Security for All

Modeling Structural Inequities to Design More Secure Systems

Elissa M. Redmiles

@eredmil1

eredmiles@cs.umd.edu

University of Zurich UZH

NSF Graduate Research Fellowship Program

HCIL

University of Maryland

MARYLAND CYBERSECURITY CENTER

facebook

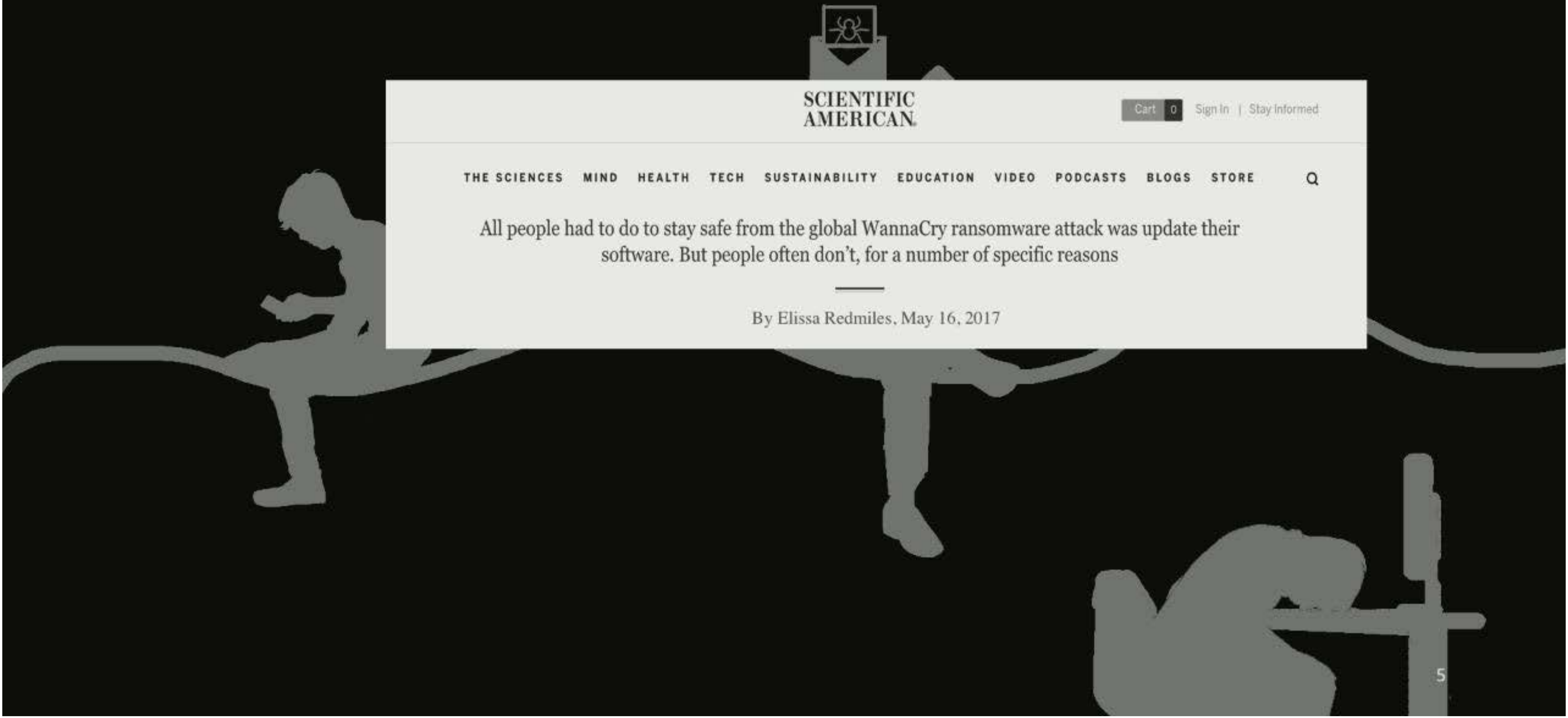
MAX PLANCK INSTITUTE FOR SOFTWARE SYSTEMS

People must make a variety of security decisions

People are not always good at making security decisions

Despite advances on core security problems, user decisions can still lead to significant security risks

Despite advances on core security problems, user decisions can still lead to significant security risks



SCIENTIFIC AMERICAN

Cart 0 Sign In | Stay Informed

THE SCIENCES MIND HEALTH TECH SUSTAINABILITY EDUCATION VIDEO PODCASTS BLOGS STORE

All people had to do to stay safe from the global WannaCry ransomware attack was update their software. But people often don't, for a number of specific reasons

By Elissa Redmiles, May 16, 2017

Despite advances on core security problems, user decisions can still lead to significant security risks

SCIENTIFIC AMERICAN

Cart 0 Sign In | Stay Informed

THE SCIENCES MIND HEALTH TECH SUSTAINABILITY EDUCATION VIDEO PODCASTS BLOGS STORE

All people had to do to stay safe from the global WannaCry ransomware attack was update their software. But people often don't, for a number of specific reasons

By Elissa Redmiles, May 16, 2017

COMMUNICATIONS OF THE ACM

HOME | CURRENT ISSUE | NEWS | BLOGS | OPINION | RESEARCH | PRACTICE | CAREERS | ARCHIVE | VIDEOS

The State of Phishing Attacks

By Jason Hong

Estimates of damage caused by phishing vary widely, ranging from $61 million per year to $3 billion per year of direct losses to victims in the U.S.

Goal: keep people secure

Goal: keep people secure

Change the people

Change the systems

Goal: keep people secure

Scientifically understand people's security behavior

Change the people

Change the systems

Goal: keep people secure

Scientifically understand people's security behavior

MANUAL

Change the people

Change the systems

My focus: behavioral security

Economic
Behavioral Econ

Security Measurement
Large-scale Log Analysis

Social Scientific
Surveys & Interviews

My focus: behavioral security

Economic
Behavioral Econ

Security Measurement
Large-scale Log Analysis

Social Scientific
Surveys & Interviews

My focus: behavioral security

Economic
Behavioral Econ

Security Measurement
Large-scale Log Analysis

Social Scientific
Surveys & Interviews

Scientifically understand insecure behavior

My prior work investigates security problems using behavioral models to facilitate secure system design

Account & Device Security

[S&P16]
[EC18]
[CCS16]
[CCS18a]
[CCS18b]
[S&P19a]
[S&P19b]
[CHI17]
[TWEB18]
[WAY17]

Spam & Fake News

[CHI18]
[FAT*19a]
[FAT*19b]

Enterprise Security

[S&P18]
[BigData16]
[USENIXSec18]
Distinguished Paper

Encryption & Data Use

[USENIX Sec17]
[SOUPS18]
[ICWSM18]
[ICWSM19]
[FOCI18]

My prior work investigates security problems using behavioral models to facilitate secure system design

Account & Device Security		Spam & Fake News	Enterprise Security	Encryption & Data Use
[S&P16]	[S&P19a]	[CHI18]	[S&P18]	[USENIX Sec17]
[EC18]	[S&P19b]	[FAT*19a]	[BigData16]	[SOUPS18]
[CCS16]	[CHI17]	[FAT*19b]	[USENIXSec18] *Distinguished Paper*	[ICWSM18]
[CCS18a]	[TWEB18]			[ICWSM19]
[CCS18b]	[WAY17]			[FOCI18]

My prior work investigates security problems using behavioral models to facilitate secure system design

Account & Device Security

[S&P16]
[EC18]
[CCS16]
[CCS18a]
[CCS18b]
[S&P19a]
[S&P19b]
[CHI17]
[TWEB18]
[WAY17]

Spam & Fake News

[CHI18]
[FAT*19a]
[FAT*19b]

Enterprise Security

[S&P18]
[BigData16]
[USENIXSec18]
Distinguished Paper

Encryption & Data Use

[USENIX Sec17]
[SOUPS18]
[ICWSM18]
[ICWSM19]
[FOCI18]

My prior work investigates security problems using behavioral models to facilitate secure system design

Account & Device Security

[S&P16] [S&P19a]

[EC18] [S&P19b]

[CCS16] [CHI17]

[CCS18a] [TWEB18]

[CCS18b] [WAY17]

Spam & Fake News

[CHI18]

[FAT*19a]

[FAT*19b]

Enterprise Security

[S&P18]

[BigData16]

[USENIXSec18]
Distinguished Paper

Encryption & Data Use

[USENIX Sec17]

[SOUPS18]

[ICWSM18]

[ICWSM19]

[FOCI18]

My prior work investigates security problems using behavioral models to facilitate secure system design

Account & Device Security

[S&P16]
[EC18]
[CCS16]
[CCS18a]
[CCS18b]
[S&P19a]
[S&P19b]
[CHI17]
[TWEB18]
[WAY17]

Spam & Fake News

[CHI18]
[FAT*19a]
[FAT*19b]

Enterprise Security

[S&P18]
[BigData16]
[USENIXSec18]
Distinguished Paper

Encryption & Data Use

[USENIX Sec17]
[SOUPS18]
[ICWSM18]
[ICWSM19]
[FOCI18]

Goal: keep people secure

Scientifically understand people's security behavior

Is security behavior suitable for scientific study?

The user is going to pick
dancing pigs over **security** every time.

-- McGraw and Felten / Schneier

Today's Agenda: finding a model of best fit for security behavior & balancing structural inequities in security

Model of best fit for security behavior

Balancing structural inequities in real systems

Epistemology of methods

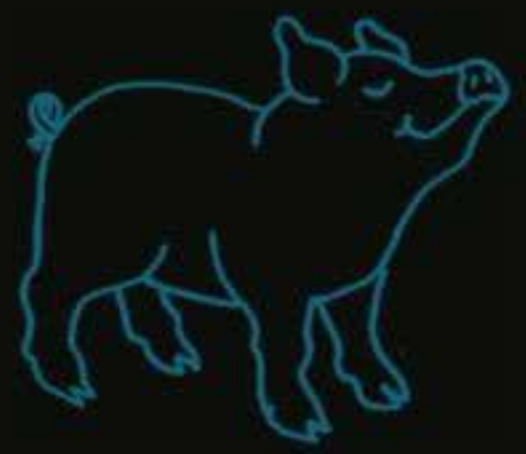
Today's Agenda: finding a model of best fit for security behavior & balancing structural inequities in security

Model of best fit for security behavior

Balancing structural inequities in real systems

Epistemology of methods

Potential model for security behavior: rational choice



The user is going to pick
dancing pigs over **security** every time.

