RUHR-UNIVERSITÄT BOCHUM

"Someone Definitely Used 0000": Strategies, Performance, and User Perception of Novice Smartphone-Unlock PIN-Guessers

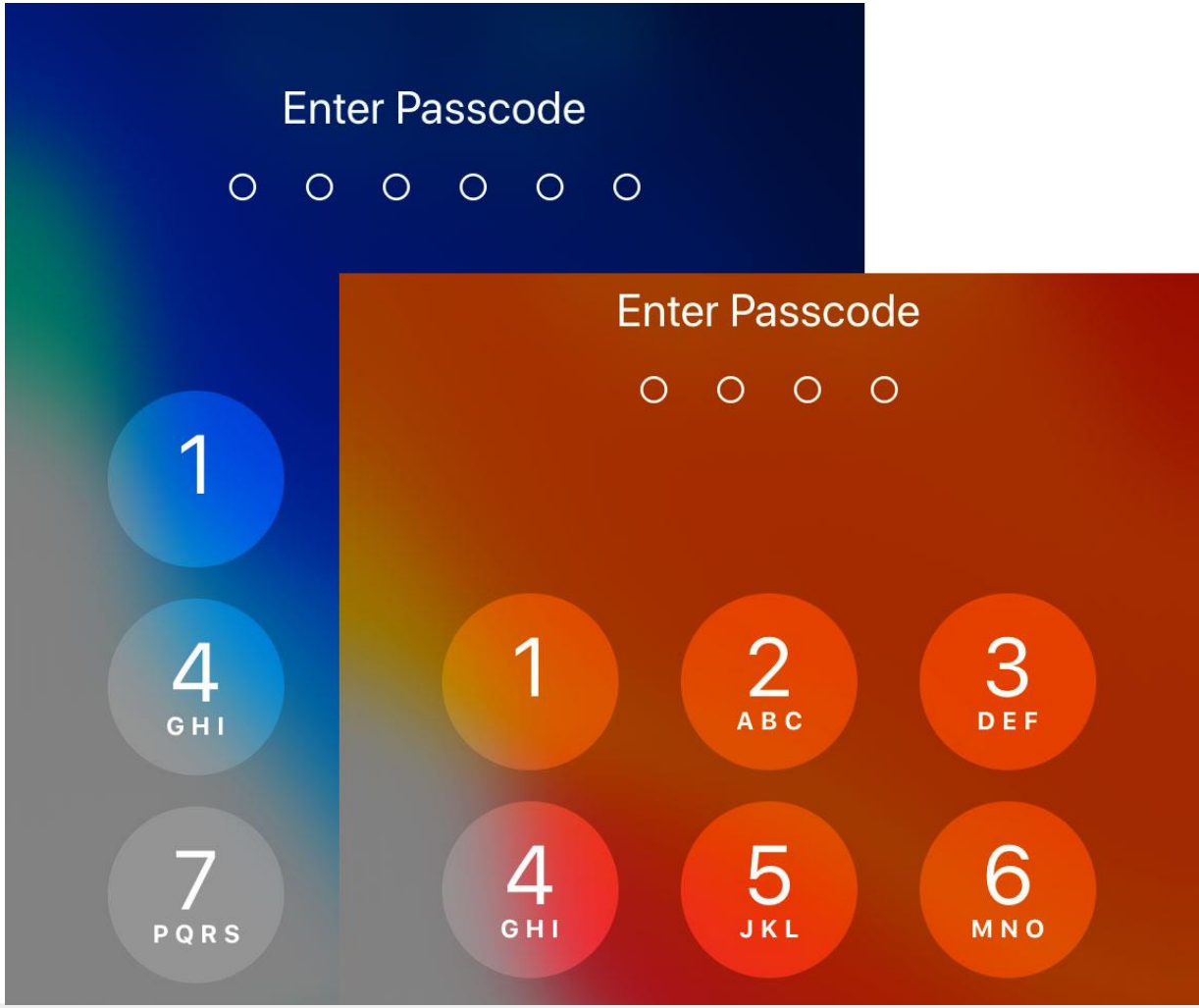
Daniel V. Bailey, Collins W. Munyendo, Hunter A. Dyer, Miles Grant, Philipp Markert, Adam J. Aviv

EuroUSEC, Copenhagen, Denmark, October 16, 2023

User Authentication

Decades of research on authentication and still we struggle with the same issues around **security** and **usability**. [1]

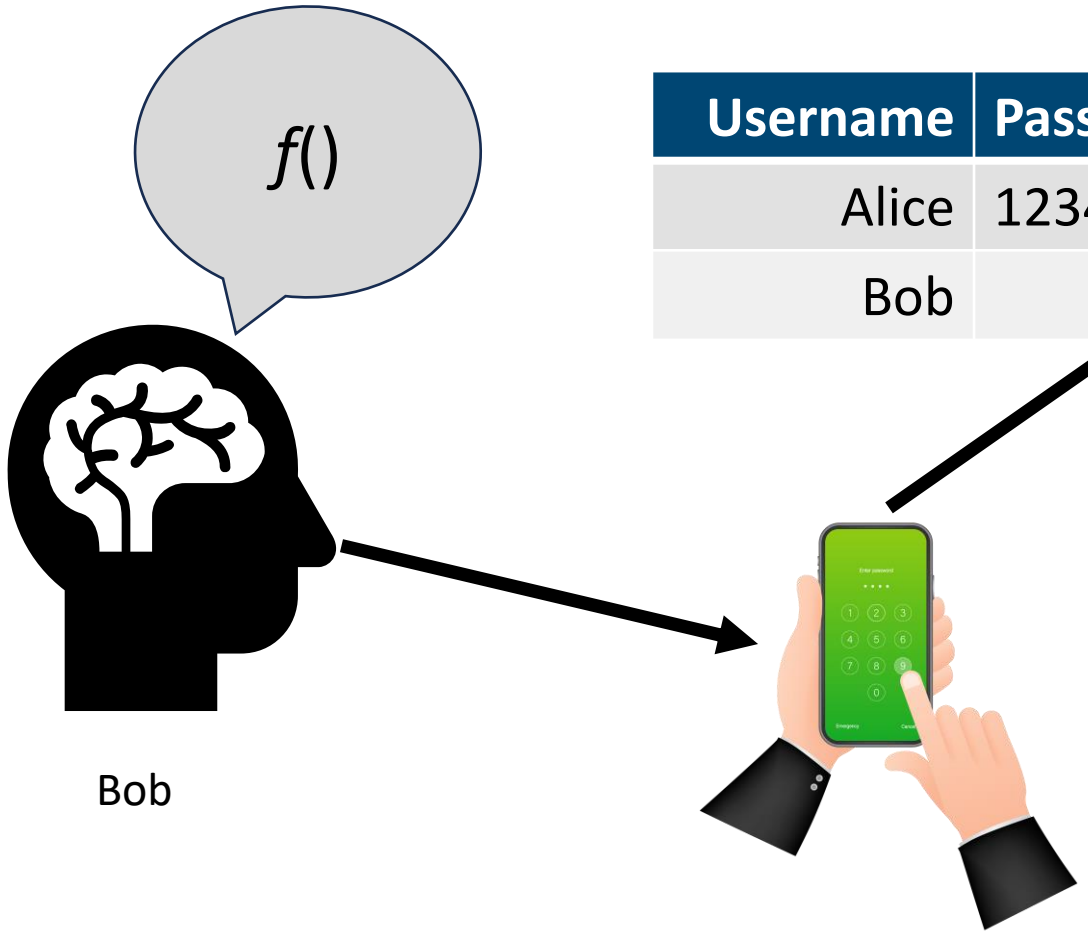
Knowledge-Based Authentication (KBA)



