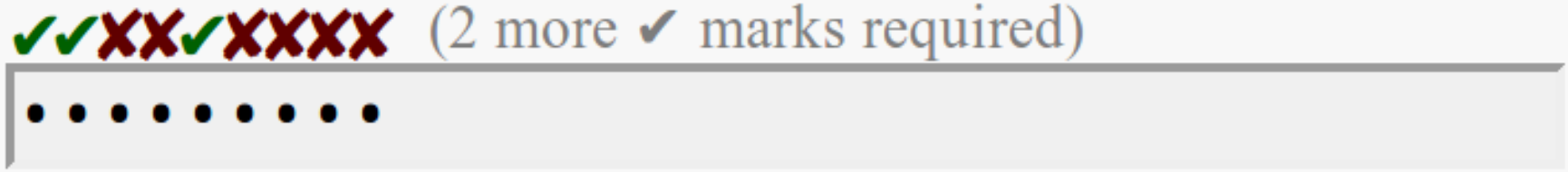
Telepathwords

- Explanations show why password is guessable



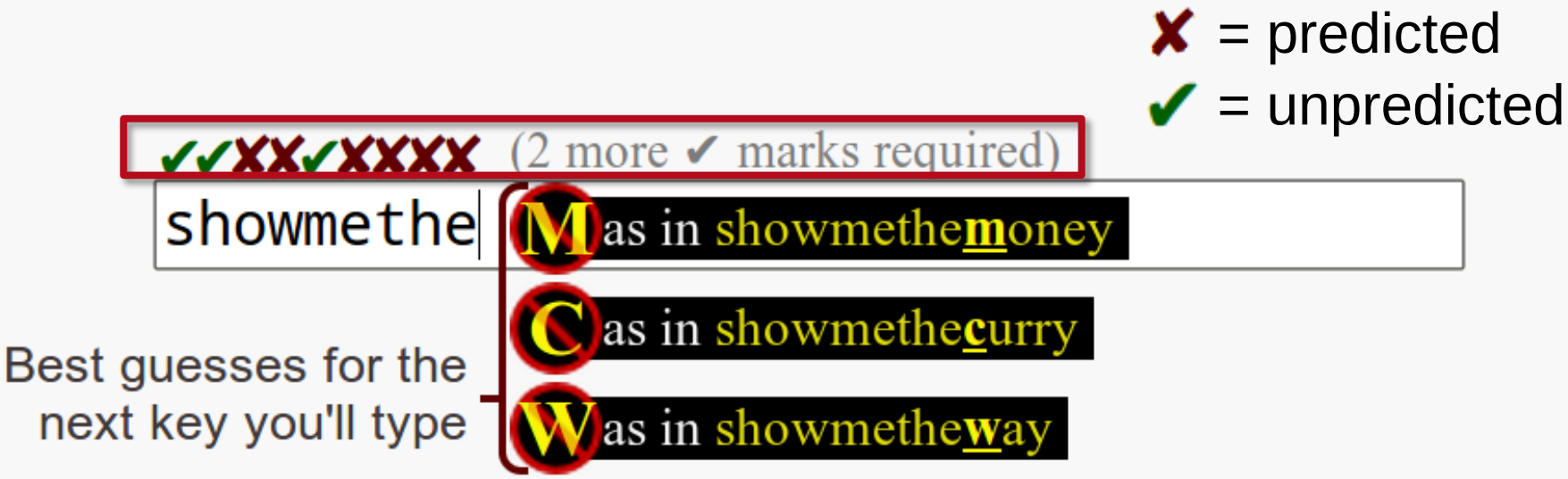
Telepathwords

- Prediction display can be turned off



Telepathwords

- Feedback bar scores password by predictions



- Can enforce a policy by requiring a minimum number of unpredicted characters

Related work

- Character requirements haven’t changed much in 35 years since Morris and Thompson
  - Proposed 6-character minimum [1979]
  - Very little security improvement [Bonneau 2012]

Related work

- Password meters
  - Meters are typically based on character requirements [Ur et al. 2012]
  - No consistency across meters [de Carnevalet and Mannan 2013]

Related work

- Password meters
  - Meters are typically based on character requirements [Ur et al. 2012]
  - No consistency across meters [de Carnevalet and Mannan 2013]
- ***zxcvbn*** entropy estimator [Wheeler 2012]

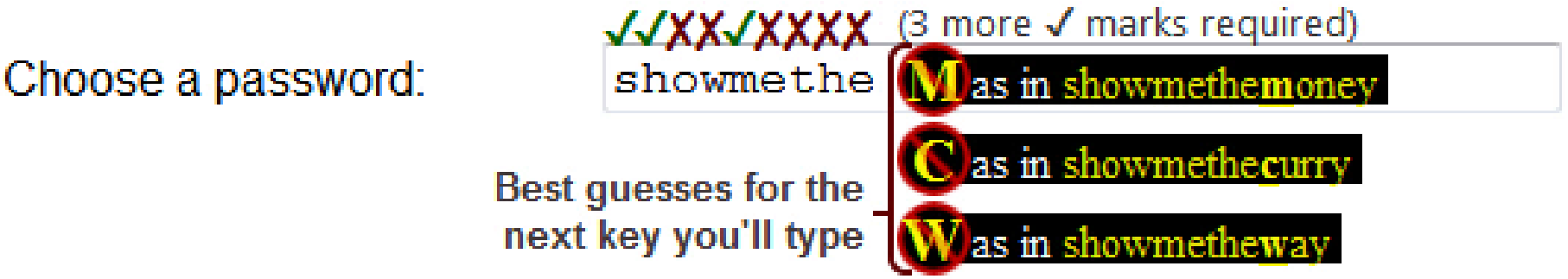
Password meters don’t explain their scores

Generating predictions

- Multiple, weighted “predictors” produce next-character guesses and likelihood scores

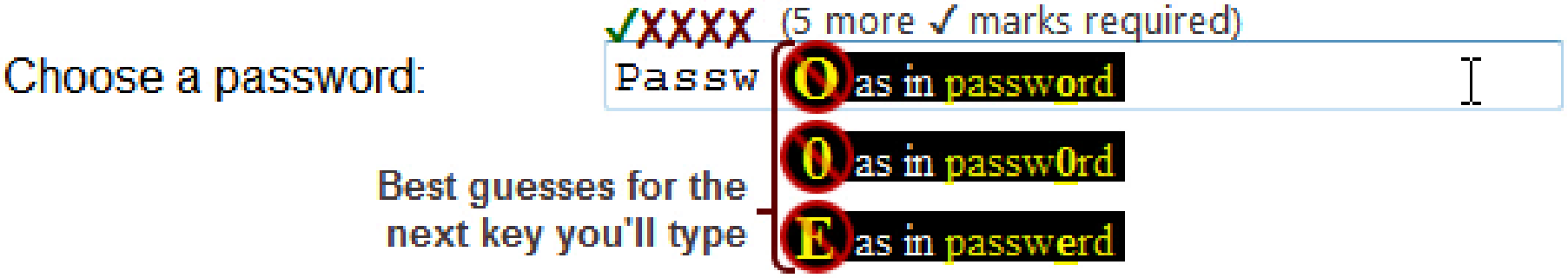
Generating predictions

- Search query n-grams



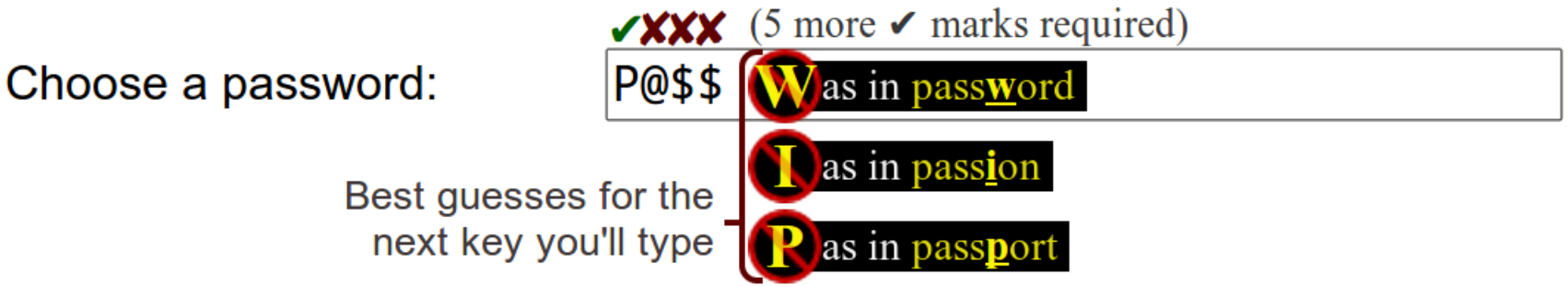
Generating predictions

- Password sets



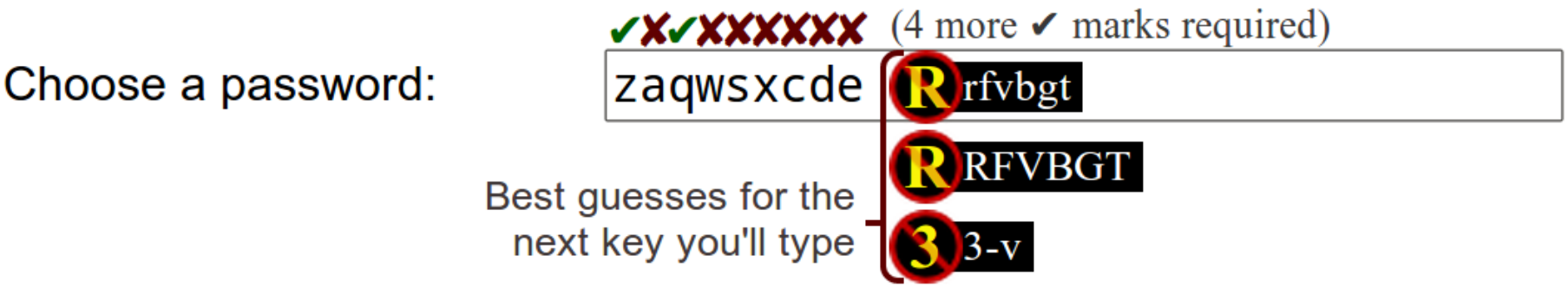
Generating predictions

- Common substitutions (s -> $, a -> @, etc.)