-- McGraw and Felten / Schneier

The user is rationally ignoring security advice
because **the costs outweigh the benefits.**
-- Herley, 2009

To test the rationality hypothesis, we need controlled experiments to observe tradeoffs between cost & risk

Experimentation

Security Measurement

Survey Methodology

Designed a novel, scalable behavioral-economics experimentation system for security behavior

Online experimental system: simple bank account

Account holds study compensation

- Account has explicit **risk** of being hacked

Redmiles, E.M., Mazurek, M.L., and Dickerson, J.P. *Dancing Pigs or Externalities? Measuring the Rationality of Security Decisions.* ACM **Economics & Computation (EC)** 2018.

Featured on

Schneier on Security

Designed a novel, scalable behavioral-economics experimentation system for security behavior

Online experimental system: simple bank account

Account holds study compensation

- Account has explicit **risk** of being hacked

Users make a security choice: enable/don't enable 2FA

2FA lowers **risk** of hacking

- Increases **cost** (time and effort) to complete study

Redmiles, E.M., Mazurek, M.L., and Dickerson, J.P. *Dancing Pigs or Externalities? Measuring the Rationality of Security Decisions.* ACM **Economics & Computation (EC)** 2018.

Featured on

Schneier on Security

Designed a novel, scalable behavioral-economics experimentation system for security behavior

Online experimental system: simple bank account
Account holds study compensation
- Account has explicit **risk** of being hacked

123456

Users make a security choice: enable/don't enable 2FA
2FA lowers **risk** of hacking
- Increases **cost** (time and effort) to complete study

$$$

Participants stand to lose money
Amazon Mechnical Turk (Crowd Worker) participants
- Earn money from small time increments

Redmiles, E.M., Mazurek, M.L., and Dickerson, J.P. *Dancing Pigs or Externalities? Measuring the Rationality of Security Decisions.* ACM **Economics & Computation (EC)** 2018.

Featured on
Schneier on Security

Participants interact with simulation system
We observe their responses to security prompts

Create account
bank.cs

UMD Website Study

Login

Bank

Study Details

Contact

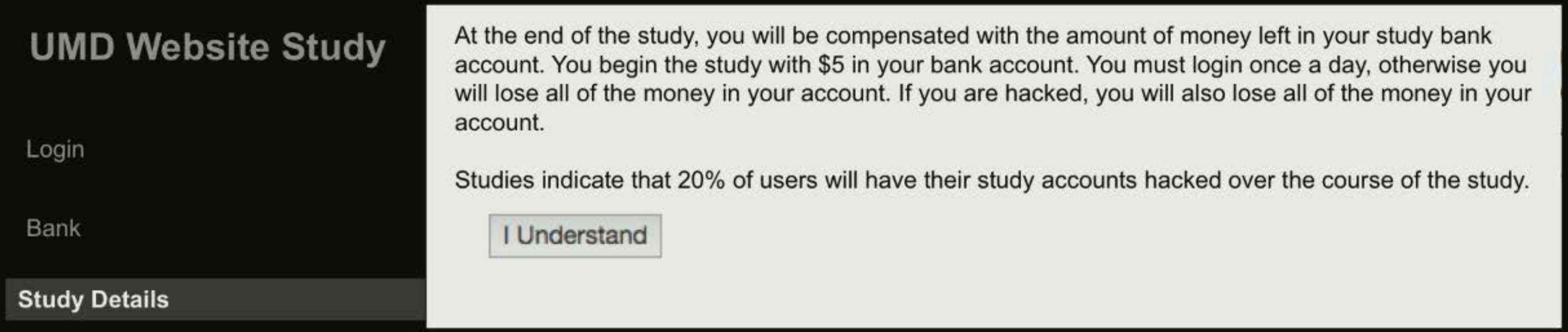
MTurk ID:

Password:

Confirm Password:

Submit

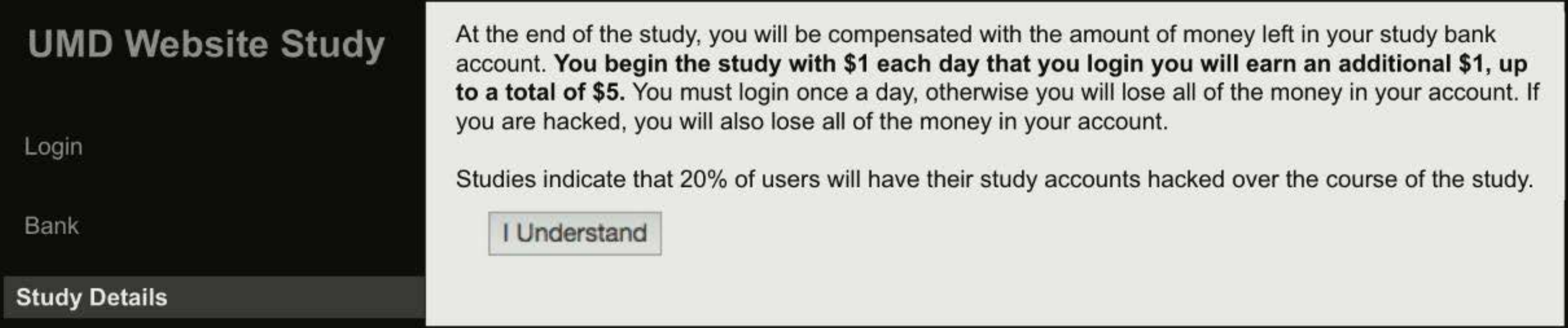
Participants interact with simulation system
We observe their responses to security prompts



Participants interact with simulation system
We observe their responses to security prompts

Create account on bank.cs

Learn risk of hacking (H)



Participants interact with simulation system
We observe their responses to security prompts

Create account on bank.cs

↓

Learn risk of hacking (H)

UMD Website Study

Login

Bank

Study Details

At the end of the study, you will be compensated with the amount of money left in your study bank account. **You begin the study with $1 each day that you login you will earn an additional $1, up to a total of $5.** You must login once a day, otherwise you will lose all of the money in your account. If you are hacked, you will also lose all of the money in your account.

Studies indicate that 20% of users will have their study accounts hacked over the course of the study.

I Understand

H = 1%, 20%, or 50%

Participants interact with simulation system
We observe their responses to security prompts

Create account on bank.cs

Learn risk of hacking (H)

Learn protection offered by 2FA (P)

Participants interact with simulation system
We observe their responses to security prompts

Create account on bank.cs

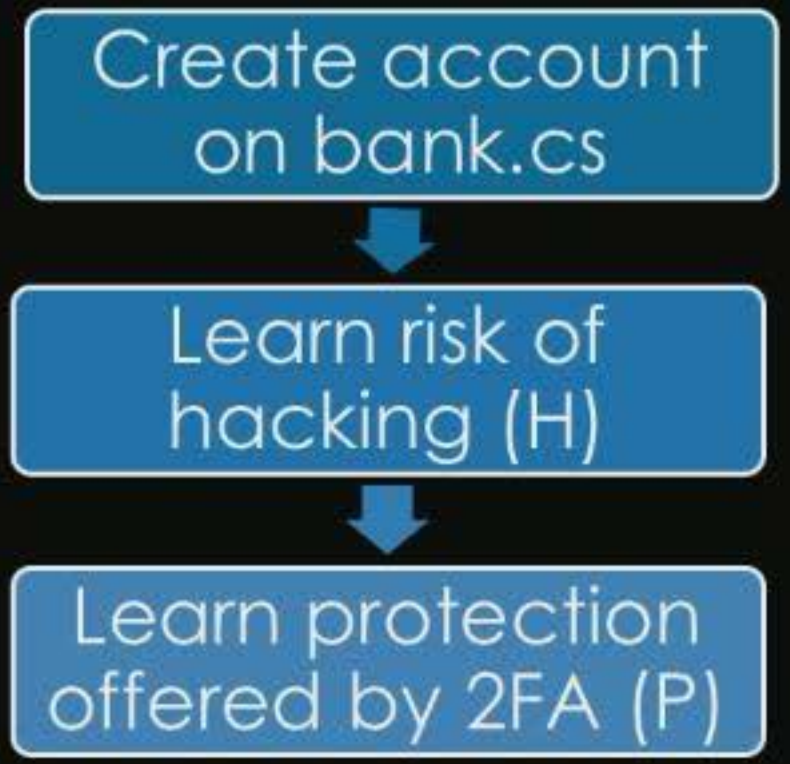
↓

Learn risk of hacking (H)

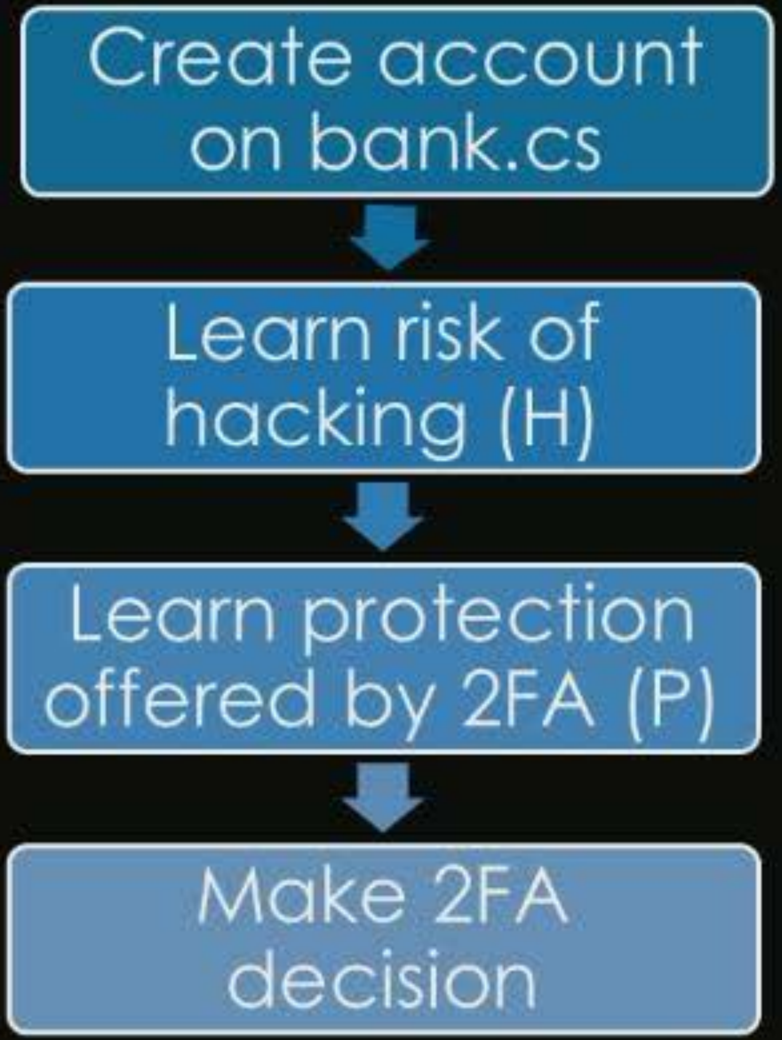
↓

Learn protection offered by 2FA (P)