[Ref. 1] Herley et al.: **Passwords: If We're So Smart, Why are We Still Using Them?** (FC 2009)

RUB

Basic Setup: User-Chosen KBA Research



Username	Password
Alice	123456
Bob	

We don’t know *f*()! What can we do?

Define setting for *f*()

↓

Find new dataset like *f*()

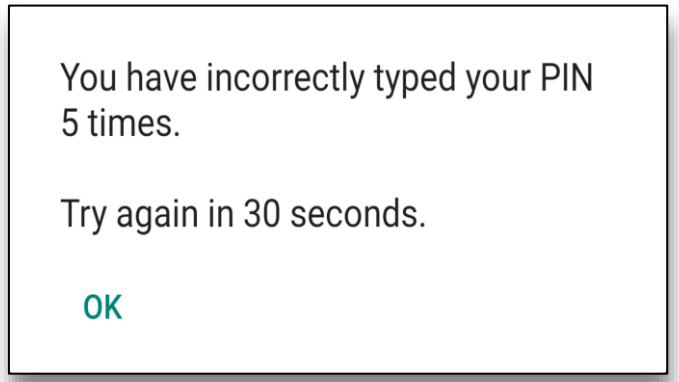
↓

Make new guessing heuristic

↓

Compare against previous heuristics

Focus on Throttled Attacker



No side channels!

Mobile guesser has a
- limited number of attempts

Guesses	Android	iOS
10	30sec	1h 36m
100	10h 45min	-

Attacker is *trawling*
- No personal info about target
- Attacker happy to unlock anyone's device

Example: Phones sold at US police auctions [2]

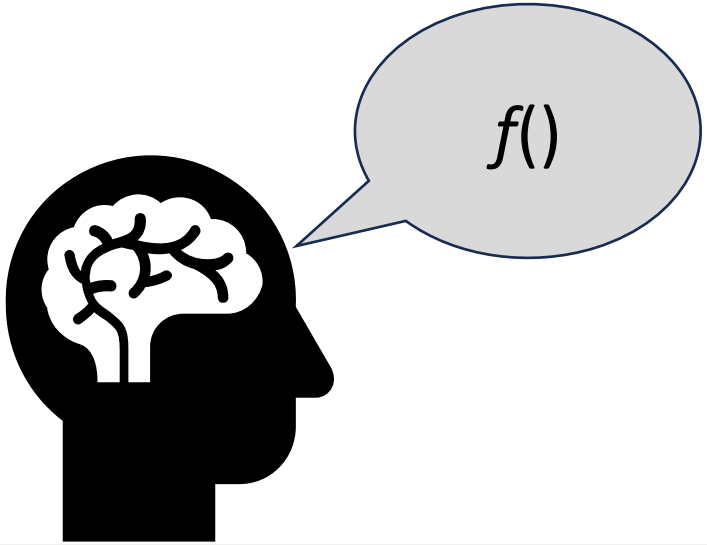
[Ref. 2] Roberts et al.: "**Blue Is the New Black (Market): Privacy Leaks and Re-Victimization from Police-Auctioned Cellphones**" (IEEE S&P '23)

KBA Research Basic Methodology

Previous proxy: Digits from password leaks [3]

- 6-digit PINs from alphanumeric passwords
- 4-digit PINs from Amitay's
  - "Big Brother Camera Security" iOS app

Count	Password
1 044 164	123456
176 120	password
88 076	12345678
78 720	111111



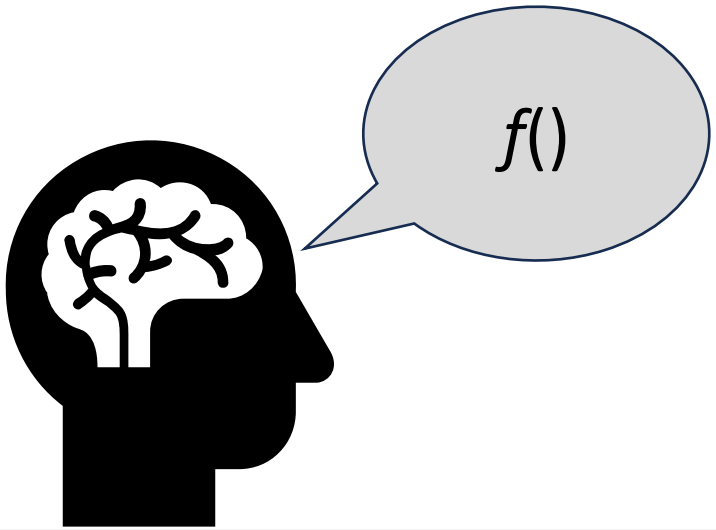
[Ref. 3] Markert et al.: "**On the Security of Smartphone Unlock** PINs" (ACM TOPS '21)

What About “Real” Attackers?

Do the previous datasets model a real-world threat?

PIN-guessing is something of a “target of opportunity”

A lost/stolen/unattended phone



Define setting for *f*()

↓

Find new dataset like *f*()

↓

Make new guessing heuristic

↓

Contribution: New “proxy” dataset from users playing the attacker

↓

Compare against previous heuristics

Novice Guesser Research Questions

RQ1: How do novices guess?

RQ2: How do they compare to prior datasets?