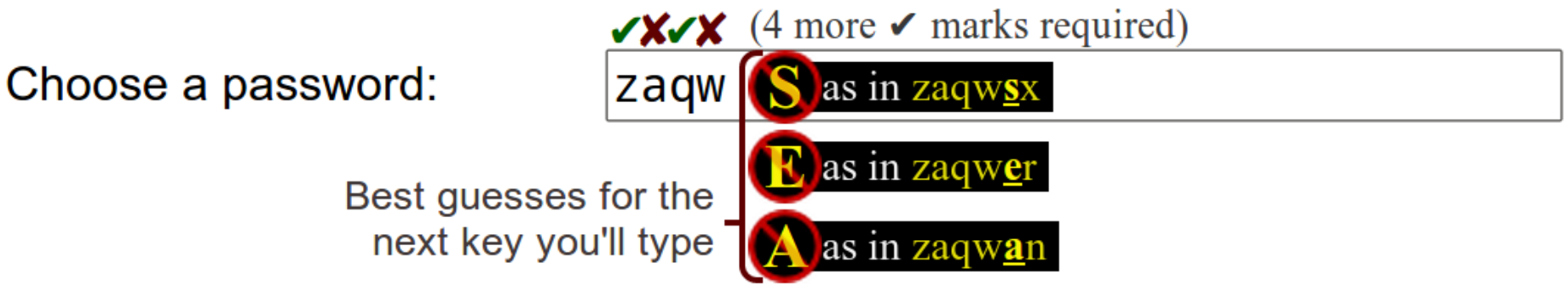
Generating predictions

- Keyboard patterns



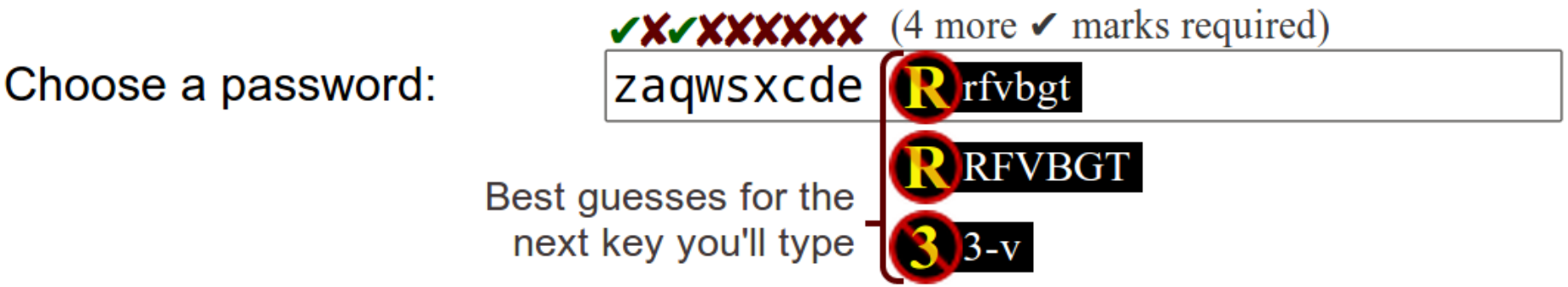
Generating predictions

- Keyboard patterns



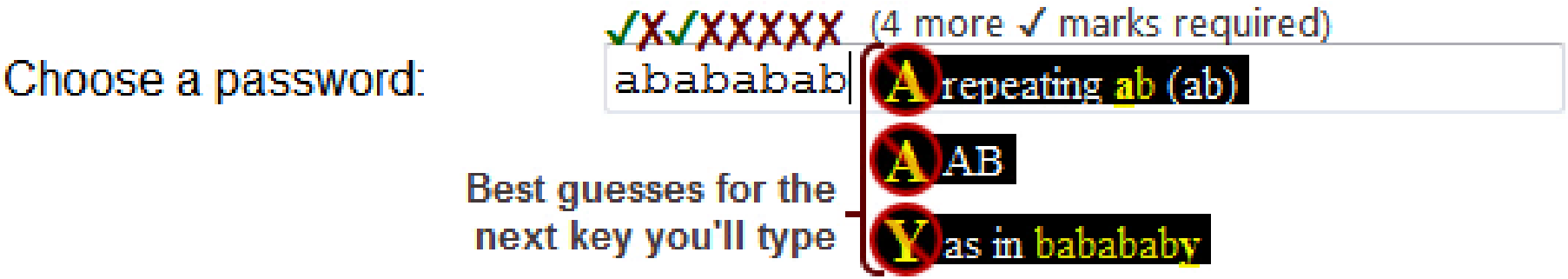
Generating predictions

- Keyboard patterns



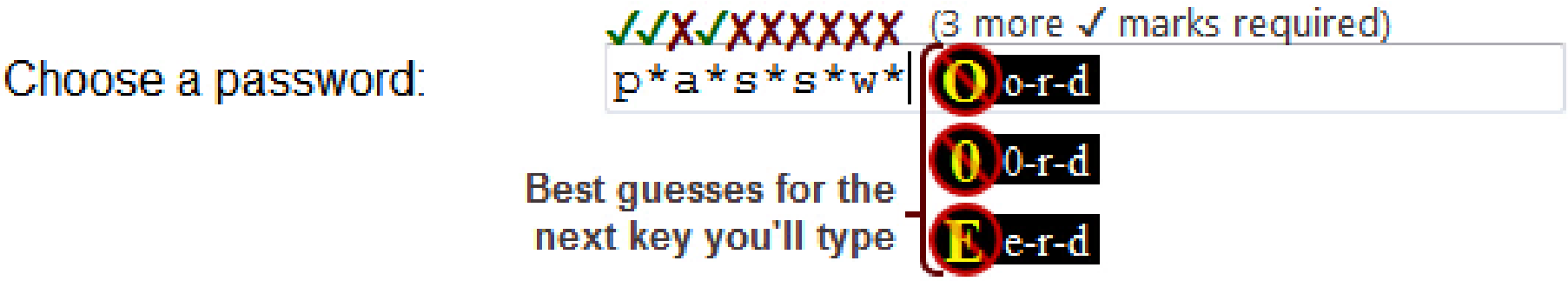
Generating predictions

- Repeating patterns



Generating predictions

- Interleaving strings



Generating predictions

- Can cover many behaviors and easily add

- Many possible ways to implement, ours is just one example

Evaluation

- December 2013 - deployed as a public website
  https://telepathwords.research.microsoft.com

- Feburary 2014 - Mechanical Turk study
  - CMU branded, using Javascript API

Policies

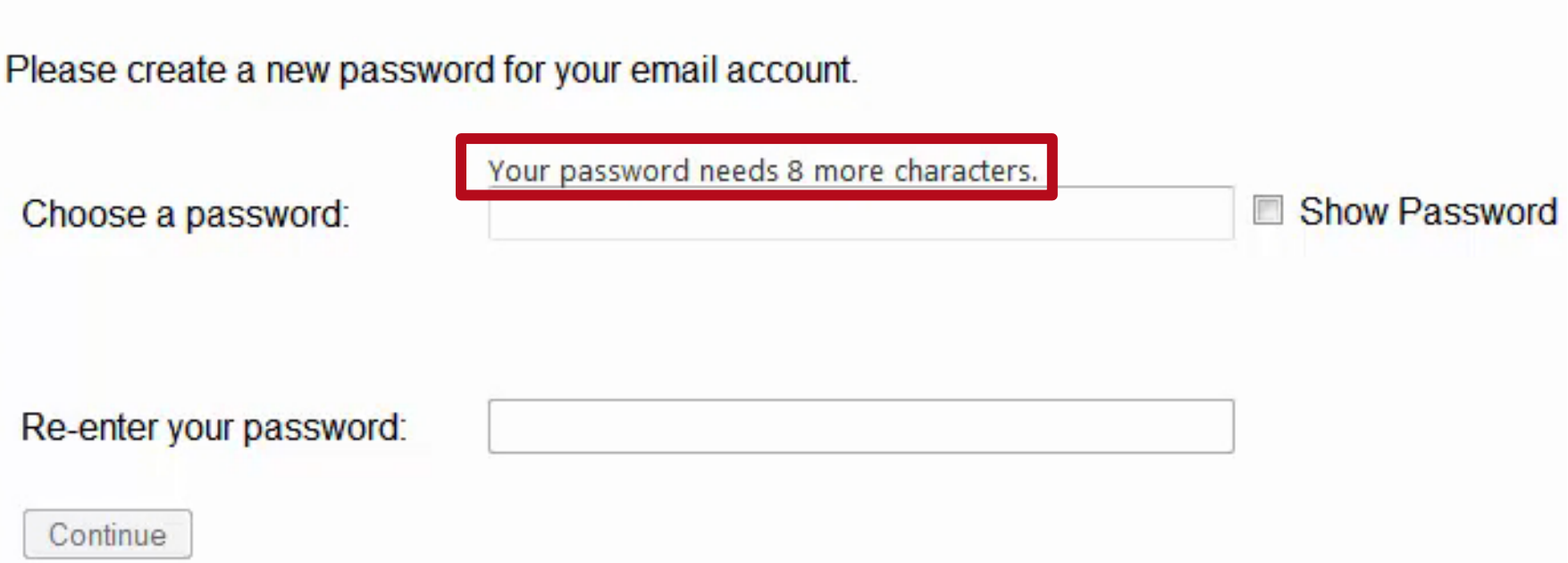
- 6 policy conditions (2 Telepathwords)

- All conditions included some visual feedback

Condition: basic8

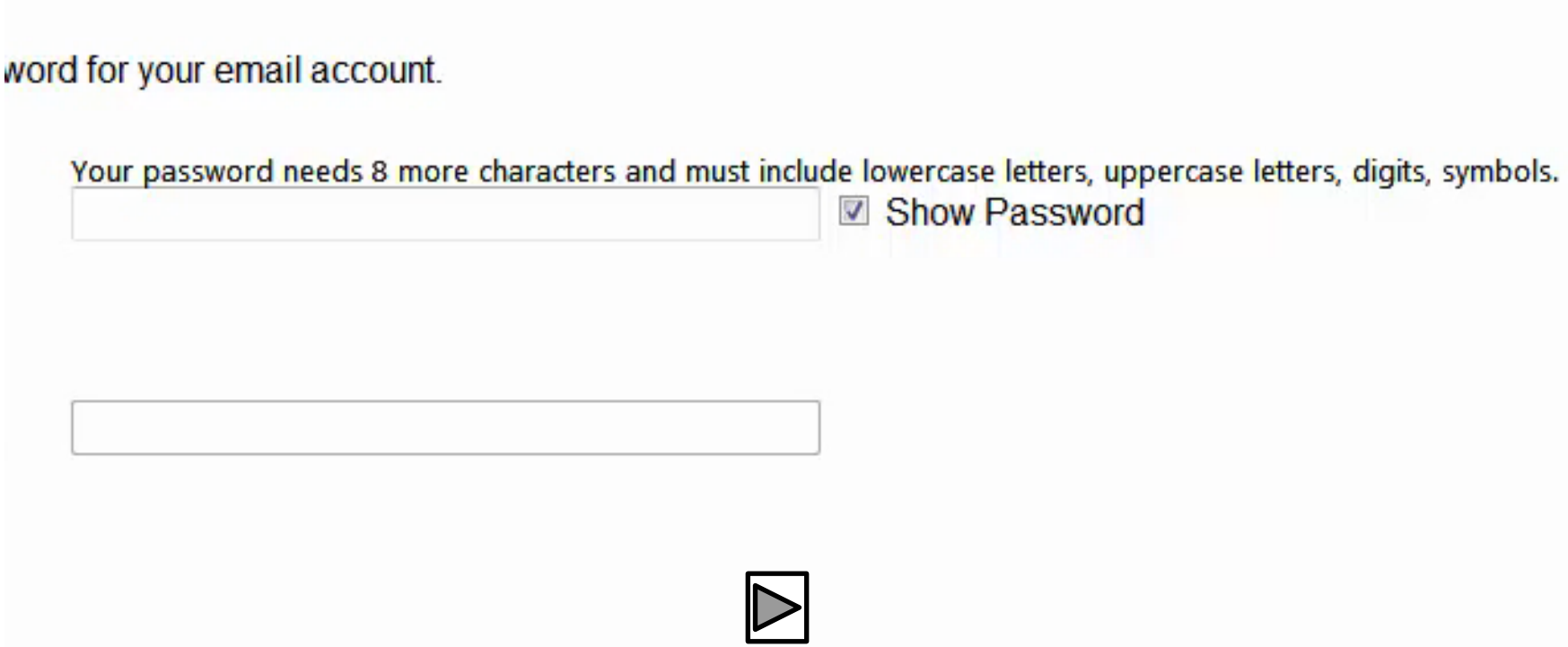
Requirements

- Minimum 8 characters



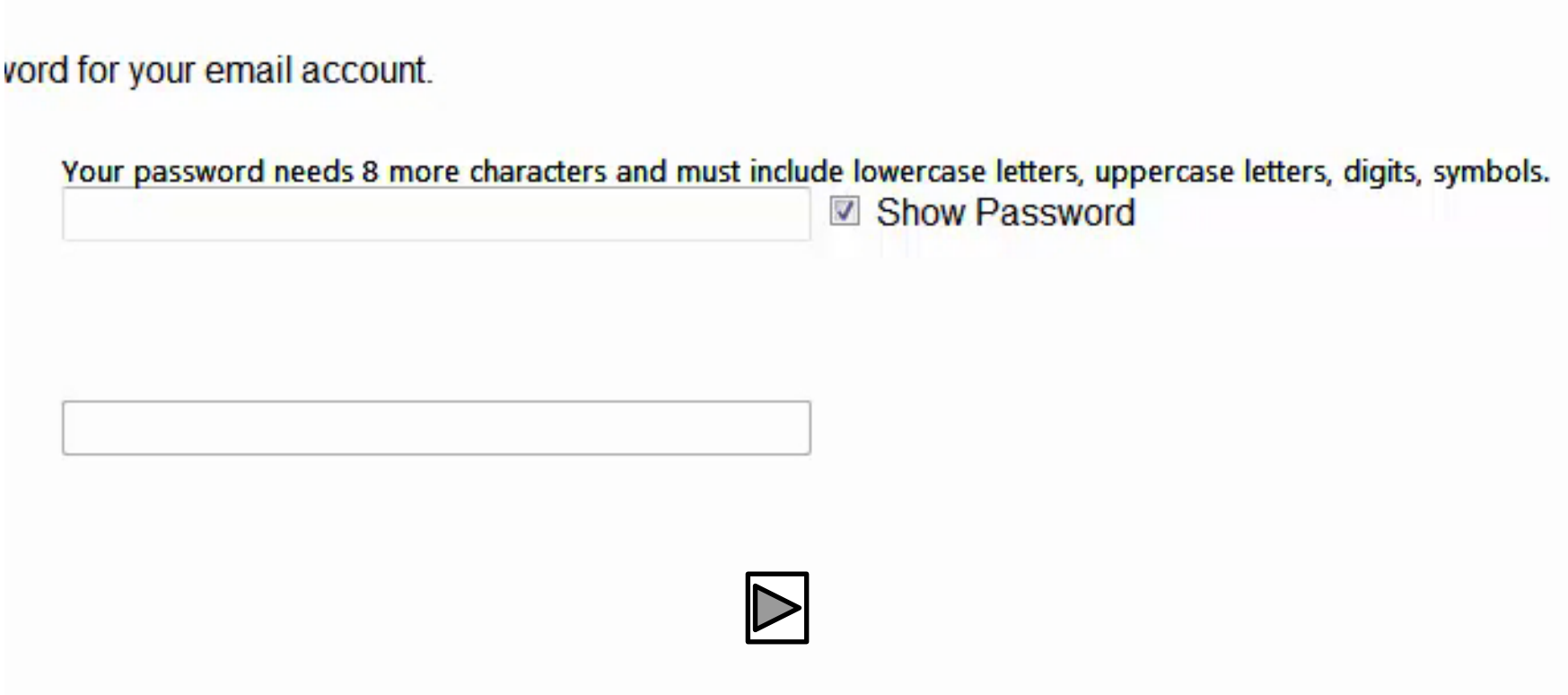
Condition: 3class8

- Minimum 8 characters
- Must contain at least 3 of the following: lowercase, uppercase, digit, symbol