Participants interact with simulation system
We observe their responses to security prompts



Participants interact with simulation system
We observe their responses to security prompts



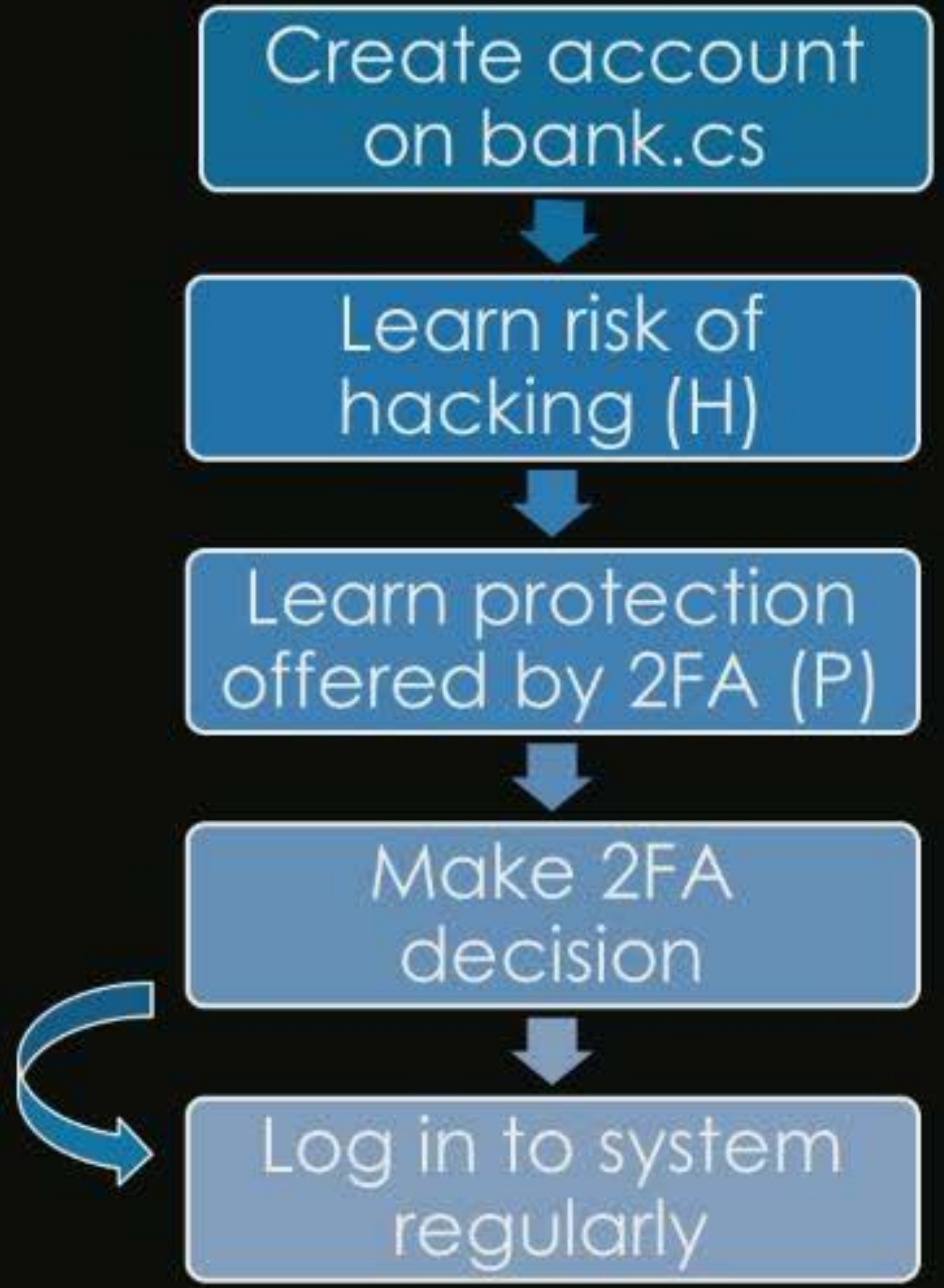
Participants interact with simulation system
We observe their responses to security prompts

UMD Website Study

You will lose all of your money if you do not login before January 19, 2018, 5:02pm EST.

Bank: $5

Participants interact with simulation system
We observe their responses to security prompts



UMD Website Study

You will lose all of your money if you do not login before January 19, 2018, 5:02pm EST.

Bank: $5

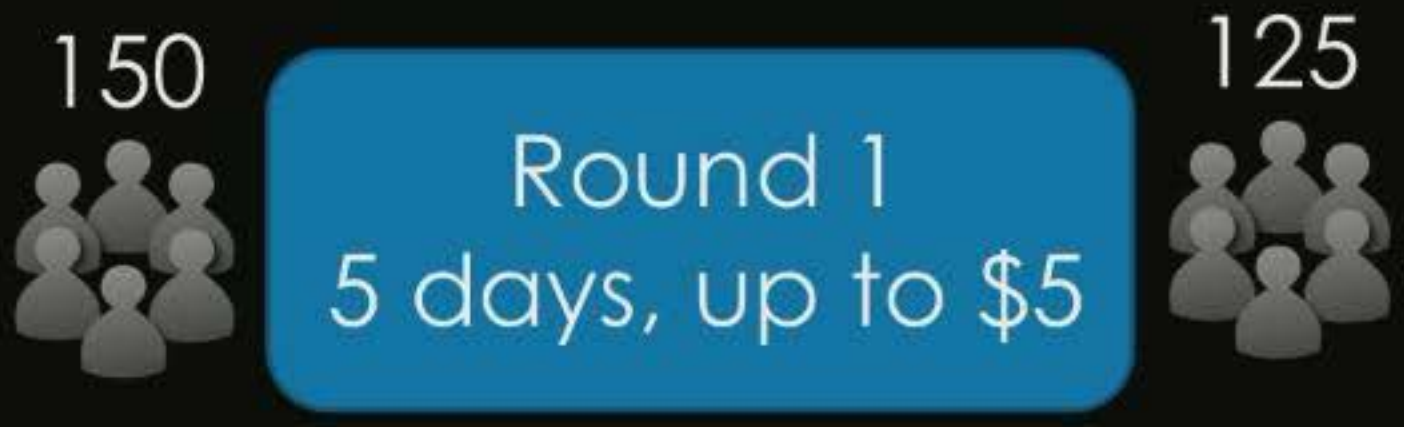
Observed 2FA behavior in two controlled experiments