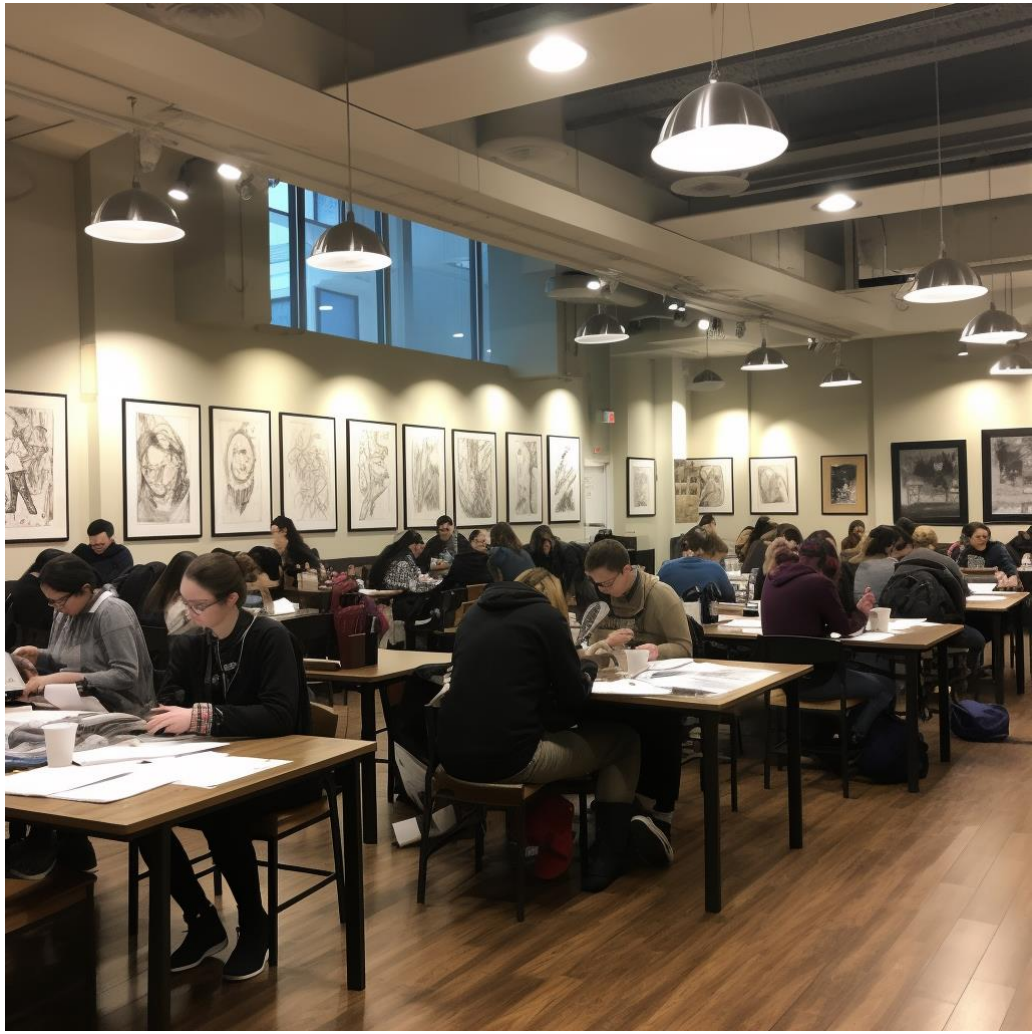
RQ3: What scenarios are they concerned about?

User Study Overview

Inspired by Uellenbeck, et al.[4]

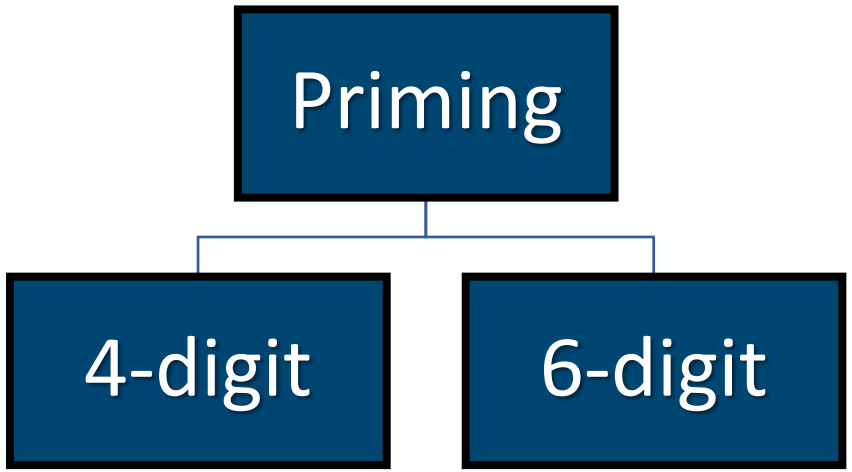
Our version:

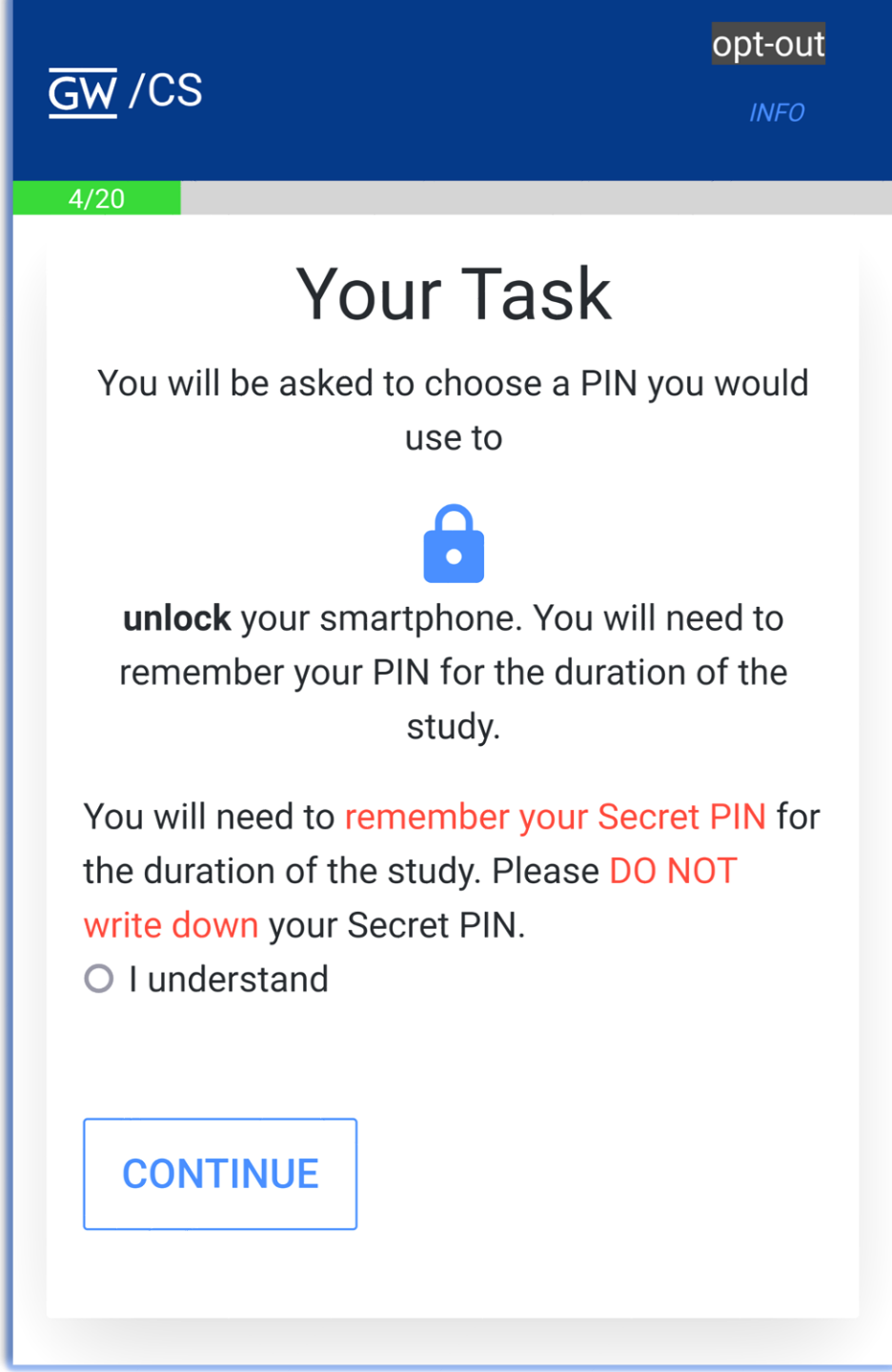
- Online (*n* = 210)
- Pick a “secret” PIN
- 5 guesses → Get a cash bonus for success