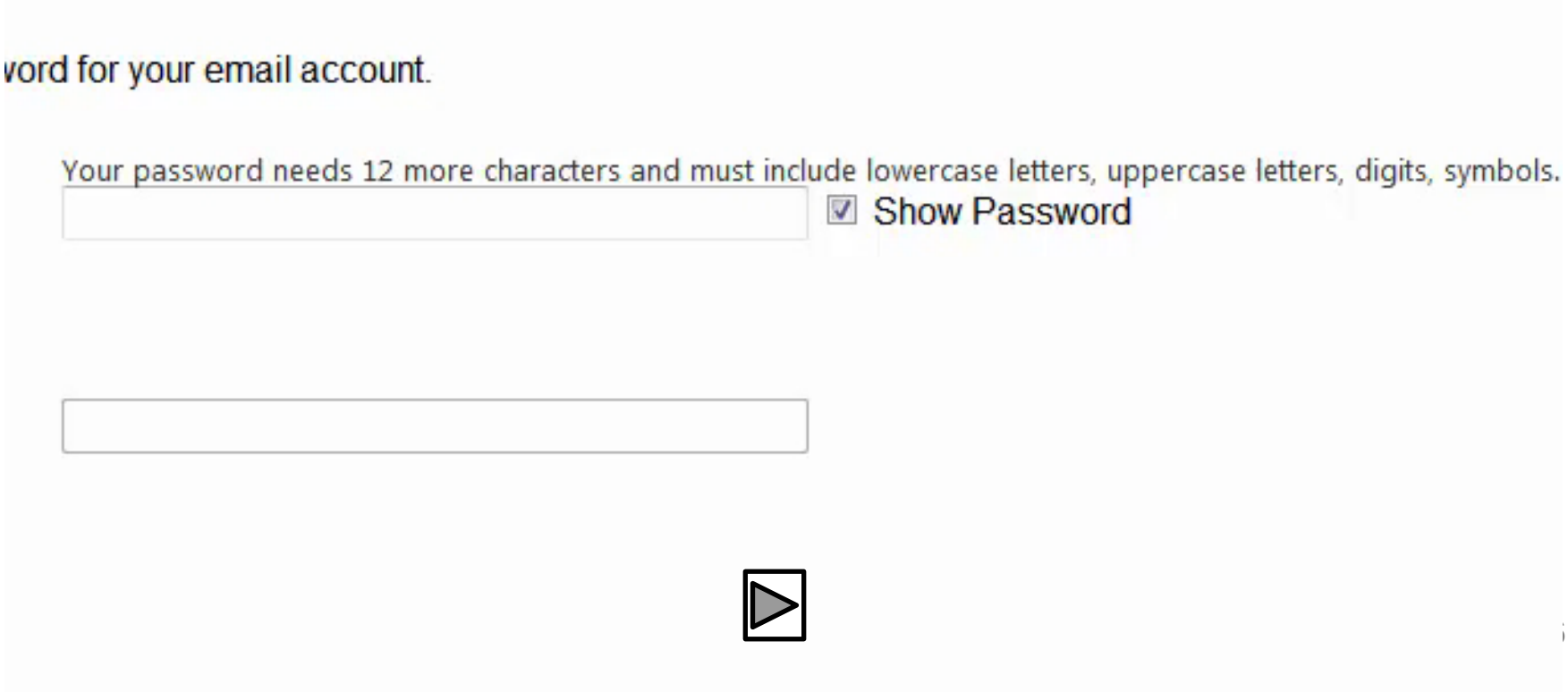
Condition: 3class8-d

- Minimum 8 characters
- Must contain at least 3 of the following: lowercase, uppercase, digit, symbol
- Letters in password must not be in a dictionary



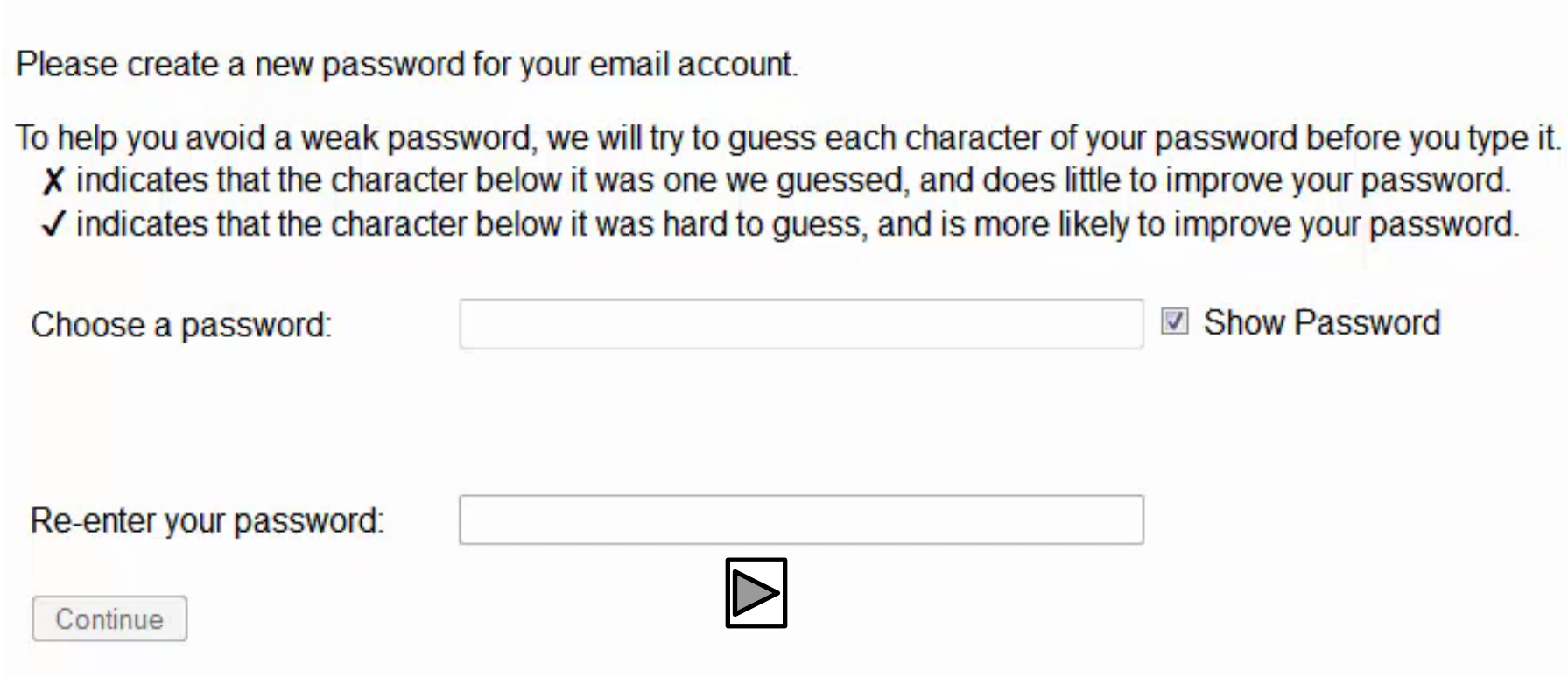
Condition: 3class12

- Minimum 12 characters
- Must contain at least 3 of the following: lowercase, uppercase, digit, symbol

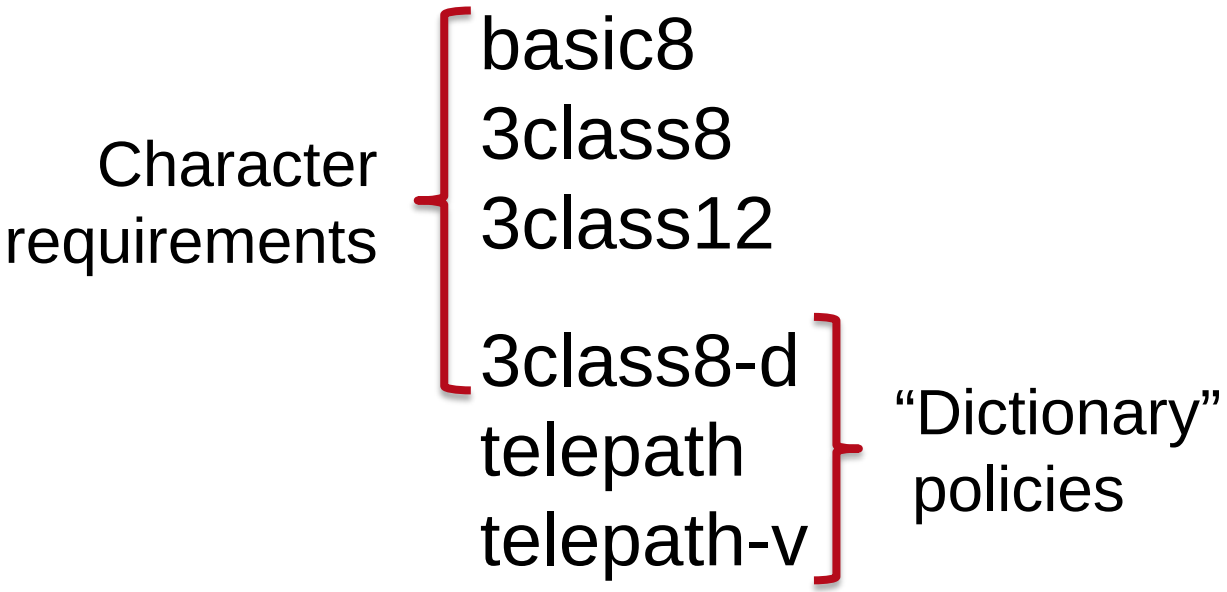


Conditions: telepath, telepath-v

- Minimum 6 unpredicted characters
- “Show Password” checked by default in telepath-v



Conditions


Character requirements
basic8
3class8
3class12
3class8-d
telepath
telepath-v
“Dictionary” policies

Evaluation

- N = 2,844 (started) / 2,560 (finished)
- Median age = 27 (limited to 18+)
- 60% male, 44% with Bachelor's or above
- Required 95% acceptance rate and U.S. location

Study design

- Hypothetical email scenario for password creation

Steps:

1. Create a password under a randomly assigned condition
2. Take a survey
3. Recall password
4. Return in two days

Policy metrics

Security

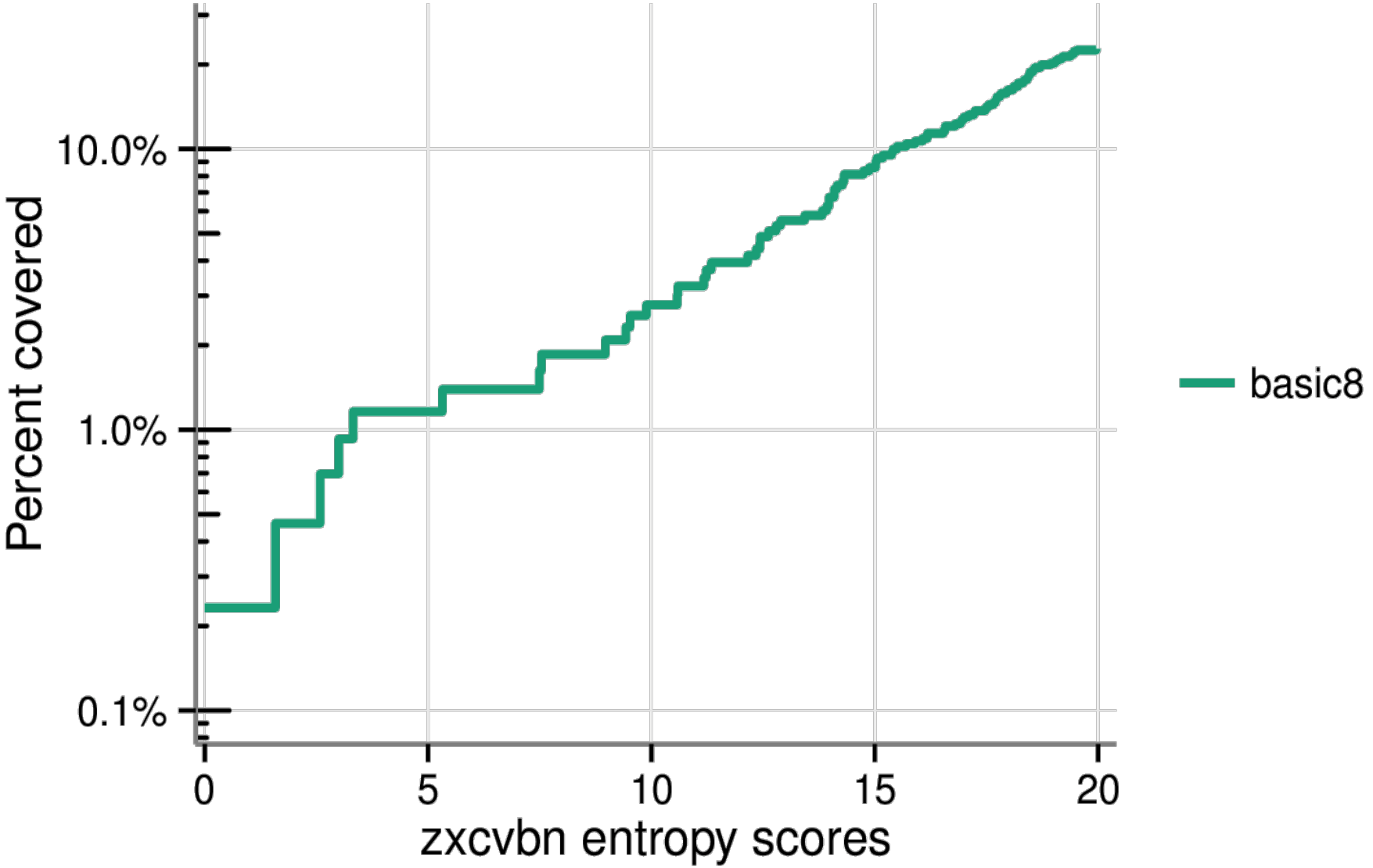
- Weir+ guessability
  - Refinement of [Weir et al. 2009]
  - Minimum number of guesses needed for single password, 2.5%, 5%, and 10%
- ***zxcvbn*** entropy estimate
  - Min-entropy; 2.5th, 5th, 10th percentiles
- Probability metrics were not viable