Observed 2FA behavior in two controlled experiments

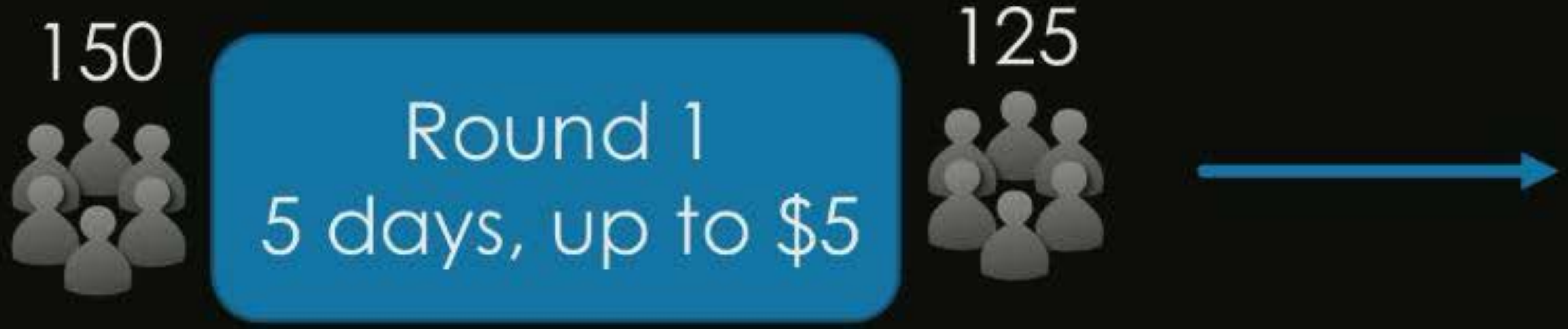
150

Round 1
5 days, up to $5

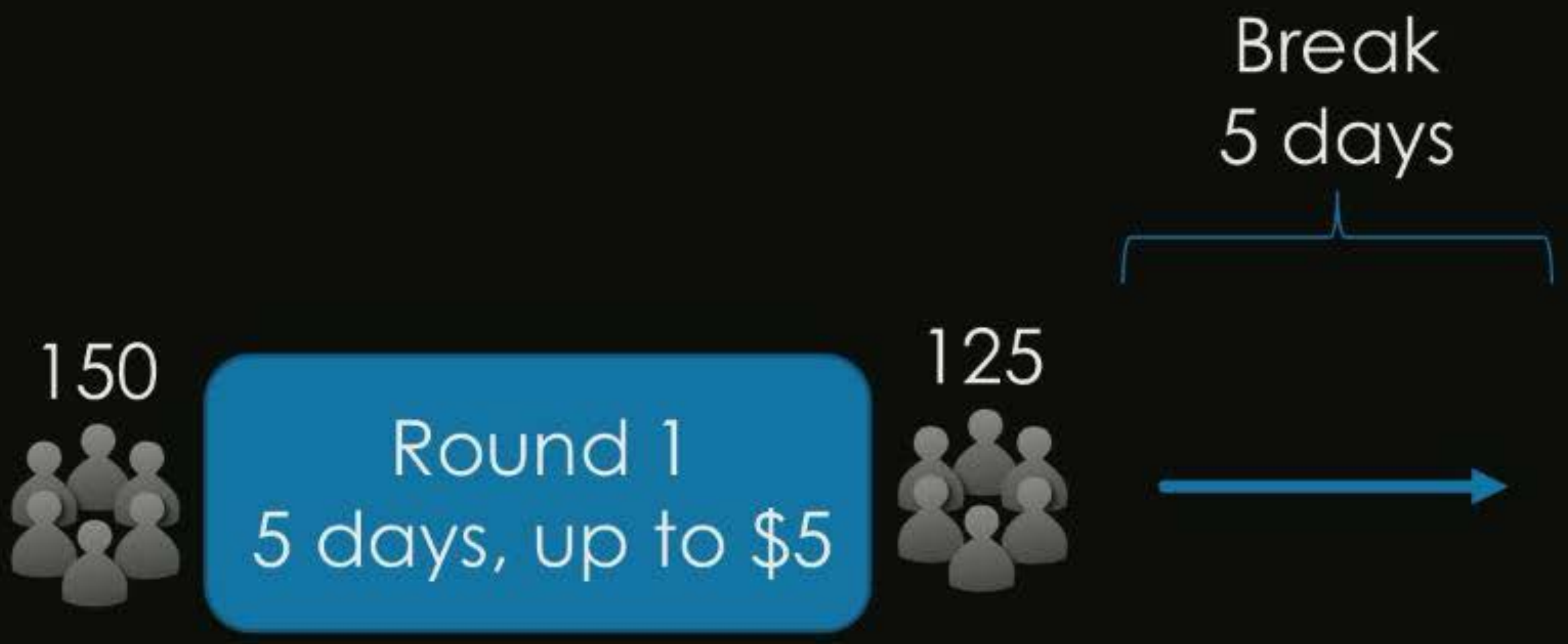
Observed 2FA behavior in two controlled experiments



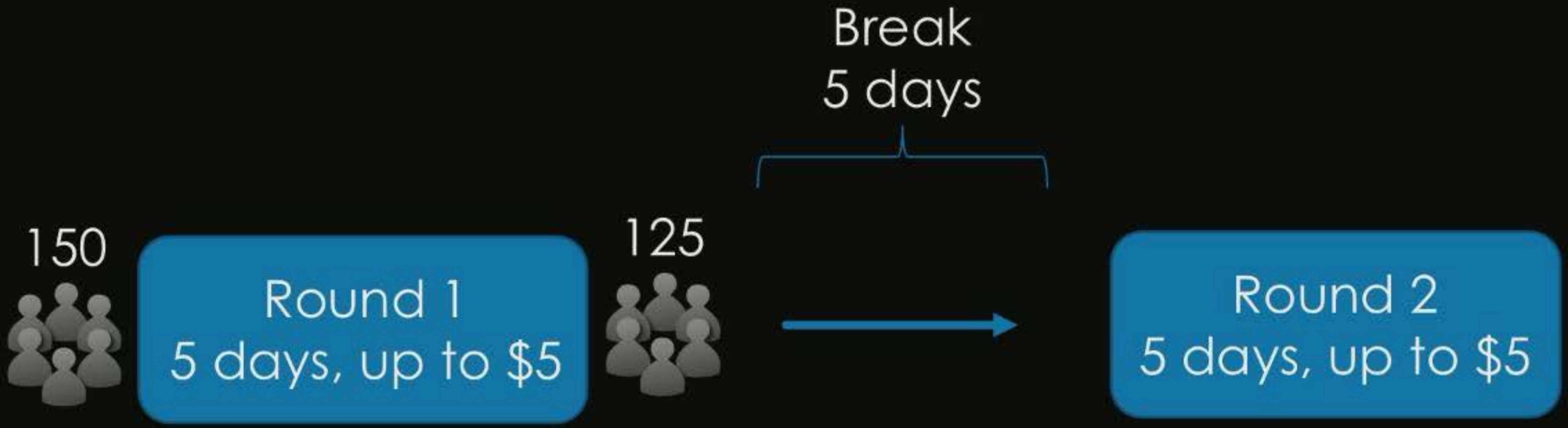
Observed 2FA behavior in two controlled experiments



Observed 2FA behavior in two controlled experiments



Observed 2FA behavior in two controlled experiments



Observed 2FA behavior in two controlled experiments

Break
5 days

150

Round 1
5 days, up to $5

125

Round 2
5 days, up to $5

107

Observed 2FA behavior in two controlled experiments

Break
5 days

150

Round 1
5 days, up to $5

125

Round 2
5 days, up to $5

107

H=1%, P=50%, Endow | Earn H=1%, P=90% Endow	H=20%, P=50%, Endow | Earn	H=50%, P=50%, Endow H=50%, P=90% Endow | Earn

Only 52% of participants enabled 2FA.

Testing the rationality hypothesis: were users rational in their 2FA choices?

Cost is defined as wage-earning time loss

$$C_{2FA} = \left(T_{signup} + \sum T_{login}\right) * wage_{MTurk}$$

Expected Value of 2FA is defined the $$$ savings if a hack occurred

$$EV_{2FA} = P[H * Max_{bank}]$$

Rational 2FA use: the expected value of the users' choice is greater than the cost

Testing the rationality hypothesis: were users rational in their 2FA choices?

Cost is defined as wage-earning time loss

$$C_{2FA} = \left(T_{signup} + \sum T_{login}\right) * wage_{MTurk}$$

Utility of 2FA is defined the $$$ savings if a hack occurred

$$EV_{2FA} = P[H * Max_{bank}]$$

Rational 2FA use: the expected value of the users' choice is greater than the cost

Testing the rationality hypothesis: were users rational in their 2FA choices?

Cost is defined as wage-earning time loss

$$C_{2FA} = \left(T_{signup} + \sum T_{login}\right) * wage_{MTurk}$$

Utility of 2FA is defined the $$$ savings if a hack occurred

$$EV_{2FA} = P[H * Max_{bank}]$$

Rational 2FA use: the expected value of the users' choice is greater than the cost

Example: Participant in H=20%, P=50% enables 2FA

Testing the rationality hypothesis: were users rational in their 2FA choices?

Cost is defined as wage-earning time loss

$$C_{2FA} = \left(T_{signup} + \sum T_{login}\right) * wage_{MTurk}$$

Utility of 2FA is defined the $$$ savings if a hack occurred

$$EV_{2FA} = P[H * Max_{bank}]$$

Rational 2FA use: the expected value of the users' choice is greater than the cost

Example: Participant in H=20%, P=50% enables 2FA

Cost

60 (s) for 2FA portion of signup + total of 180 (s) for 2FA sign-ins

240 (s) * 4.97$/hr = $0.33

Testing the rationality hypothesis: were users rational in their 2FA choices?

Cost is defined as wage-earning time loss

$$C_{2FA} = \left(T_{signup} + \sum T_{login}\right) * wage_{MTurk}$$

Utility of 2FA is defined the $$$ savings if a hack occurred

$$EV_{2FA} = P[H * Max_{bank}]$$

Rational 2FA use: the expected value of the users' choice is greater than the cost

Example: Participant in H=20%, P=50% enables 2FA

Cost

60 (s) for 2FA portion of signup + total of 180 (s) for 2FA sign-ins

240 (s) * 4.97$/hr = $0.33

Expected Value of 2FA

Participant's P = 50%, H = 20%, they can earn up to $5

0.5(0.2*5) = $0.50

Testing the rationality hypothesis: were users rational in their 2FA choices?

Cost is defined as wage-earning time loss

$$C_{2FA} = \left(T_{signup} + \sum T_{login}\right) * wage_{MTurk}$$

Utility of 2FA is defined the $$$ savings if a hack occurred

$$EV_{2FA} = P[H * Max_{bank}]$$

Rational 2FA use: the expected value of the users' choice is greater than the cost

Example: Participant in H=20%, P=50% enables 2FA

Cost

60 (s) for 2FA portion of signup + total of 180 (s) for 2FA sign-ins

240 (s) * 4.97$/hr = $0.33

Expected Value of 2FA

Participant's P = 50%, H = 20%, they can earn up to $5

0.5(0.2*5) = $0.50

$0.50 (expected value) > $0.33 (cost)

48% strictly rational with no experience (RD1)

61% strictly rational once familiar with the system (RD2)

Significant (p<0.001), medium (V=0.578) learning effect

Some users are more rational than others: those with more skill, more system experience, and at higher risk