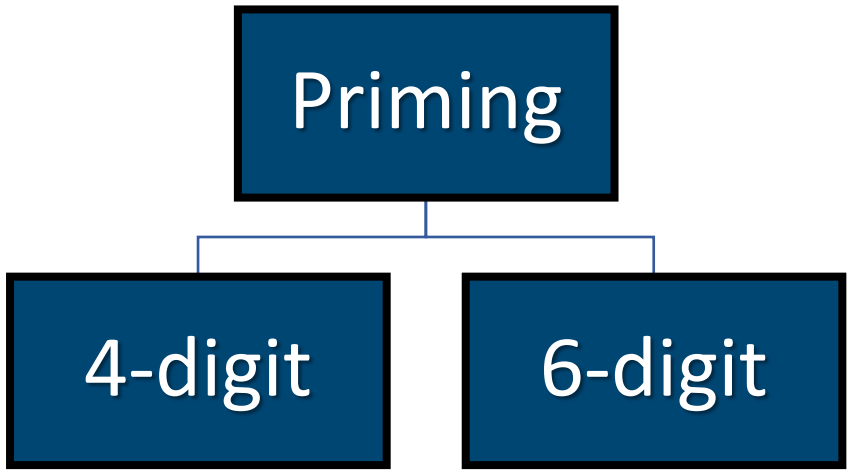
[Ref. 4] Uellenbeck, et al.: **“Quantifying the Security of Graphical Passwords: the Case of Android Unlock Patterns”**: (ACM CCS ’13)

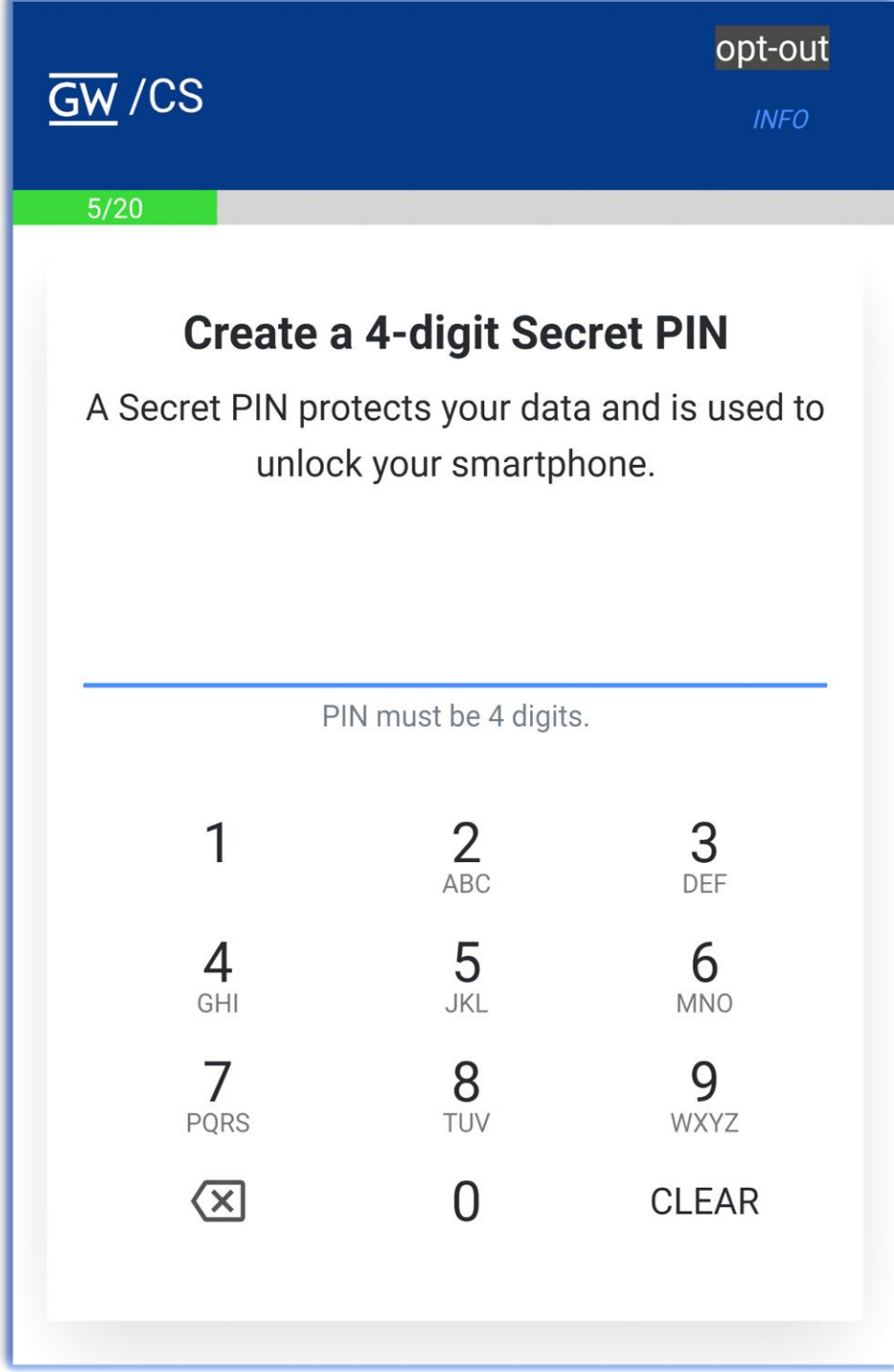
Methodology: User Study (*n* = 210)