Policy metrics

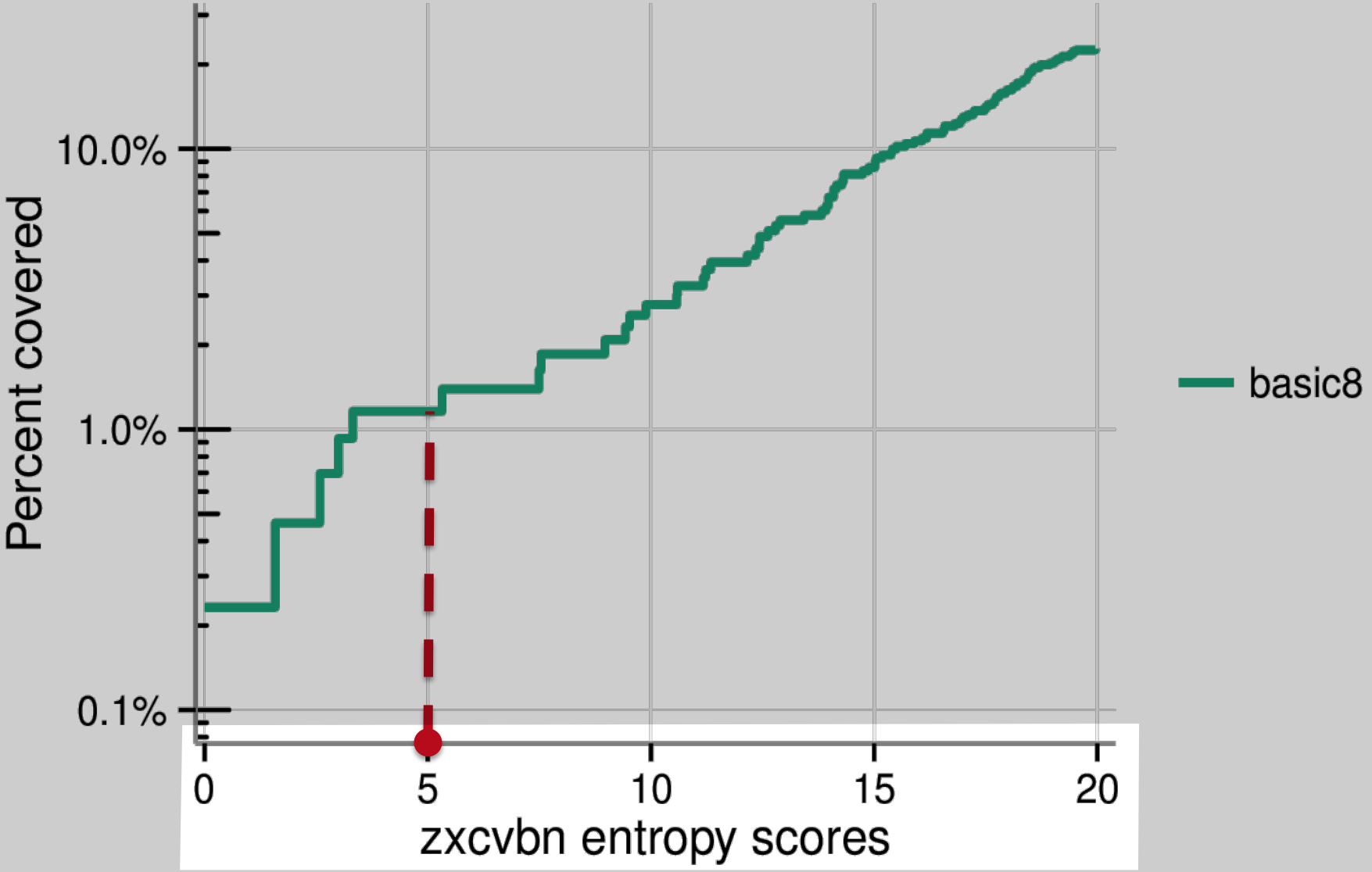
Usability

- Creation difficulty
- Did participants find it insightful?
- Recall difficulty

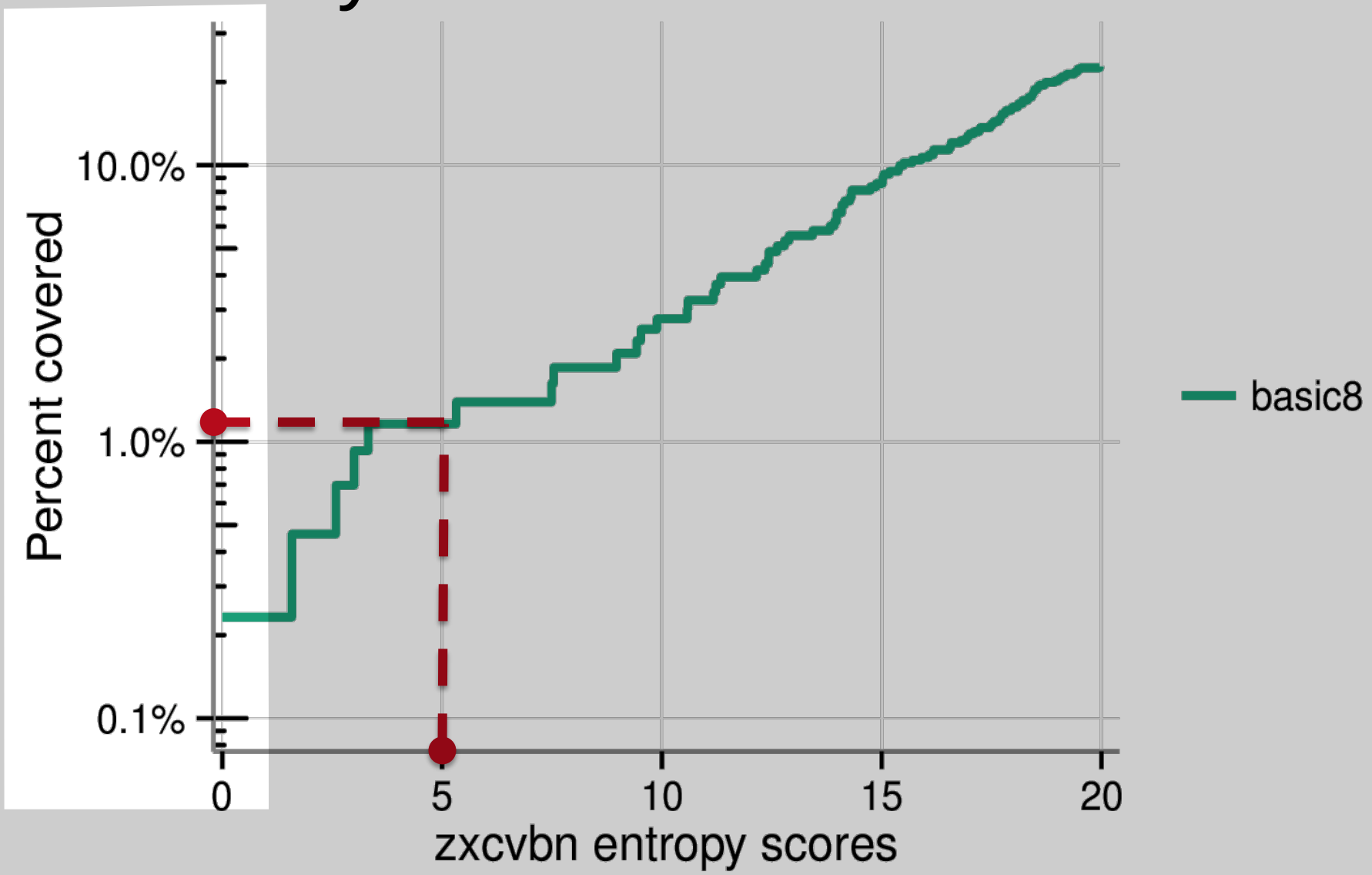
Security results



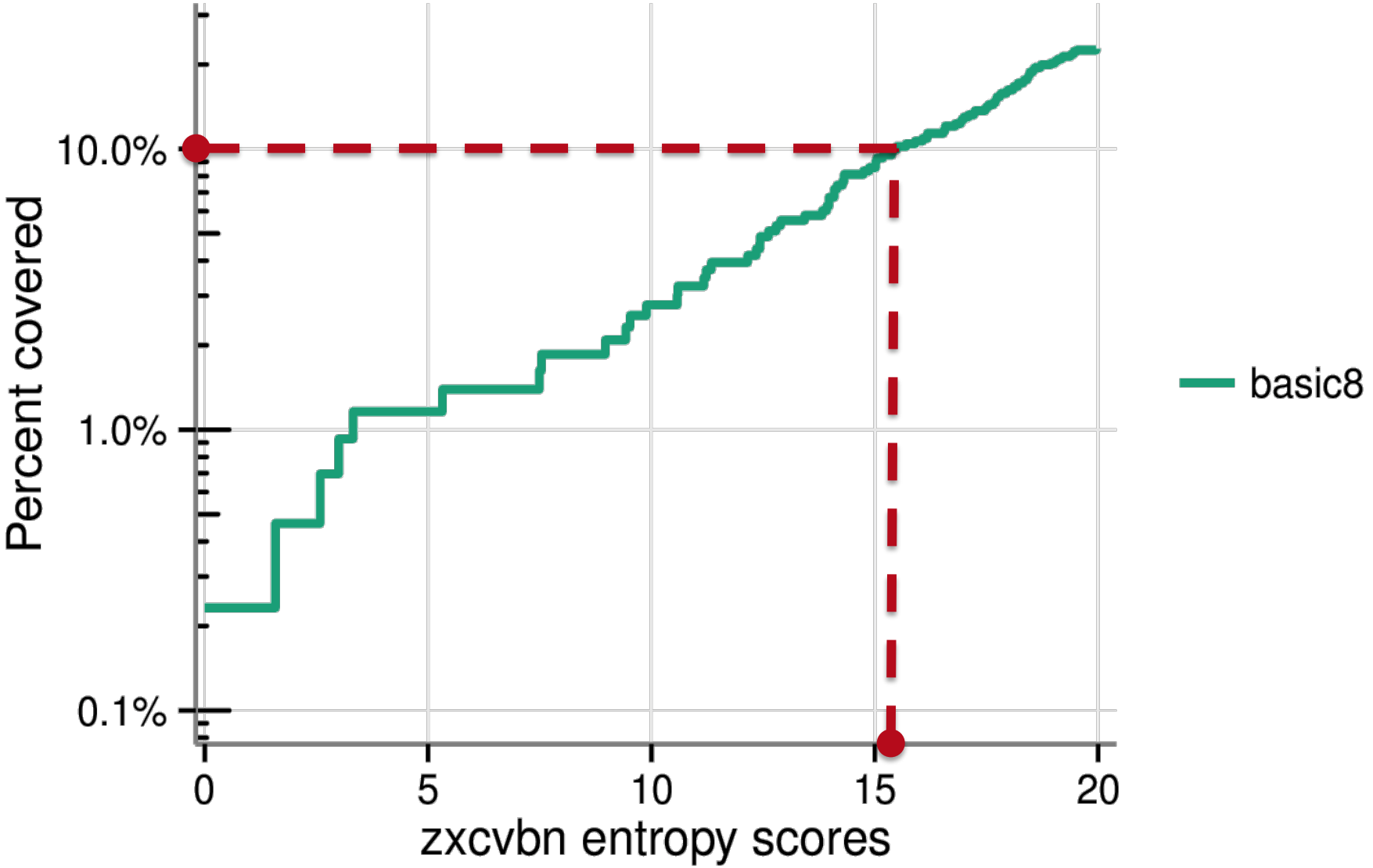
Security results



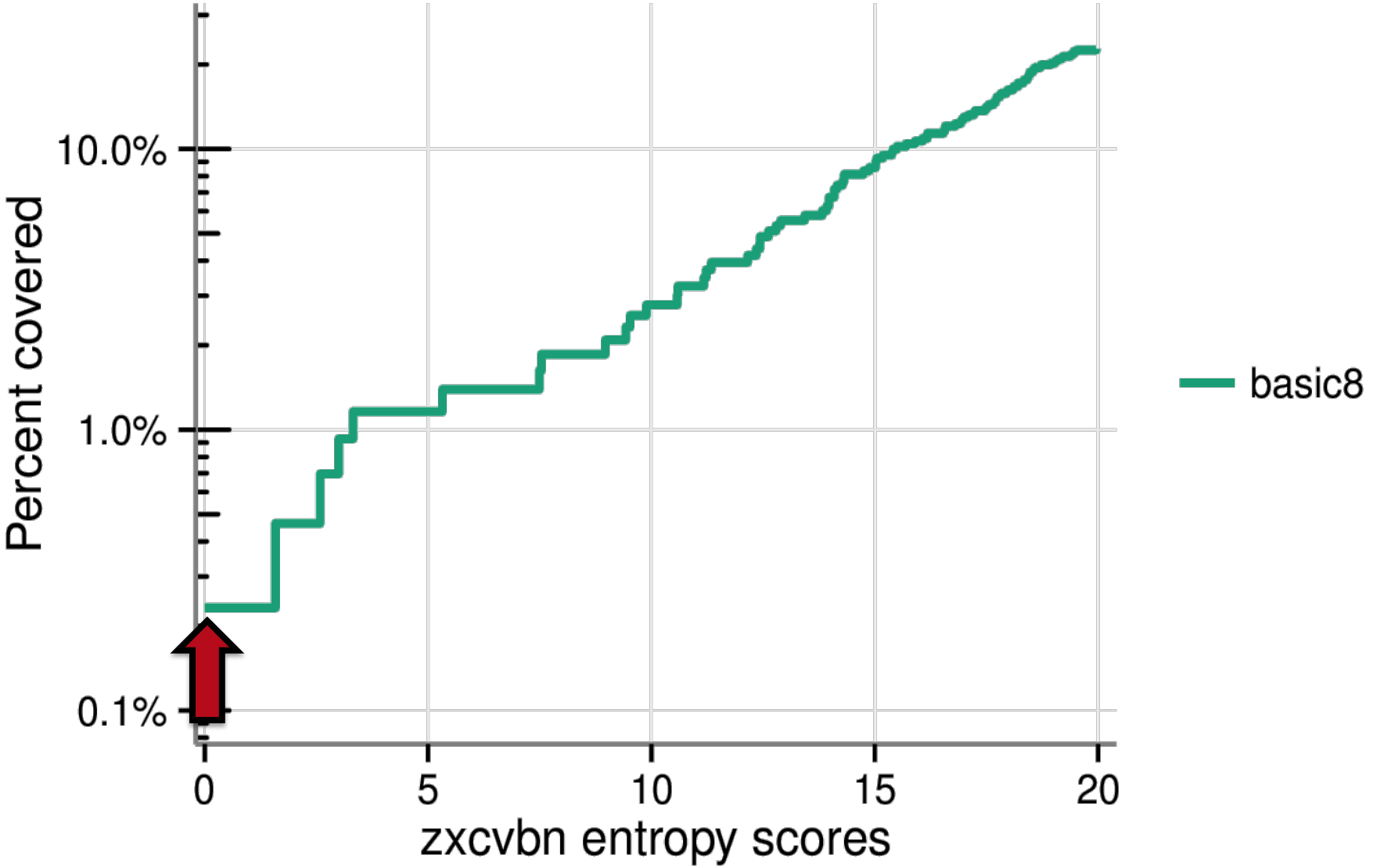