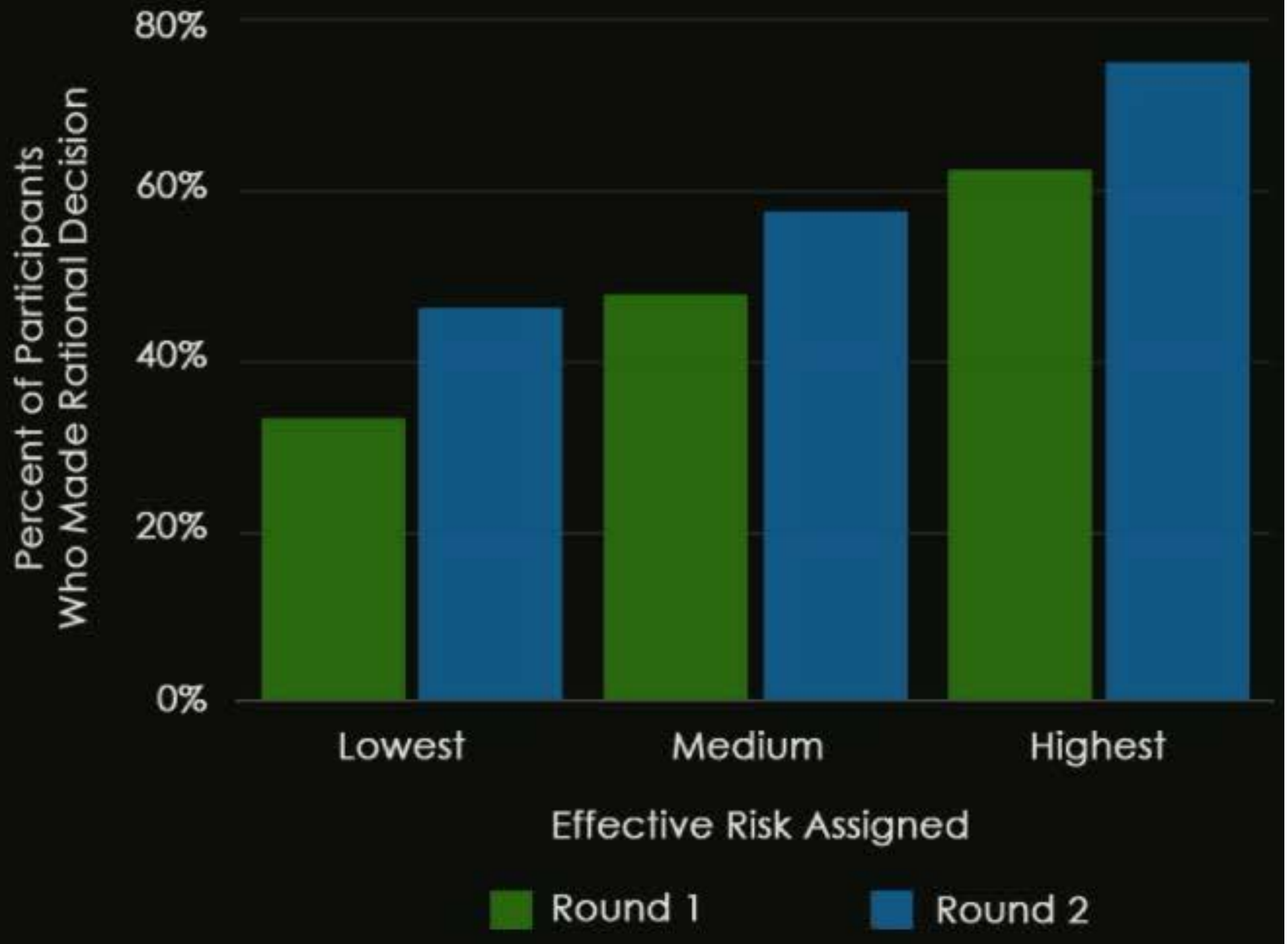
Higher internet skill 15% more likely to behave rationally

Higher security behavioral intent 3.9x more likely to behave rationally

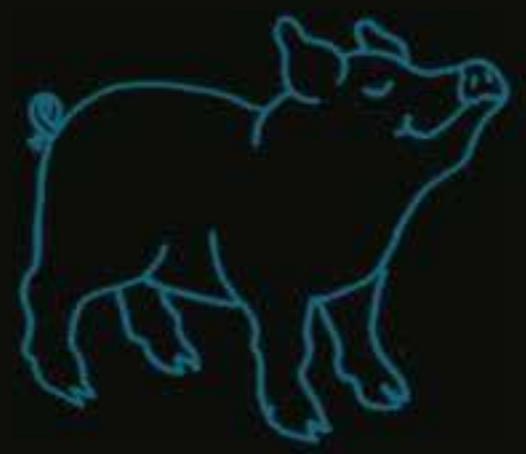
Some users are more rational than others: those with more skill, more system experience, and at higher risk

Higher internet skill 15% more likely to behave rationally

Higher security behavioral intent 3.9x more likely to behave rationally



How well does a bounded rationality model fit security behavior?



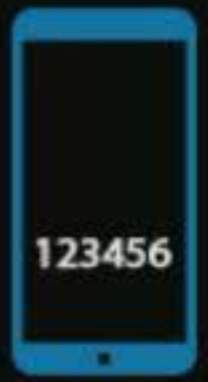
The user is going to pick
dancing pigs over **security** every time.

-- McGraw and Felten / Schneier

The user is rationally ignoring security advice
because **the costs outweigh the utility.**
-- Herley, 2009

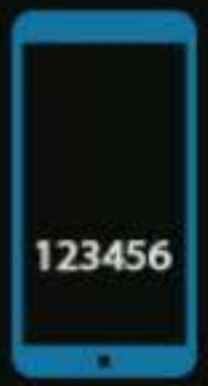
The user is a **boundedly rational security actor**
with predictable but not always utility-optimal behavior.

Testing the bounded rationality hypothesis: is there a consistent pattern in security behavior?



Enable 2FA

Testing the bounded rationality hypothesis: is there a consistent pattern in security behavior?



Enable 2FA

~

Account value

Testing the bounded rationality hypothesis: is there a consistent pattern in security behavior?

Testing the bounded rationality hypothesis: is there a consistent pattern in security behavior?

Enable 2FA

Account value

Risk
with/out 2FA

Controls
Password Strength
Internet & Security Skill
Demographics
(Gender, Age, Education)

Testing the bounded rationality hypothesis: is there a consistent pattern in security behavior?

Enable 2FA

Account value

Risk
with/out 2FA

Controls
Password Strength
Internet & Security Skill
Demographics
(Gender, Age, Education)

← Neural Net Strength Meter
Ur et al. 2017

Testing the bounded rationality hypothesis: is there a consistent pattern in security behavior?

Enable 2FA

Account value

Risk
with/out 2FA

Controls
Password Strength
Internet & Security Skill ← Validated Scales
Demographics
(Gender, Age, Education)

Hargittai & Hsieh 2013
Egelman & Peer 2015

Testing the bounded rationality hypothesis: is there a consistent pattern in security behavior?

Enable 2FA

Account value

Risk
with/out 2FA

Controls
Password Strength
Internet & Security Skill
Demographics
(Gender, Age, Education)

Account value and risk relate to behavior

Endowment: 2.3x more likely to enable 2FA

Variable	O.R.	95% C.I.	p-value
Endowment	2.32	[1.44,3.76]	<0.001*

Binomial logistic regression model. Fit with AIC backward elimination.

Account value and risk relate to behavior

Endowment: 2.3x more likely to enable 2FA

Higher **risk of hacking** more likely to enable 2FA

Higher **protection** more likely to enable 2FA

Variable	O.R.	95% C.I.	p-value
Endowment	2.32	[1.44,3.76]	<0.001*
Risk (H)	2.31	[1.22, 4.38]	0.011*
Security (P)	1.46	[1.22, 1.97]	0.043*

Binomial logistic regression model. Fit with AIC backward elimination.

Account value and risk relate to behavior

Endowment: 2.3x more likely to enable 2FA

Higher **risk of hacking** more likely to enable 2FA

Higher **protection** more likely to enable 2FA

Higher **protection** & **endowment** even more likely to enable 2FA

Variable	O.R.	95% C.I.	p-value
Endowment	2.32	[1.44,3.76]	<0.001*
Risk (H)	2.31	[1.22, 4.38]	0.011*
Security (P)	1.46	[1.22, 1.97]	0.043*
Endowment:P	3.61	[1.35, 9.67]	0.012*

Binomial logistic regression model. Fit with AIC backward elimination.

Account value and risk relate to behavior

Endowment: 2.3x more likely to enable 2FA

Higher **risk of hacking** more likely to enable 2FA

Higher **protection** more likely to enable 2FA

Higher **protection** & **endowment** even more likely to enable 2FA

Variable	O.R.	95% C.I.	p-value
Endowment	2.32	[1.44,3.76]	<0.001*
Risk (H)	2.31	[1.22, 4.38]	0.011*
Security (P)	1.46	[1.22, 1.97]	0.043*
Endowment:P	3.61	[1.35, 9.67]	0.012*

Binomial logistic regression model. Fit with AIC backward elimination.

Explains 16% of behavior variance

(McFadden Pseudo R^2)

Prior work theorizes about cognitive load; economics literature shows behavior anchoring in other domains

Enable 2FA ~ Account value + Risk (P, H) with/out 2FA +

Controls

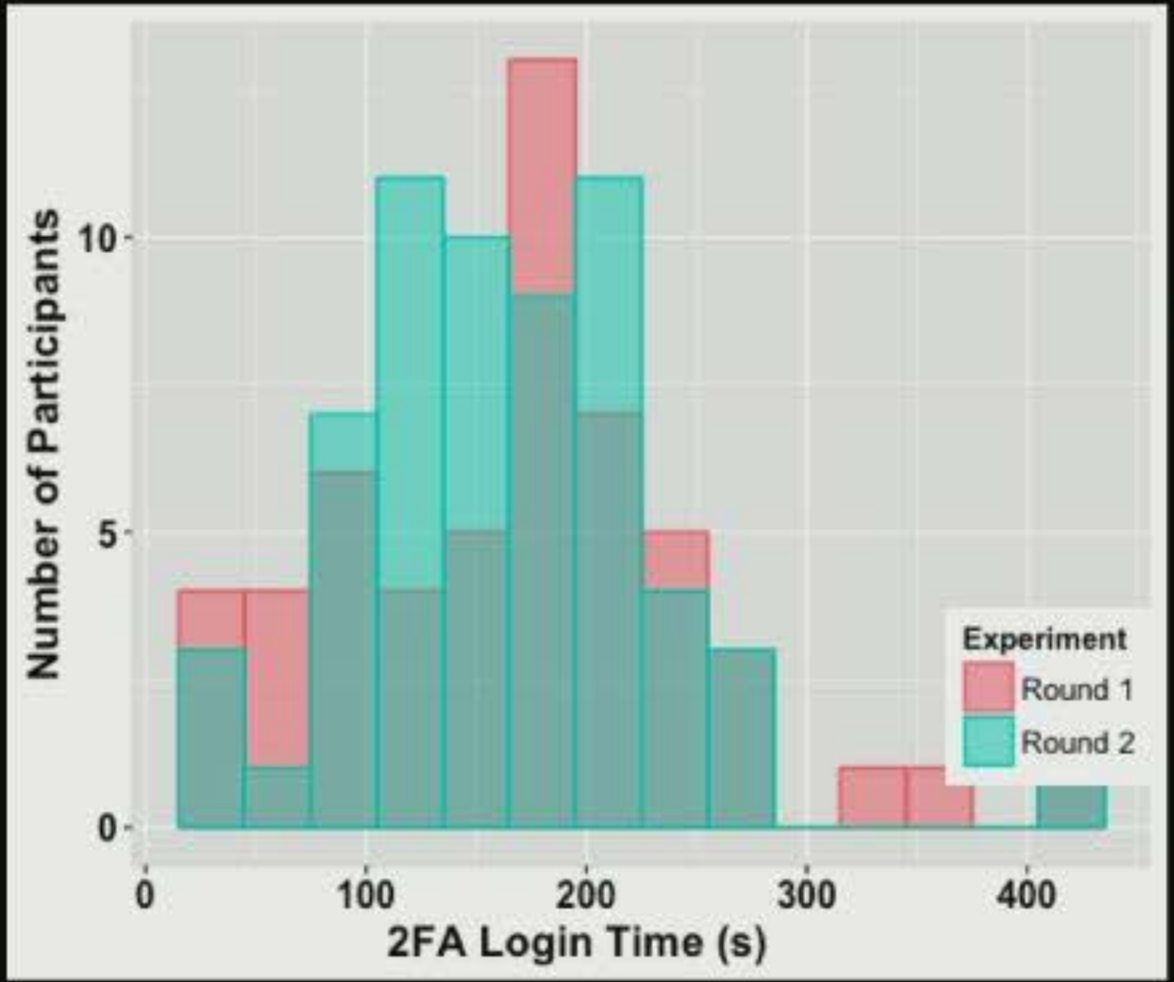
Prior work theorizes about cognitive load; economics literature shows behavior anchoring in other domains