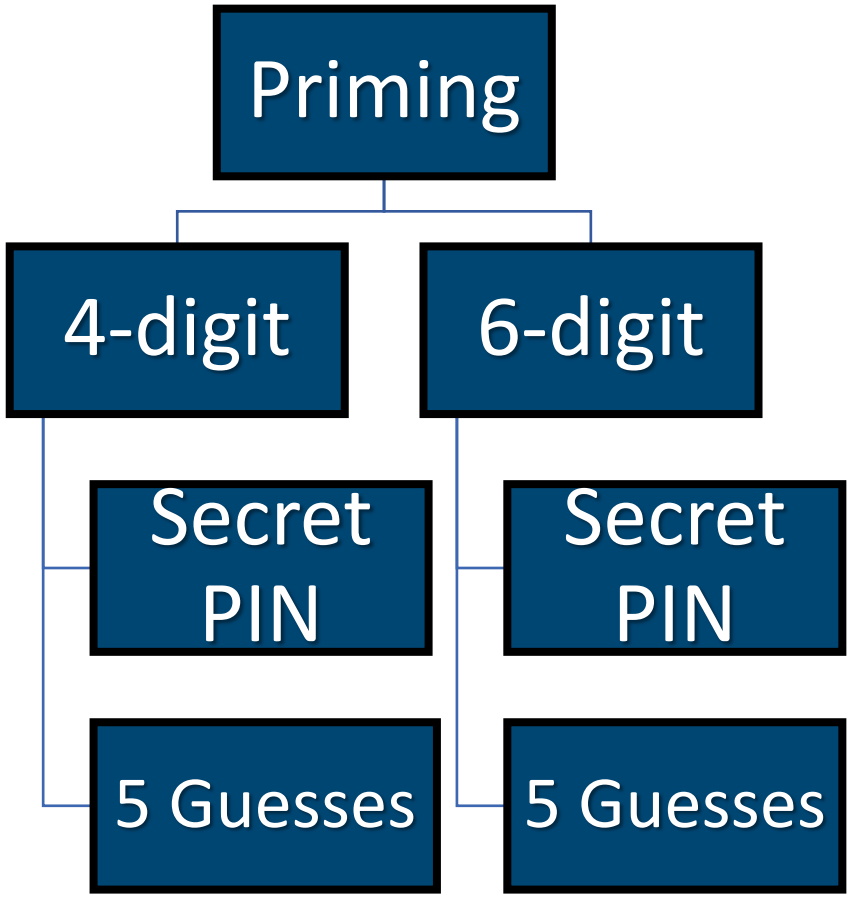


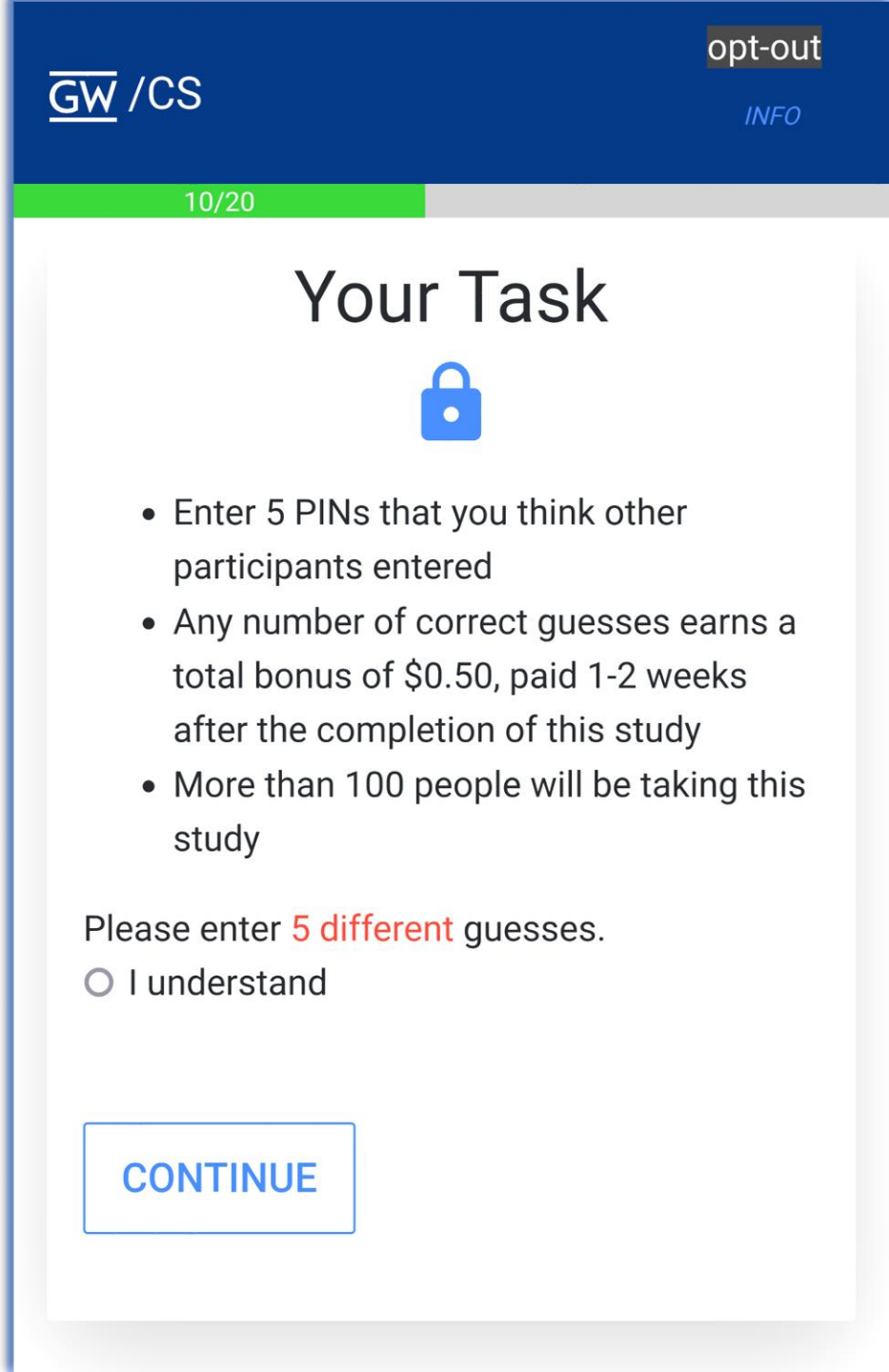
Methodology: User Study (*n* = 210)





Methodology: User Study (*n* = 210)





RQ1: How do Novices Guess?