Security results



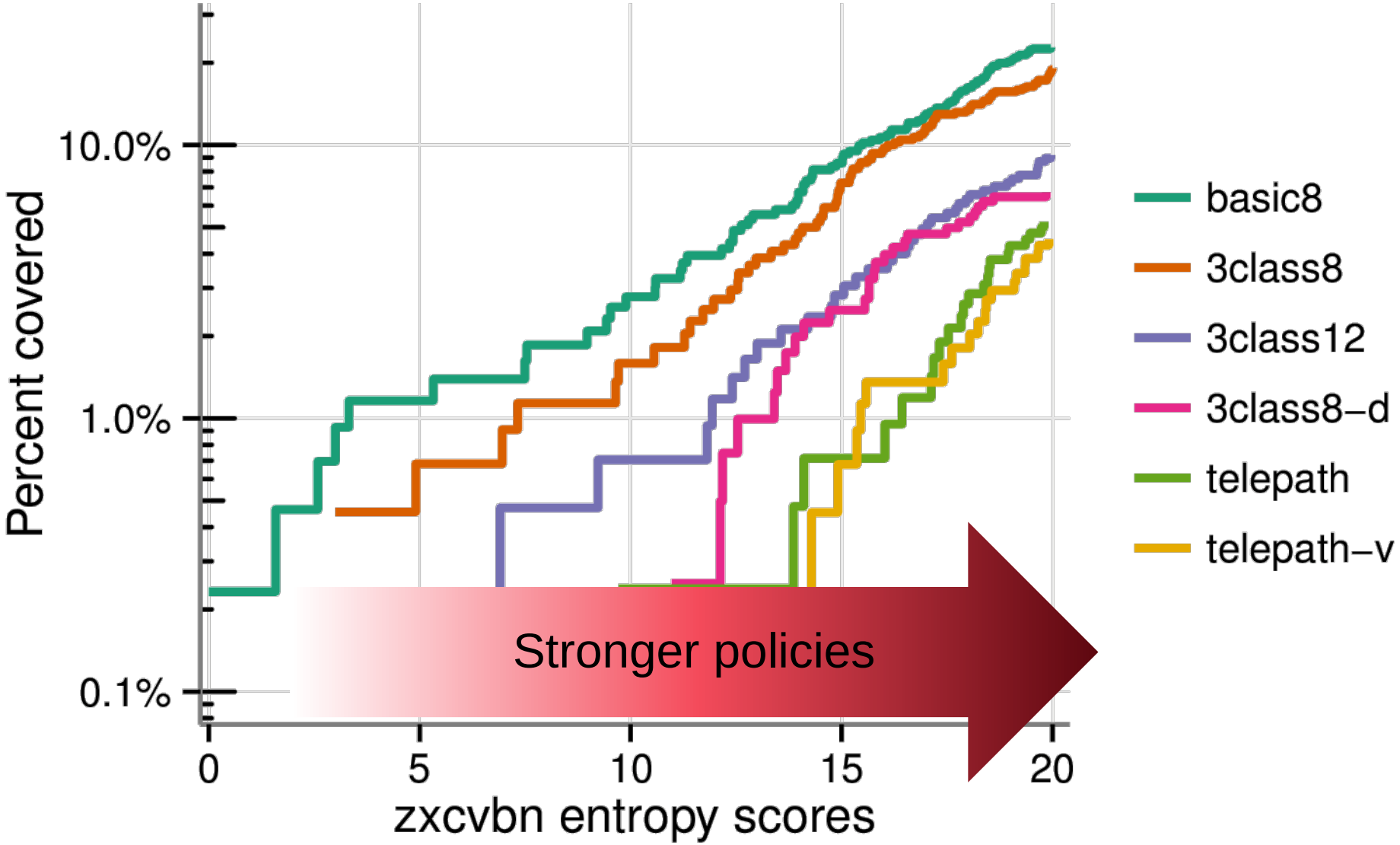
Security results



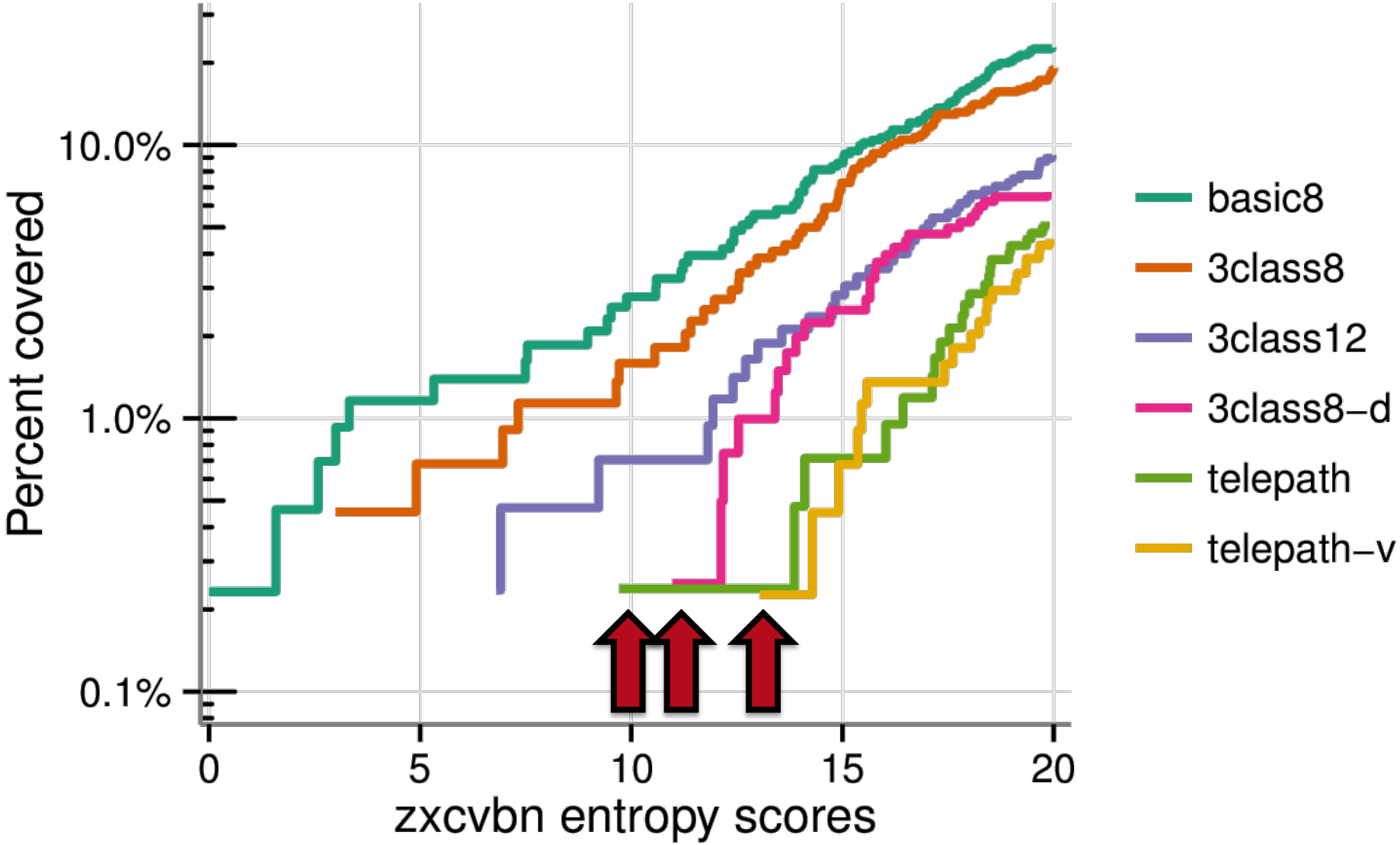
Security results



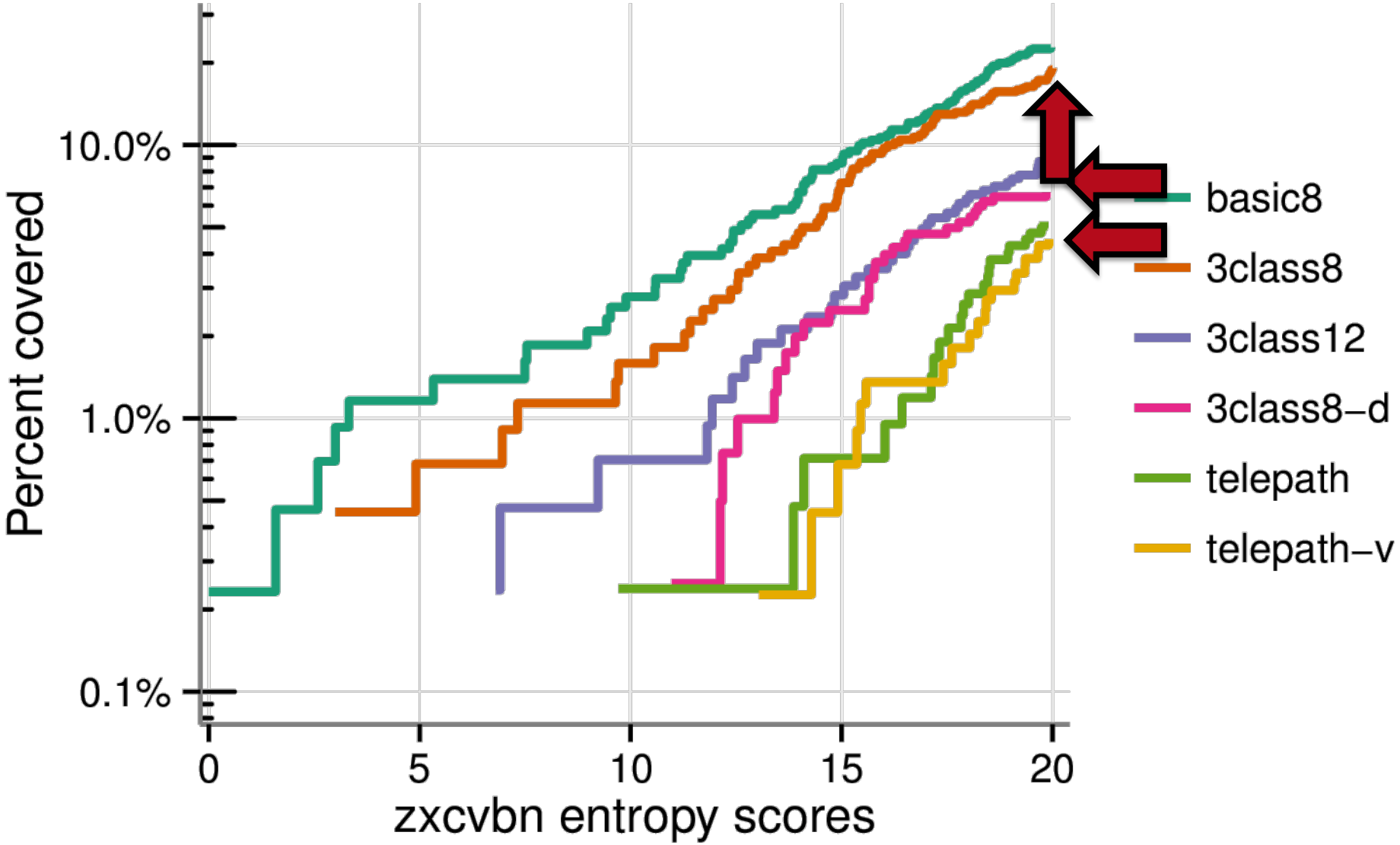
Security results



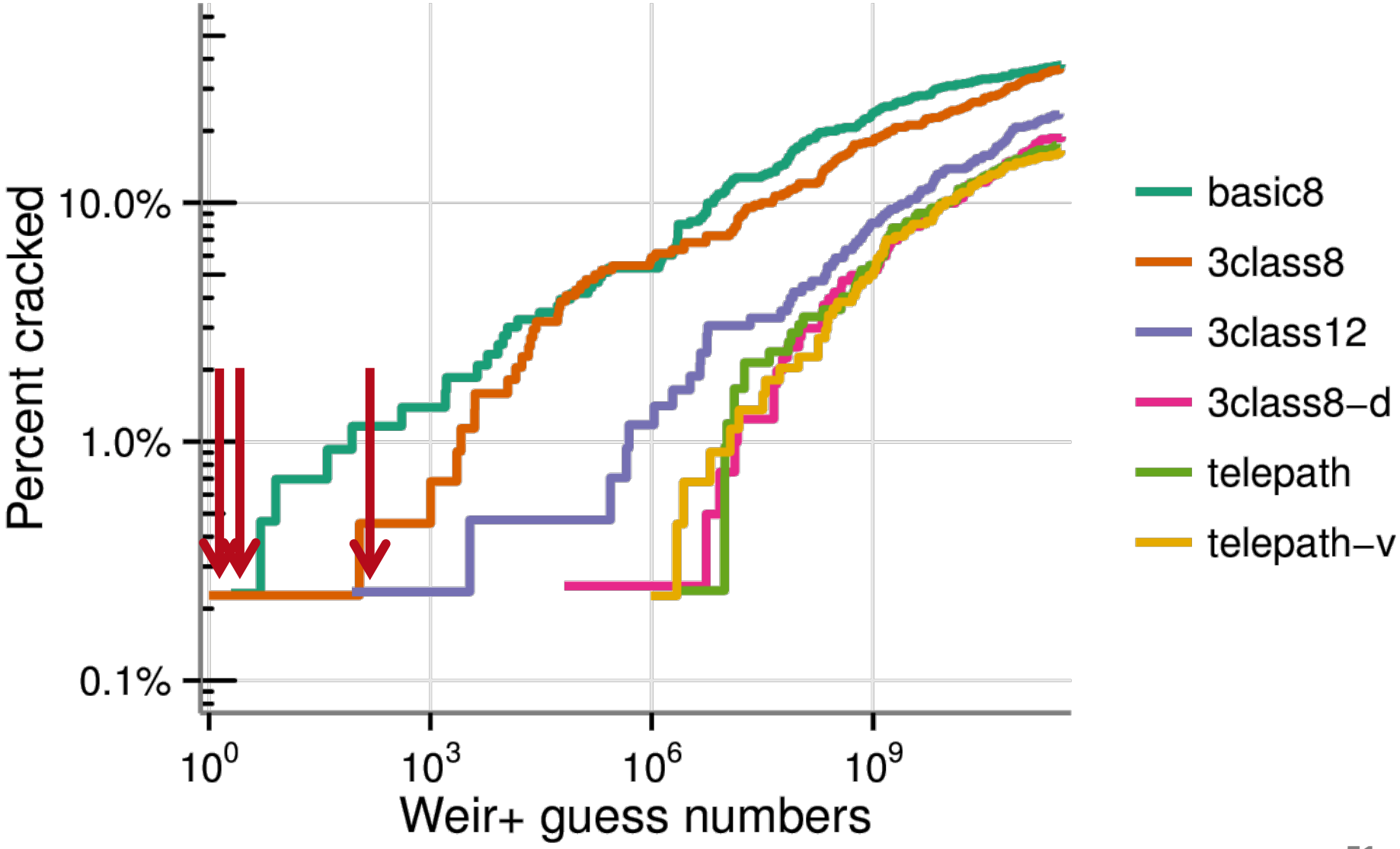
Security results



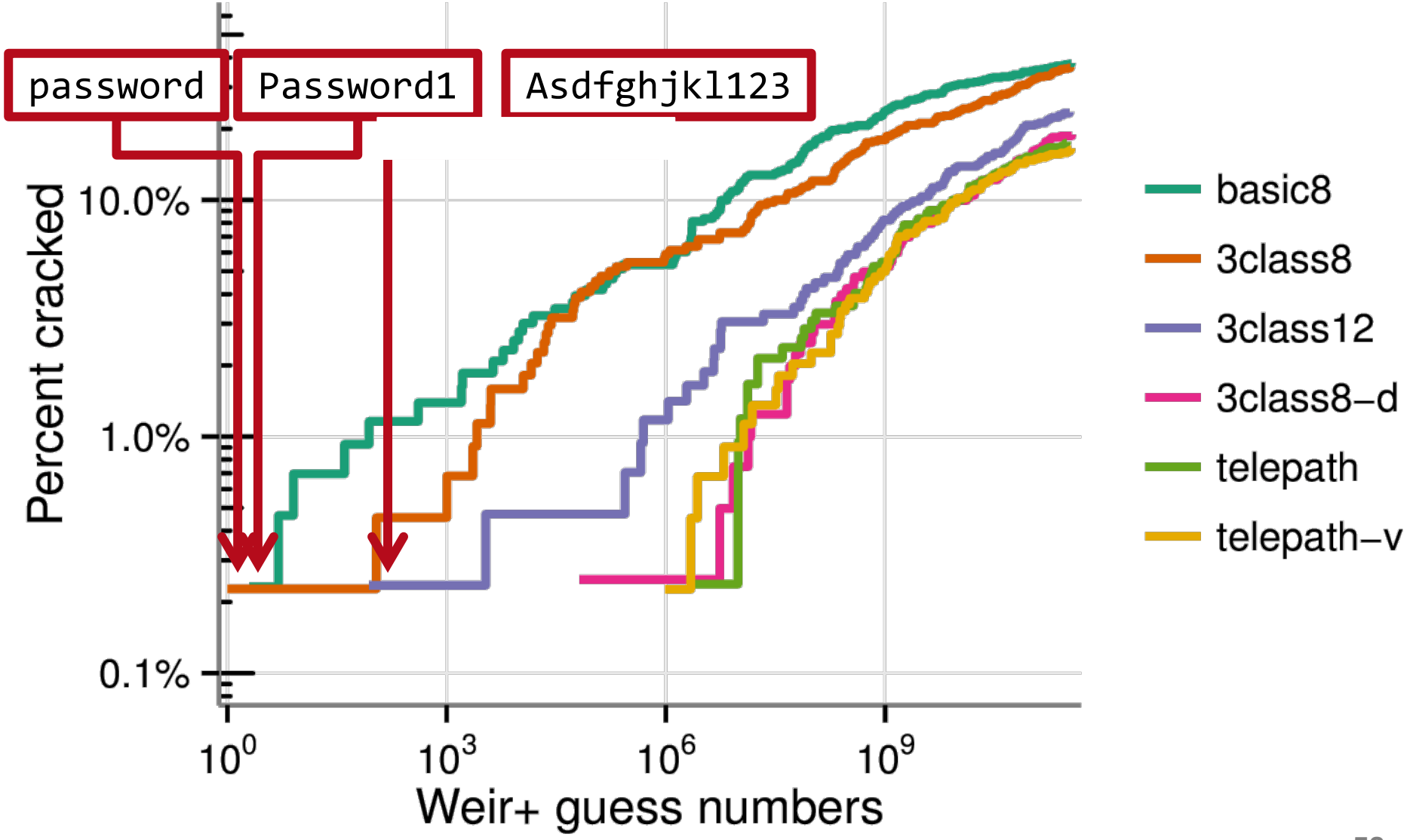