Prior work theorizes about cognitive load; economics literature shows behavior anchoring in other domains

Prior work theorizes about cognitive load; economics literature shows behavior anchoring in other domains

Enable 2FA

~



Costs
proxy:
time spent

+

Controls

Prior work theorizes about cognitive load; economics literature shows behavior anchoring in other domains

Prior work theorizes about cognitive load; economics literature shows behavior anchoring in other domains

Prior work theorizes about cognitive load; economics literature shows behavior anchoring in other domains

Experimental results suggest users are boundedly rational

Risk (H, P) + **Account Value** (Earn/Endow)

explains 9% behavior variance

Experimental results suggest users are boundedly rational

Risk (H, P) + **Account Value** (Earn/Endow)

explains 9% behavior variance

Costs + risk & account value

explains 26% behavior variance

Experimental results suggest users are boundedly rational

Risk (H, P) + **Account Value** (Earn/Endow)

explains 9% behavior variance

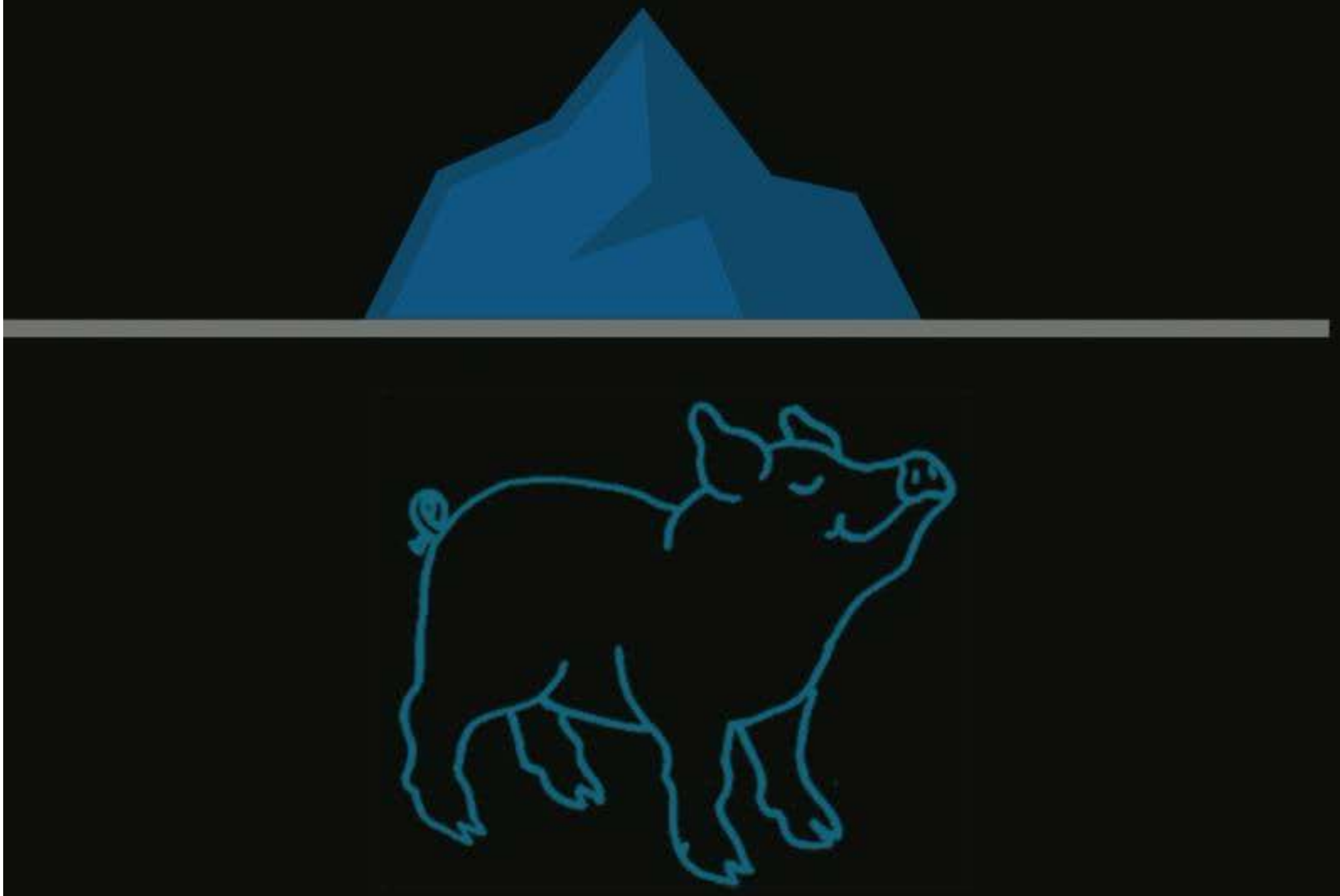
Costs + risk & account value

explains 26% behavior variance

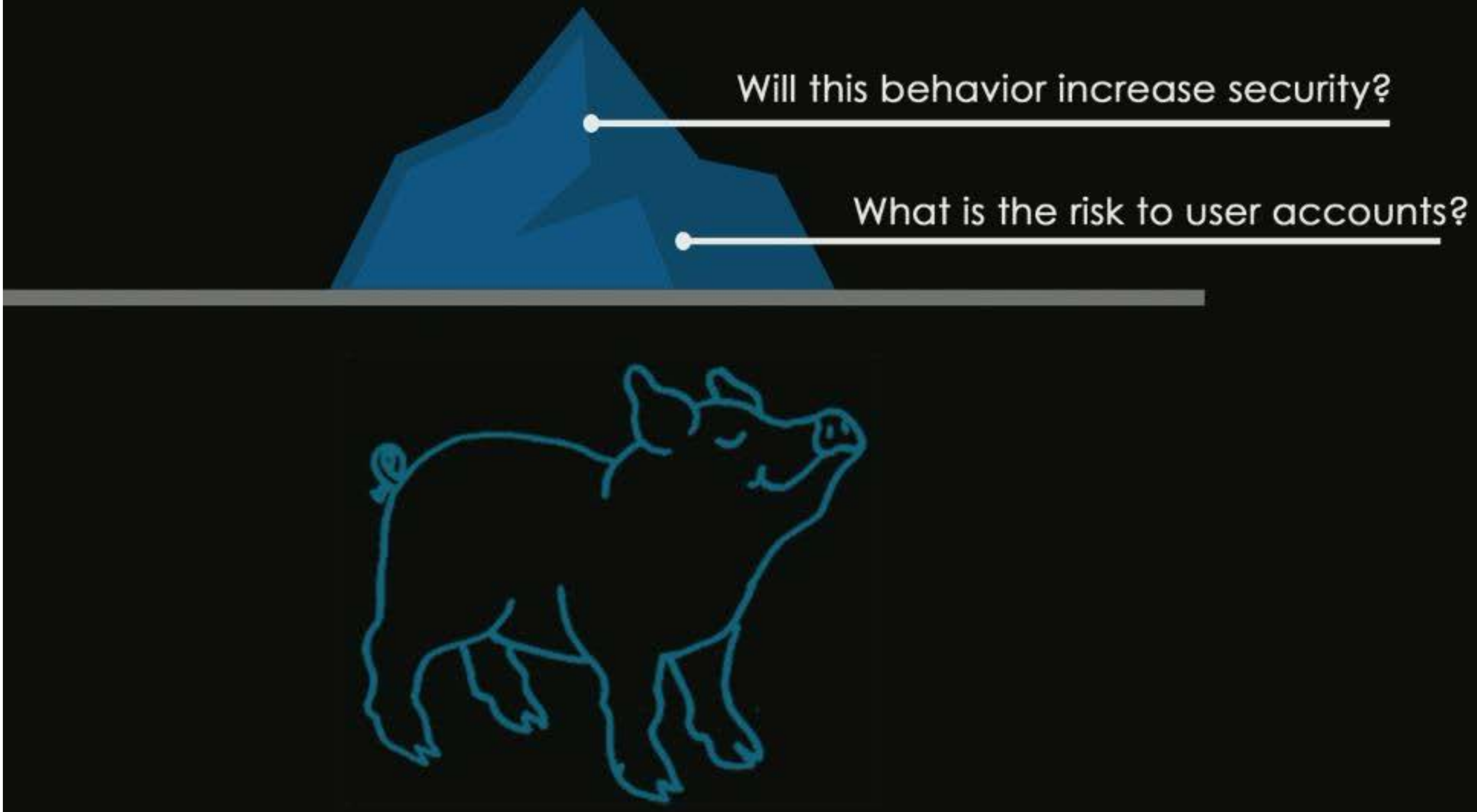
Past behavior + costs +
Risk (H, P) + Account Value (Earn/Endow)