Guessing risk concentrated in a small handful of PINs, like **0000**

- Only one-third thought their PIN would be guessed
- PINs in **bold** were guessed by 20+ attackers

Other popular guesses that were incorrect:

1111, 000000, 111111, 987654

85% of participants guessed successfully

10% of secret PINs were guessed

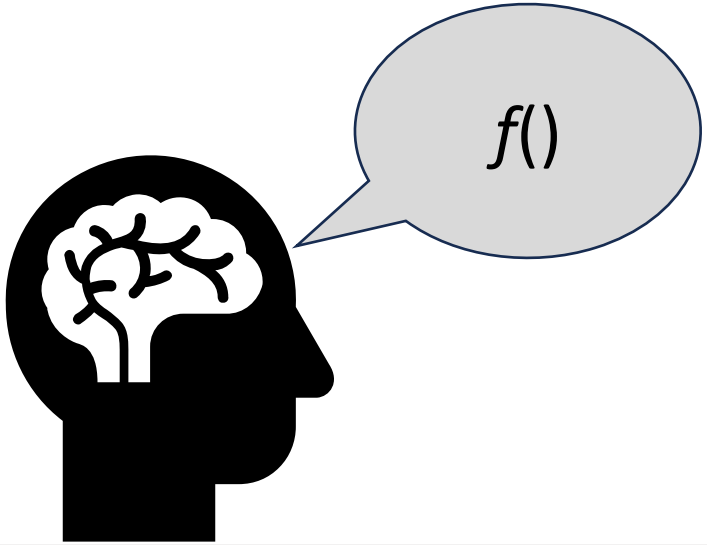
4d Guessed	6d Guessed
0000	121212
1234	**123456**
1478	134679
1990	135790
1995	159753
1997	654321
2000	
2468	
2580	
6666	

RQ2: Comparison Against Prior Datasets

What if we built an aggregate or data-driven guesser from our new list?

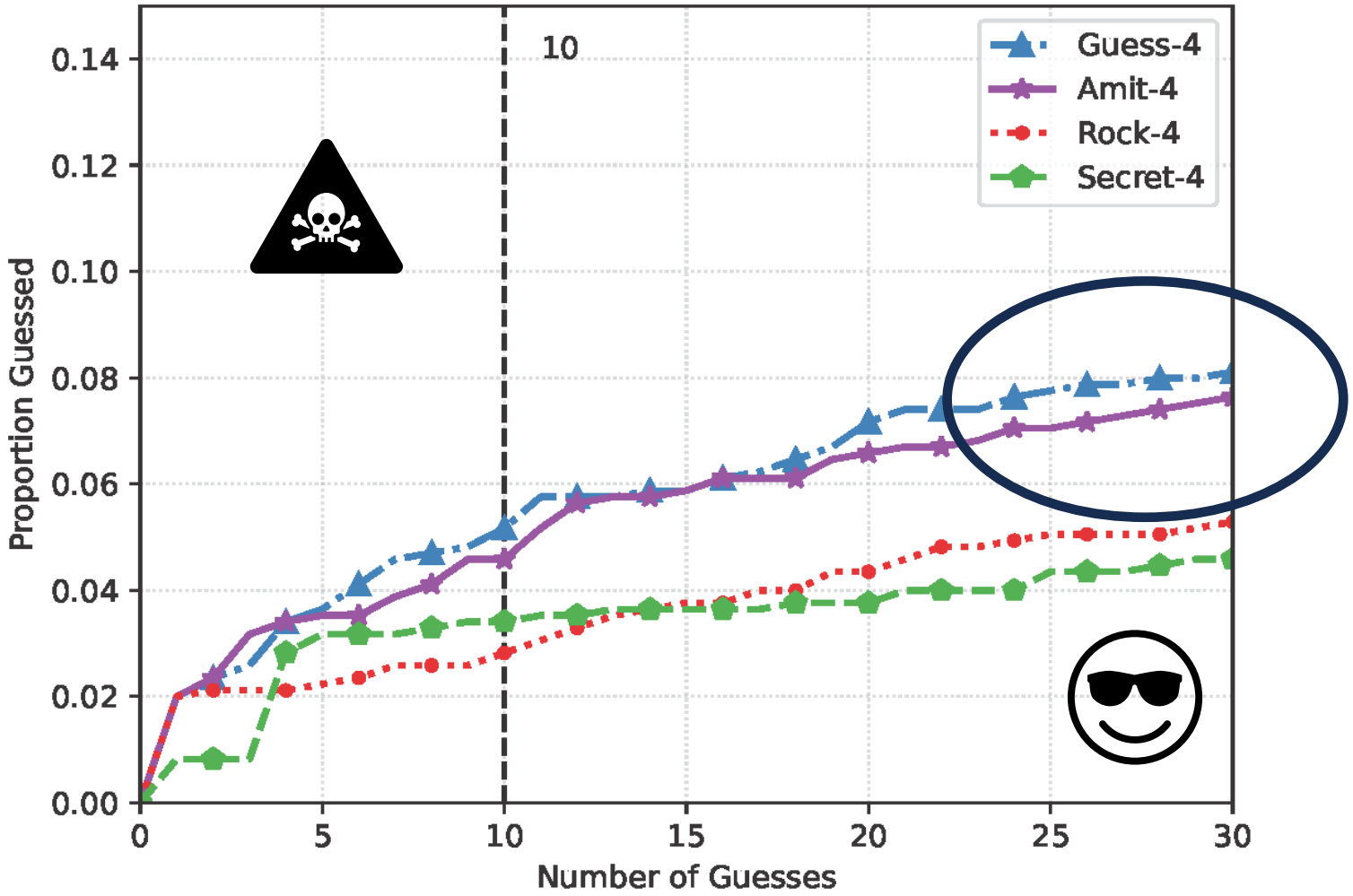
How would it compare?