Security results



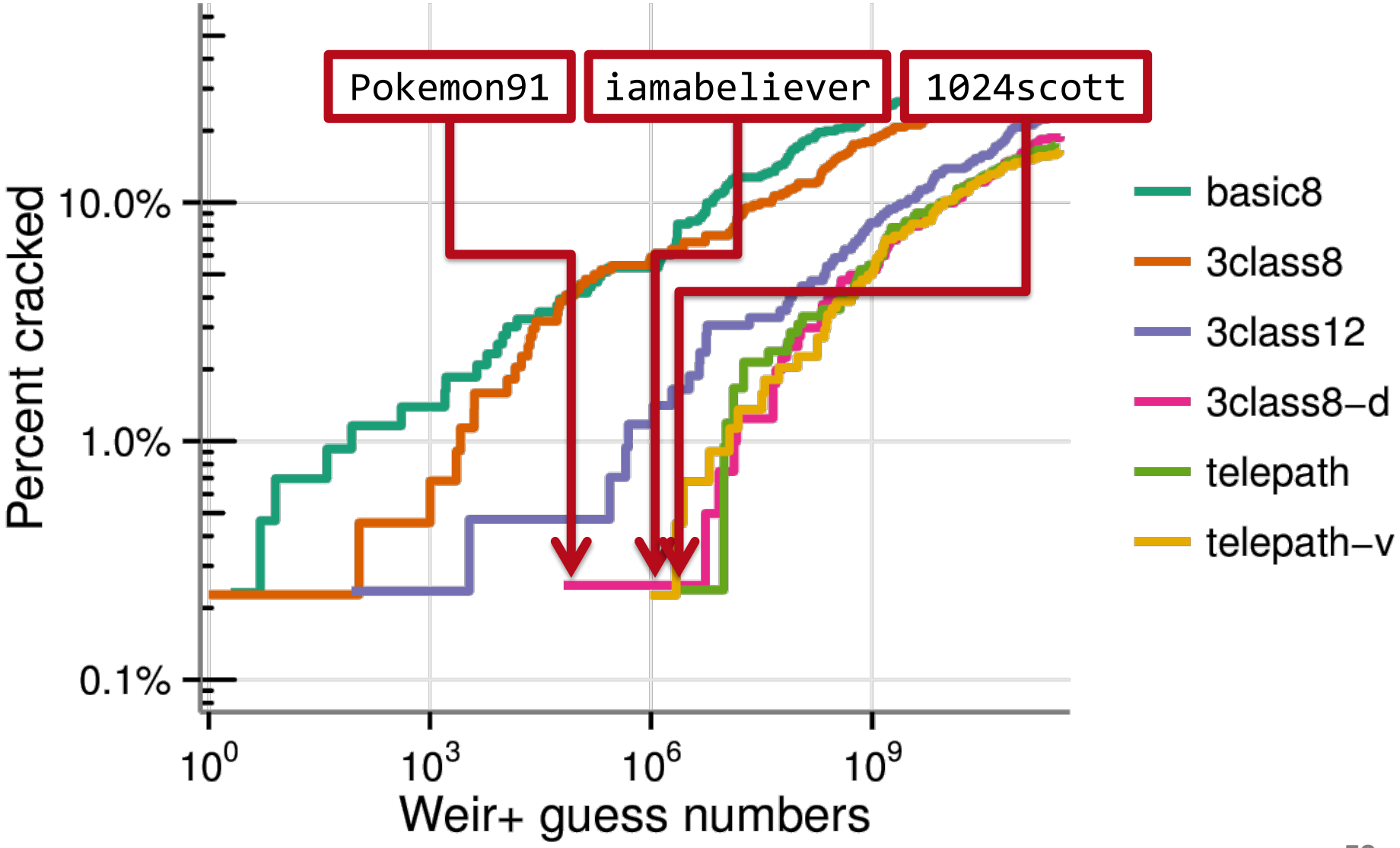
Security results



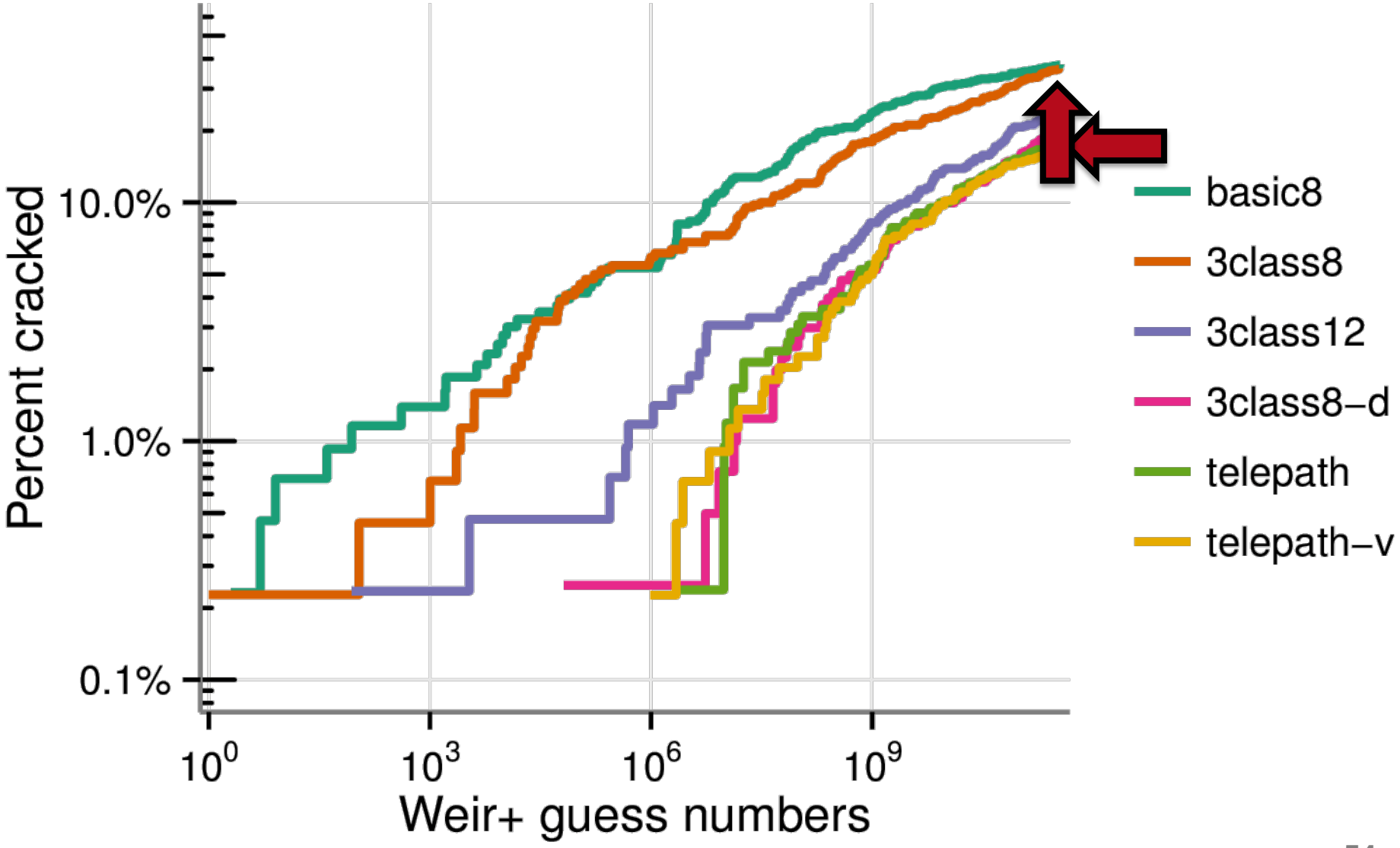
Security results



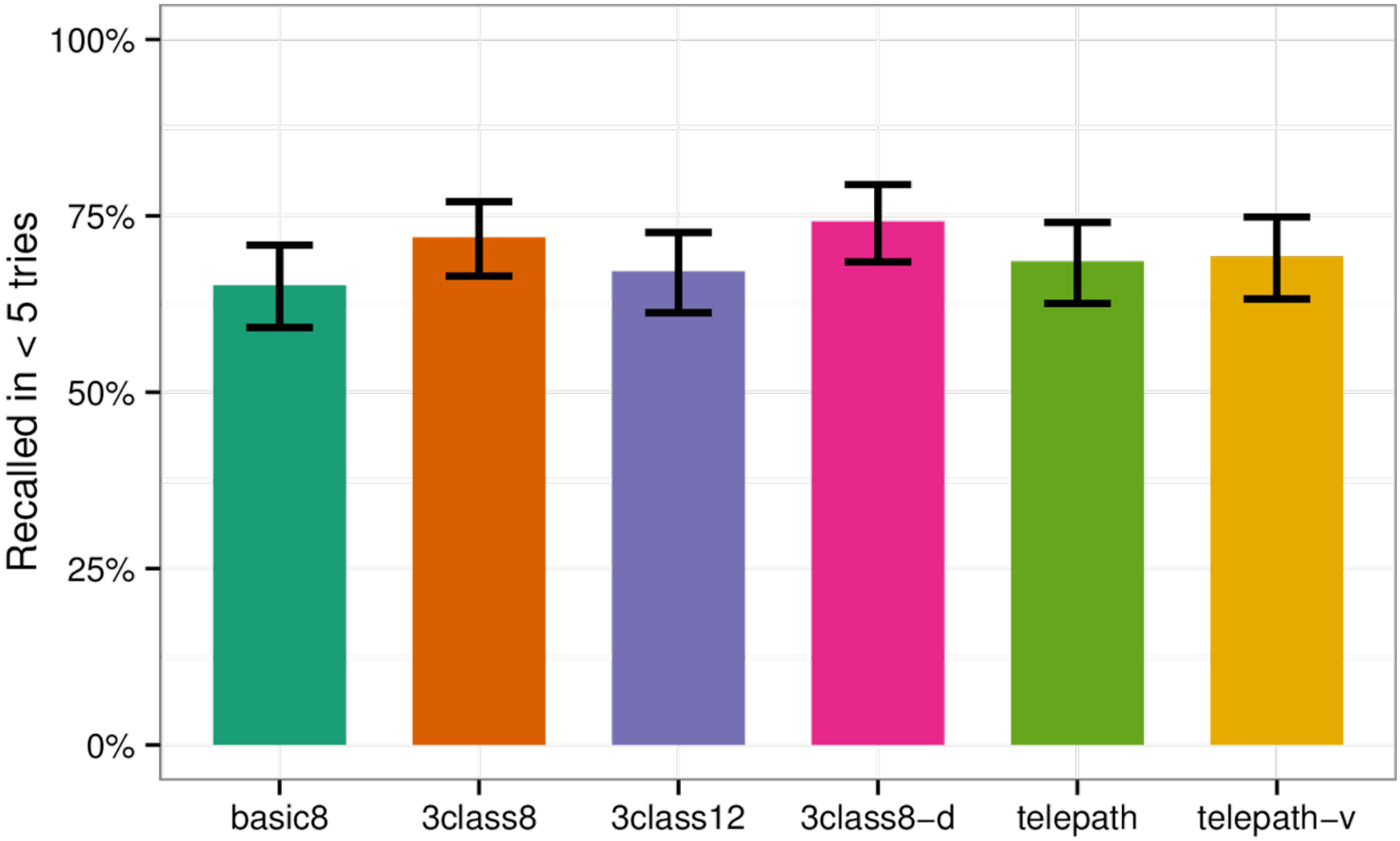
Security results



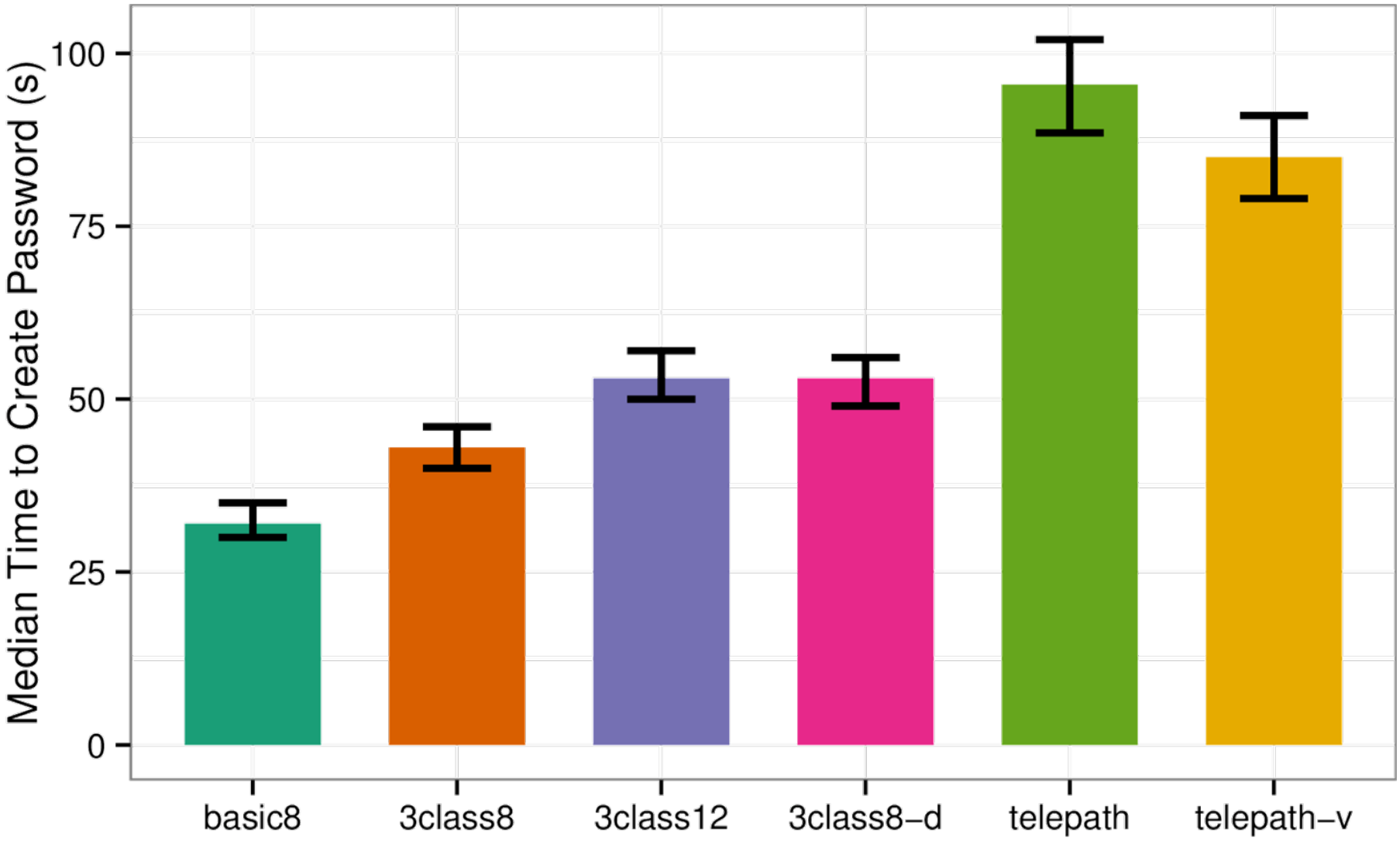
Security results



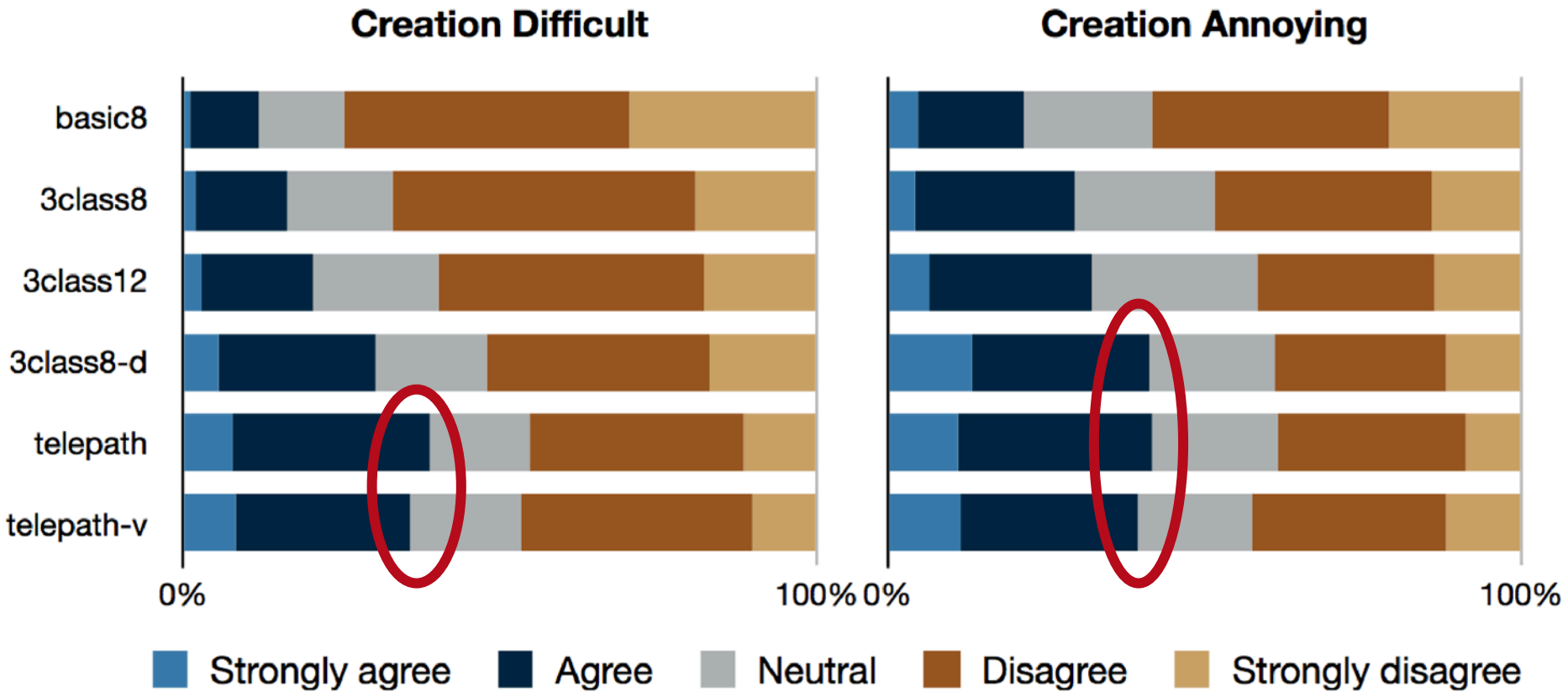