explains 61% of behavior variance

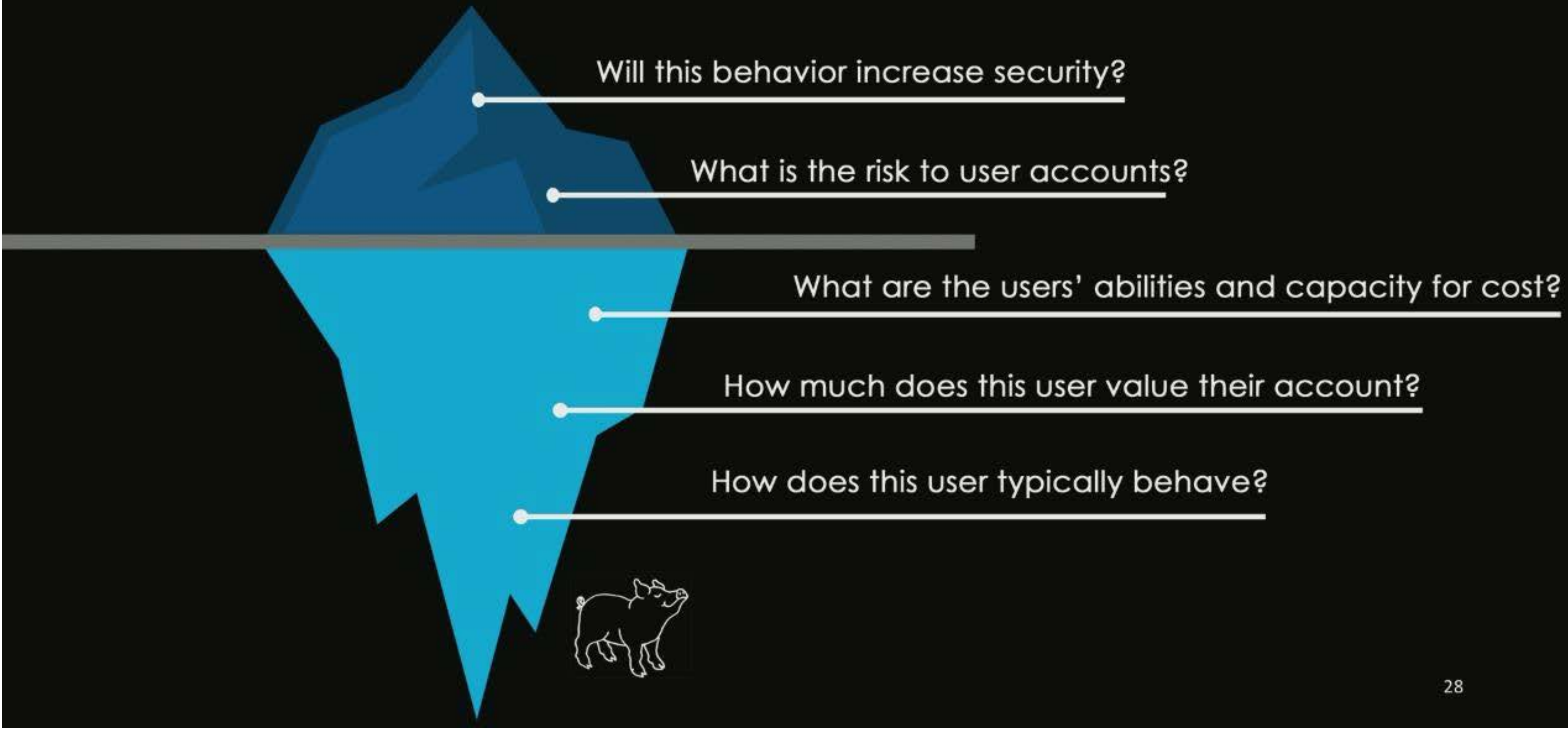
Behavioral security allows us to understand what initially looks irrational and unfixable



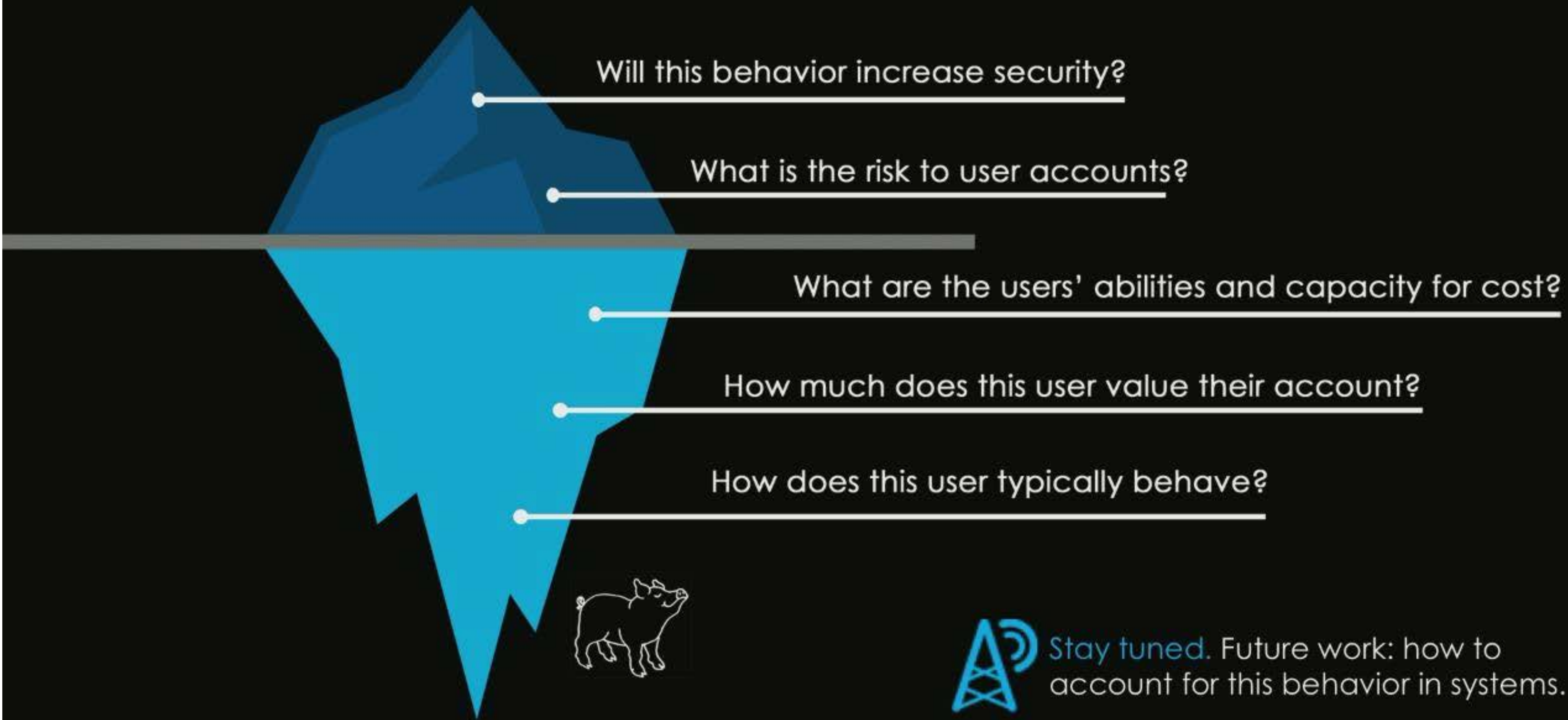
Behavioral security allows us to understand what initially looks irrational and unfixable



Behavioral security allows us to understand what initially looks irrational and unfixable



Behavioral security allows us to understand what initially looks irrational and unfixable



Stay tuned. Future work: how to account for this behavior in systems.

Today's Agenda: finding a model of best fit for security behavior & balancing structural inequities in security

Model of best fit for security behavior

Balancing structural inequities in real systems

Epistemology of methods

Systematic individual differences across security domains: structural inequities

Account & Device Security

[S&P16]
[EC18]
[CCS16]
[CCS18a]
[CCS18b]
[S&P19a]
[S&P19b]
[CHI17]
[TWEB18]
[WAY17]

Spam & Fake News

[CHI18]
[FAT*19a]
[FAT*19b]

Enterprise Security

[S&P18]
[BigData16]
[USENIXSec18]
Distinguished Paper

Encryption & Data Use

[USENIX Sec17]
[SOUPS18]
[ICWSM18]
[ICWSM19]
[FOCI18]

Structural inequities fall along many axes, not just skill

Skills and Abilities

Structural inequities fall along many axes, not just skill

Skills and Abilities

Socioeconomic Status

Structural inequities fall along many axes, not just skill

Skills and Abilities

Culture or Identity

Socioeconomic Status

Case study: Inequities in social spam susceptibility

Redmiles, E.M., Chachra, N., and Waismeyer, B. *Examining the Demand for Spam: Who Clicks?*
ACM **Conference on Human Factors in Computing Systems (CHI)** 2018.

Case study: Inequities in social spam susceptibility

Why do people fall for spam on facebook

Redmiles, E.M., Chachra, N., and Waismeyer, B. *Examining the Demand for Spam: Who Clicks?*
ACM **Conference on Human Factors in Computing Systems (CHI)** 2018.

Collaboration with Facebook to model spam susceptibility from Facebook log records

Economic
Behavioral Econ

Security Measurement
Large-scale Log Analysis

Social Scientific
Surveys & Interviews

Scientifically understand insecure behavior

Two research questions grounded in prior work on email spam & security tool adoption

RQ1: What is the quantified impact of factors suggested by prior work on email spam (gender, age, skill)?

RQ2: What is the quantified impact of inequities in social influence driven by culture (network)?