Count	Password
1 044 164	123456
176 120	password
88 076	12345678
78 720	111111

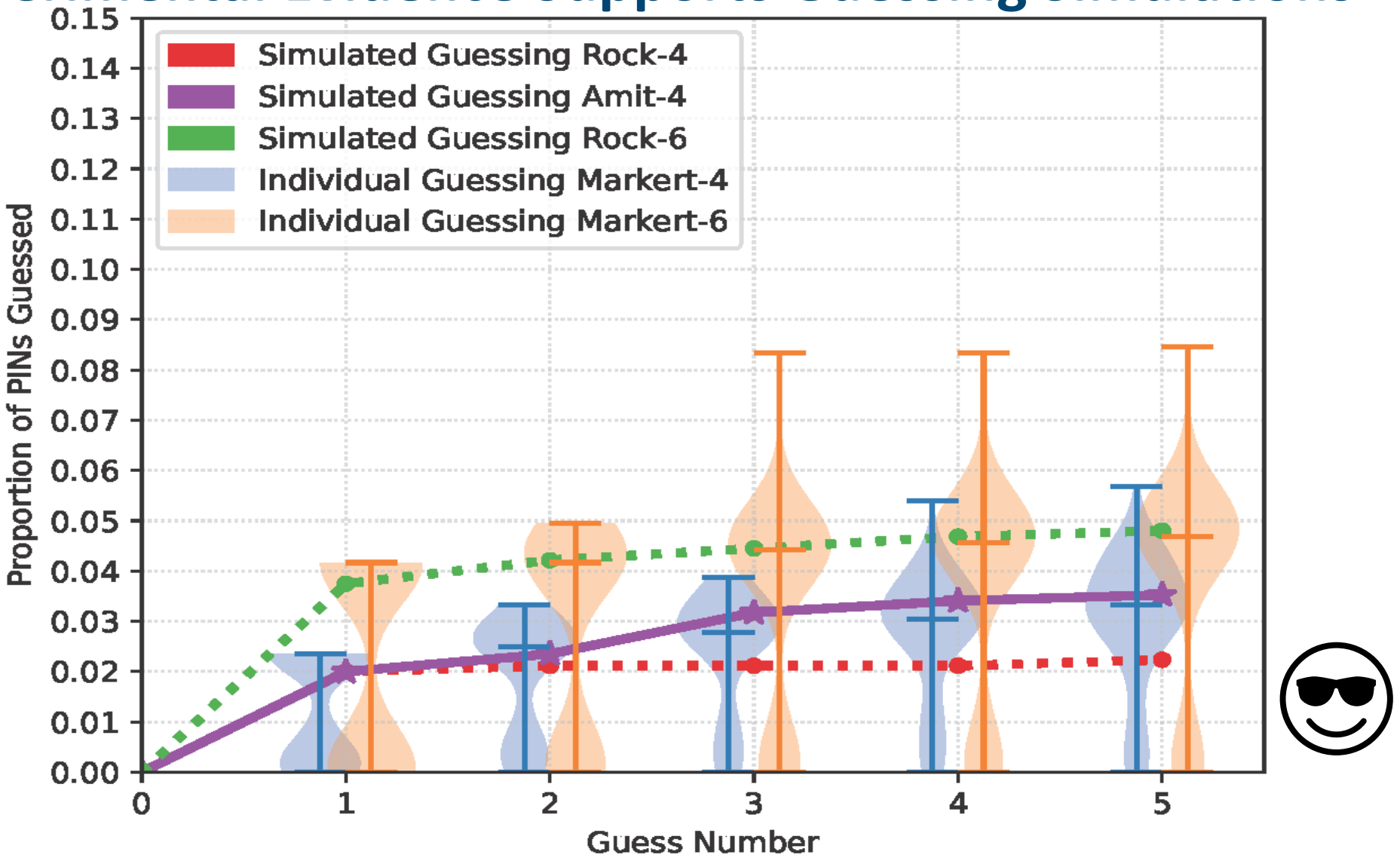


RQ2: How do They Compare with Prior Datasets?

4-digit/30 guesses: 8.1% observed vs. 7.6% simulated



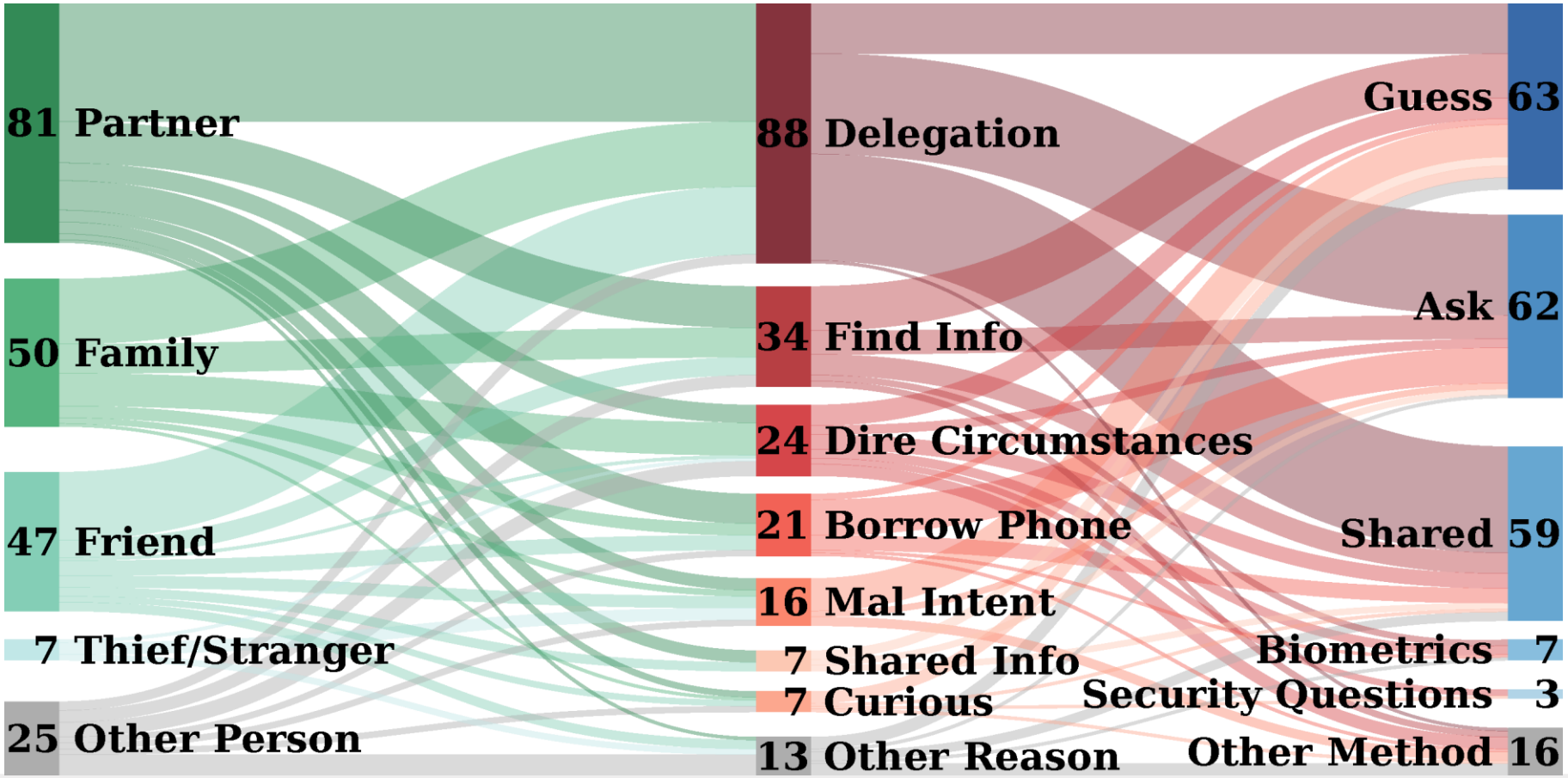
RQ2: Experimental Evidence Supports Guessing Simulations



RQ3: What Scenarios are Participants Concerned About?