Usability: Recall after 2—5 days



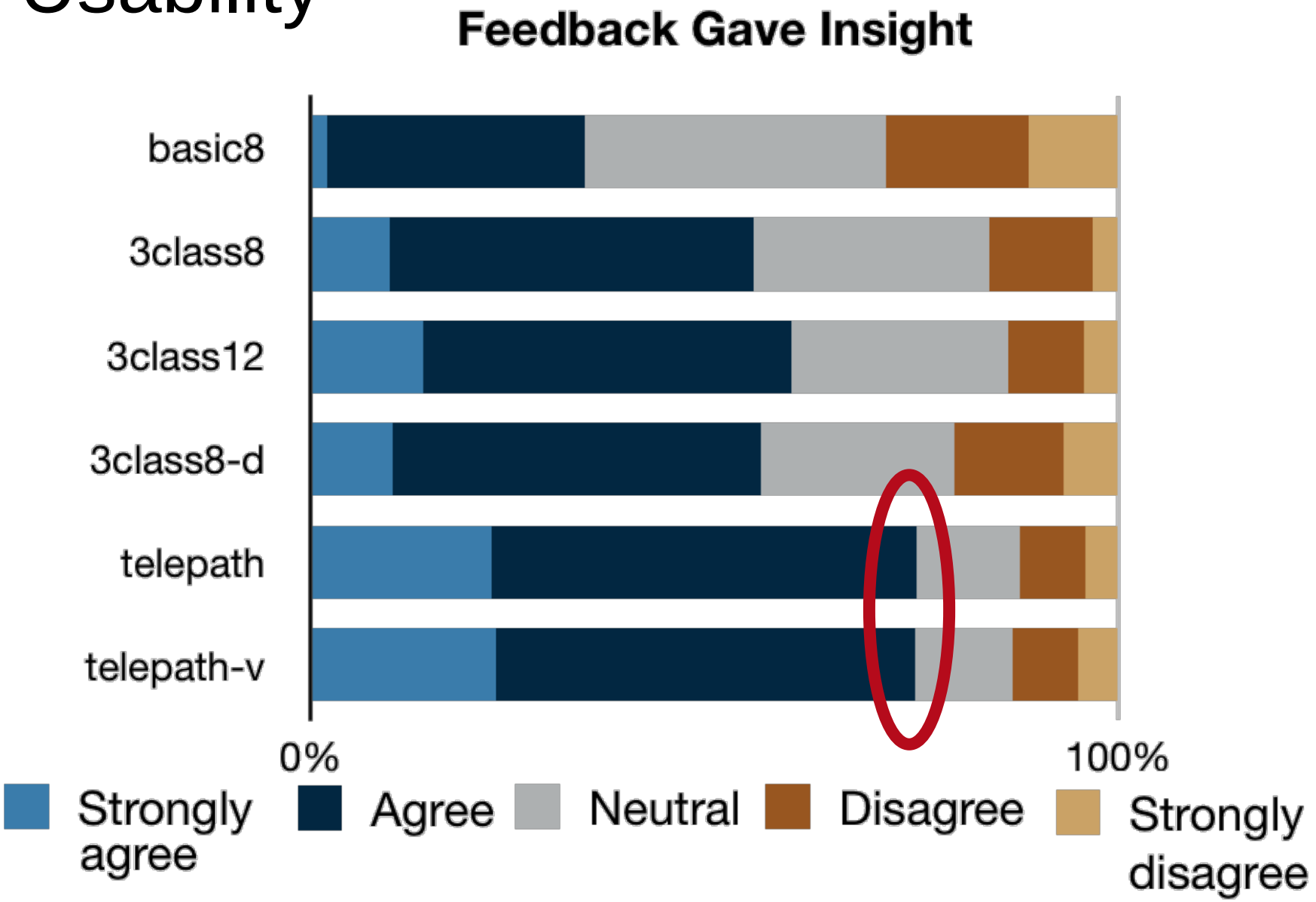
Usability: Creation time



Usability: Creation sentiment



Usability


Feedback Gave Insight
basic8
3class8
3class12
3class8-d
telepath
telepath-v
0%
100%
Strongly agree
Agree
Neutral
Disagree
Strongly disagree