Participants mostly think about close social contacts

Future work: guessing by insiders!!



Novice Guesser Research Questions: Results

RQ1: Performance of novice guessers

- 10% (21) participants' PINs were guessed
- 4-digit: 13%
- 6-digit: 7%
- …at 30 guesses

RQ2: Prior dataset comparison

- New dataset
- Comparable to prior sets

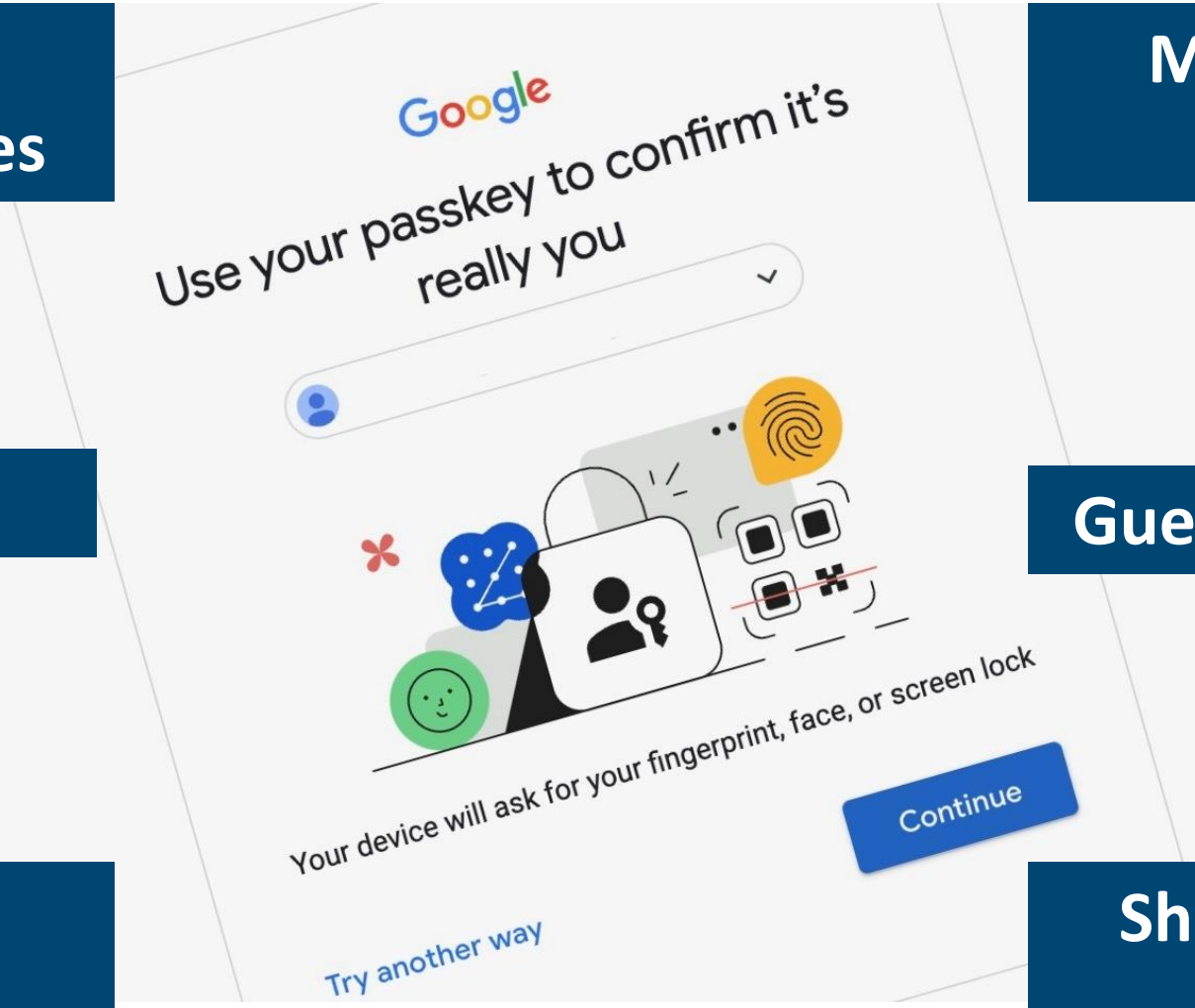
RQ3: Areas of concern

- Close social connections
- 37% attempted access
- 45% changed their PIN *to keep someone out*

Outlook/Future Work

Better messaging/nudges

More local KBA validation



Side channels

Guessing by insiders

Special populations

Sharing/revoking access

Takeaways

Define setting for *f*()

↓

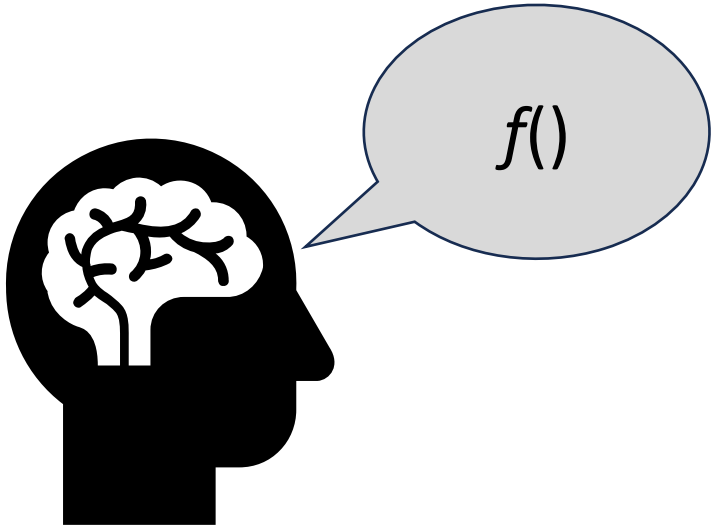
Find new dataset like *f*()

↓

Make new guessing heuristic

↓

Compare against previous heuristics



Throttling protects all but ~10% of PINs

4- and 6-digit PINs about the same

Unauthorized access is commonplace

Email: daniel.bailey@rub.de
Twitter: @pinresearcher