Conclusions

- “Dictionary” policies with real-time feedback can help users avoid weak passwords
  - Usability cost on creation
  - Telepathwords’ feedback gave insight into password strength

Conclusions

- “Dictionary” policies with real-time feedback can help users make stronger passwords
  - Usability cost on creation
  - Telepathwords’ feedback gave insight into password strength

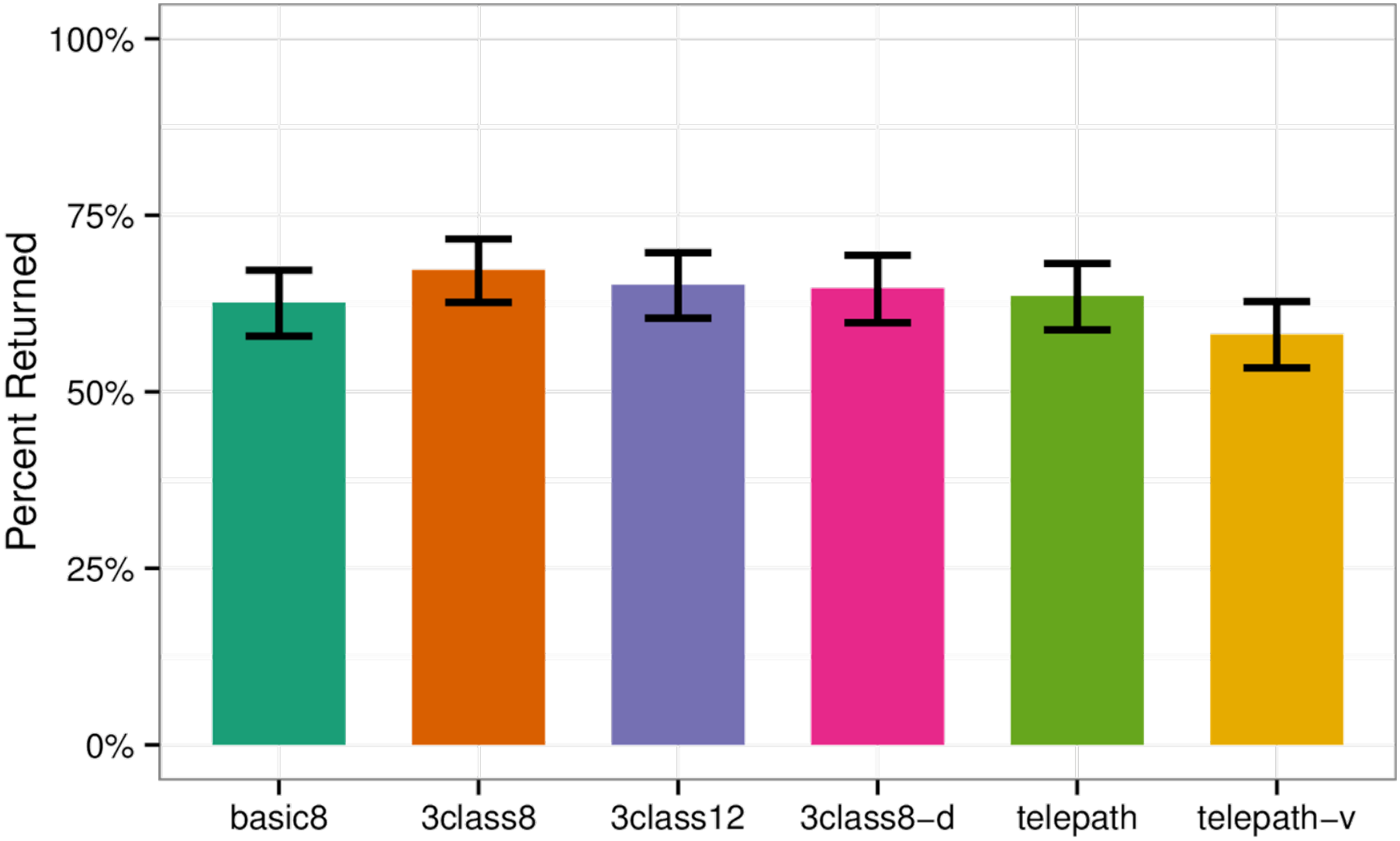
- Character-class requirements had little to no effect on security using our metrics

Carnegie Mellon University
CyLab

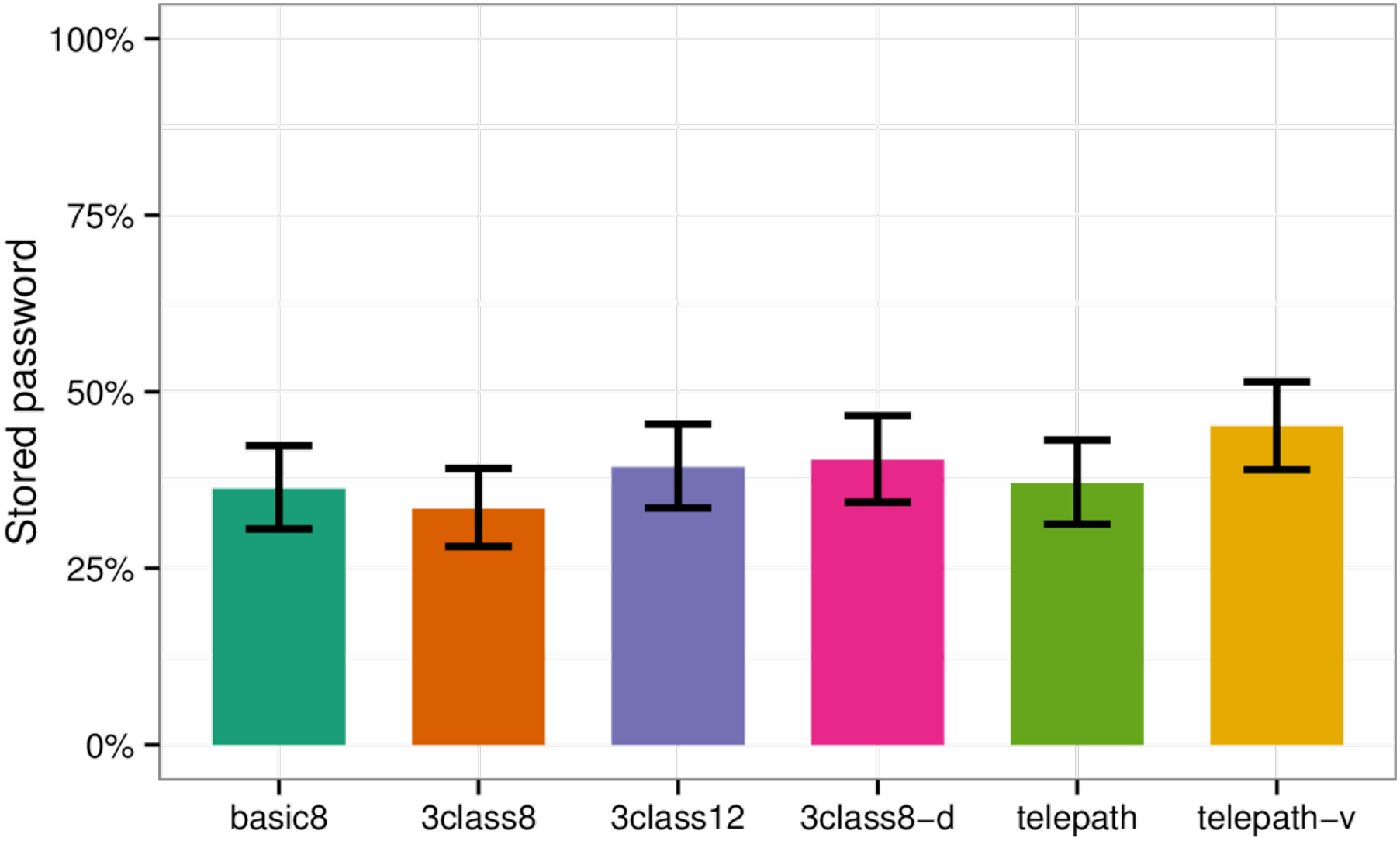
institute for SOFTWARE RESEARCH

Microsoft Research

Returned after 2-5 days

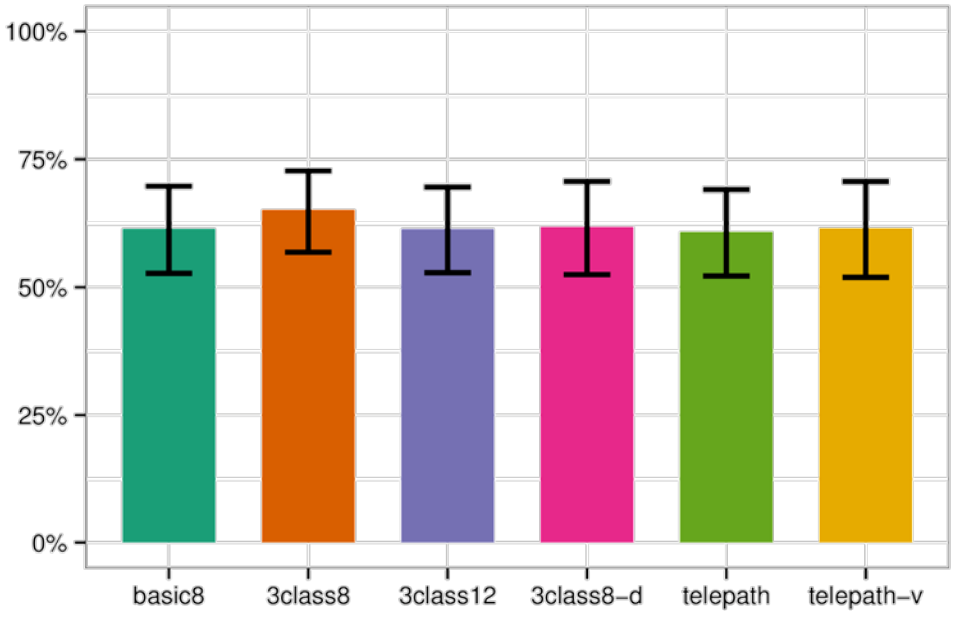


Usability: Stored password

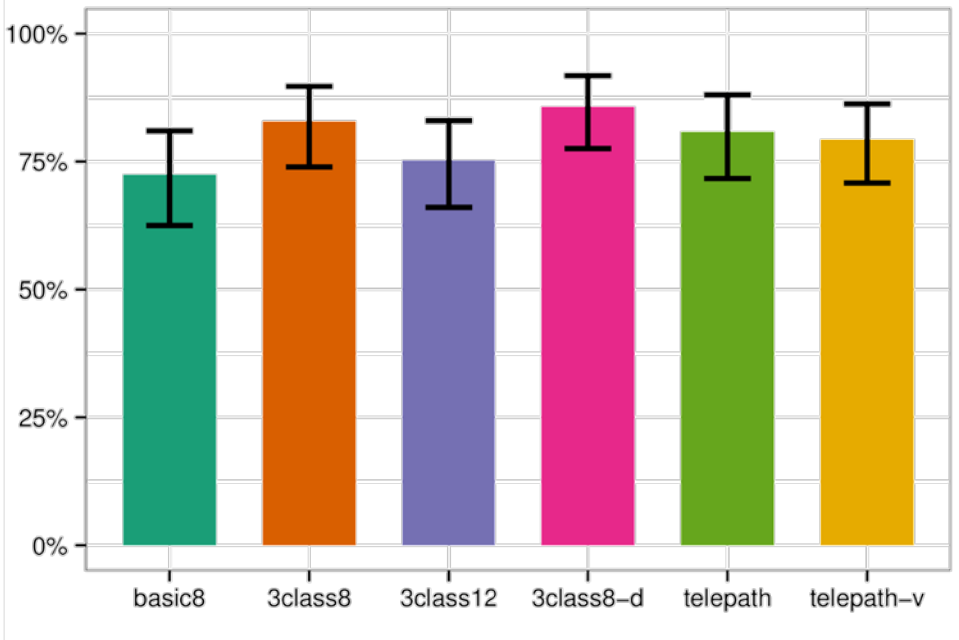


Usability: Recall after 2—5 days

Did not store password



Stored password



Usability: Toggled Show Password