Analyzed 600,000 records of user-content interactions

Spam (n=300,000)

Viewer, content pairs sampled over 20 days in July 2017

Content was spam that contained a URL

Ham (n=300,000)

Viewer, content pairs sampled over same 20 days

Content that had not been identified as spam as of 28 days later

Facebook spam is malicious or deceptive content that…

attempts to elicit illegitimate financial gain
e.g., by gathering account credentials (phishing)

distributes malware or hijacks user accounts

fails to deliver on a promised outcome
for example, content in the post (e.g., preview image) does not match the content the user receives

Analyzed 600,000 records of user-content interactions

Spam (n=300,000)

Viewer, content pairs sampled over 20 days in July 2017

Content was spam that contained a URL

Ham (n=300,000)

Viewer, content pairs sampled over same 20 days

Content that had not been identified as spam as of 28 days later

Facebook spam is malicious or deceptive content that…

attempts to elicit illegitimate financial gain
e.g., by gathering account credentials (phishing)

distributes malware or hijacks user accounts

fails to deliver on a promised outcome
for example, content in the post (e.g., preview image) does not match the content the user receives

Instantiation of four sets of features to test RQs

Demographics

Age, gender

Instantiation of four sets of features to test RQs

Demographics

Age, gender

Activity level on Facebook

L28: number of days out of the last 28 that the person was active

Instantiation of four sets of features to test RQs

Demographics

Age, gender

Activity level on Facebook

L28: number of days out of the last 28 that the person was active

Country attributes

Spam prevalence

National clicking norms (spam CTR/ham CTR by country)

Instantiation of four sets of features to test RQs

Demographics

Age, gender

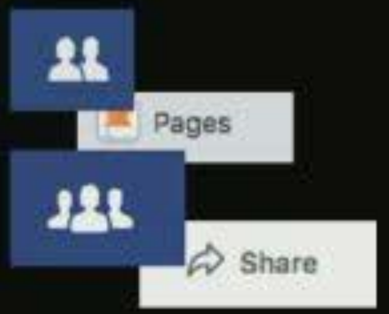
Activity level on Facebook

L28: number of days out of the last 28 that the person was active

Country attributes

Spam prevalence
National clicking norms (spam CTR/ham CTR by country)

Content attributes



User's relationship to content (friend, friend of friend, page)
Whether the content was reshared

Predict whether viewer *v* clicked on piece of content c

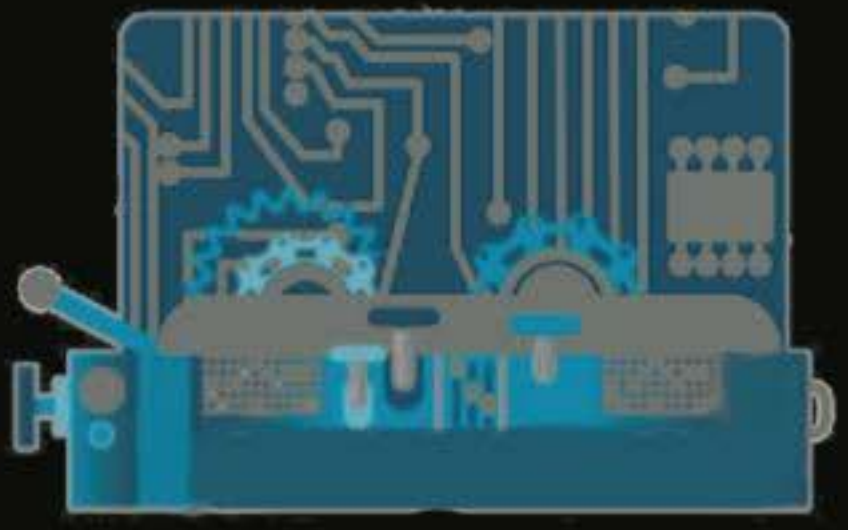
Features

Predict whether viewer *v* clicked on piece of content c

Features

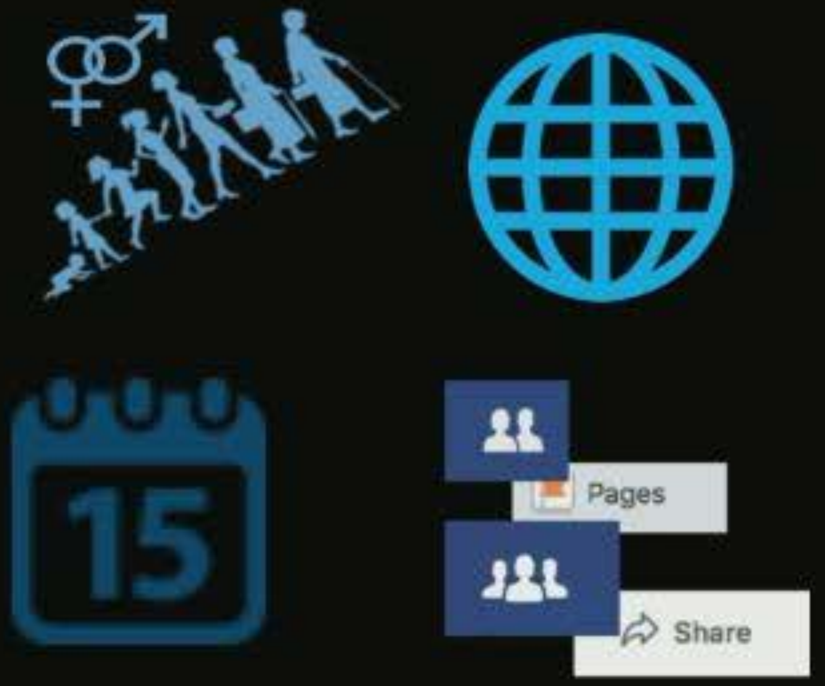
Logistic Regression

80:20 Train-Test Split



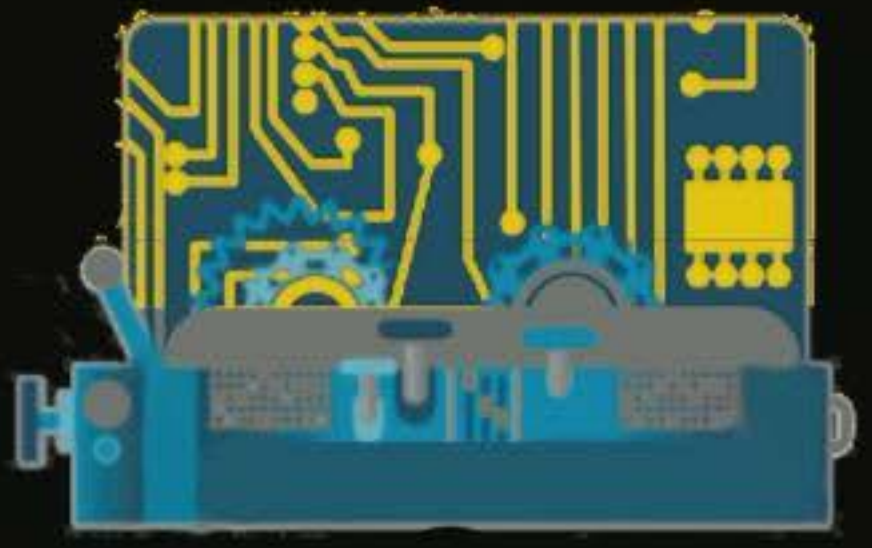
Predict whether viewer *v* clicked on piece of content c

Features



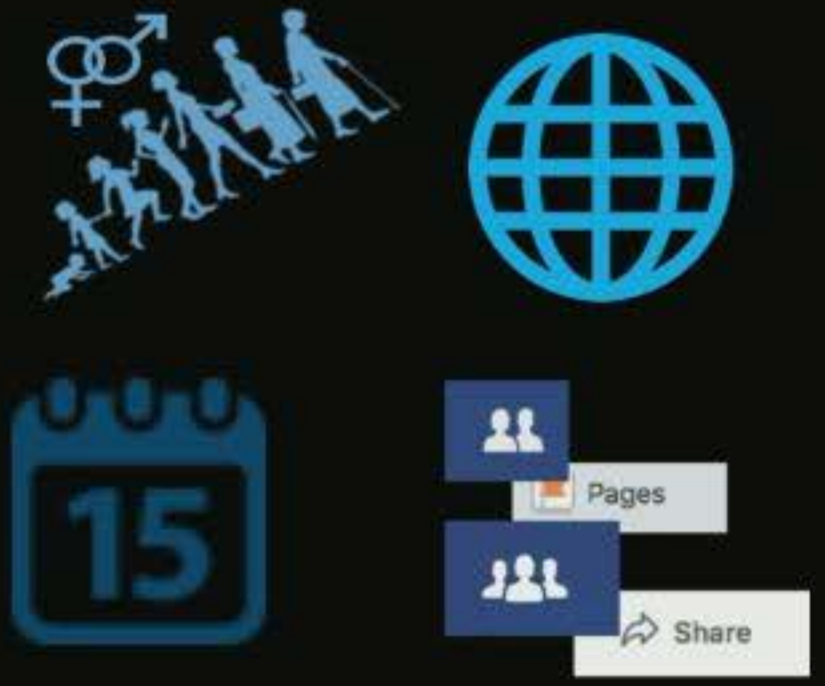
Logistic Regression

80:20 Train-Test Split



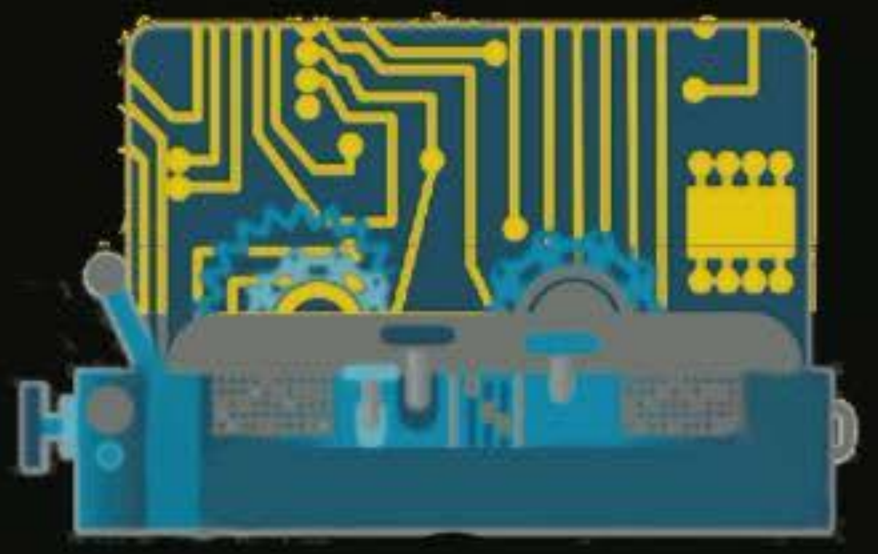
Predict whether viewer *v* clicked on piece of content c

Features



Logistic Regression

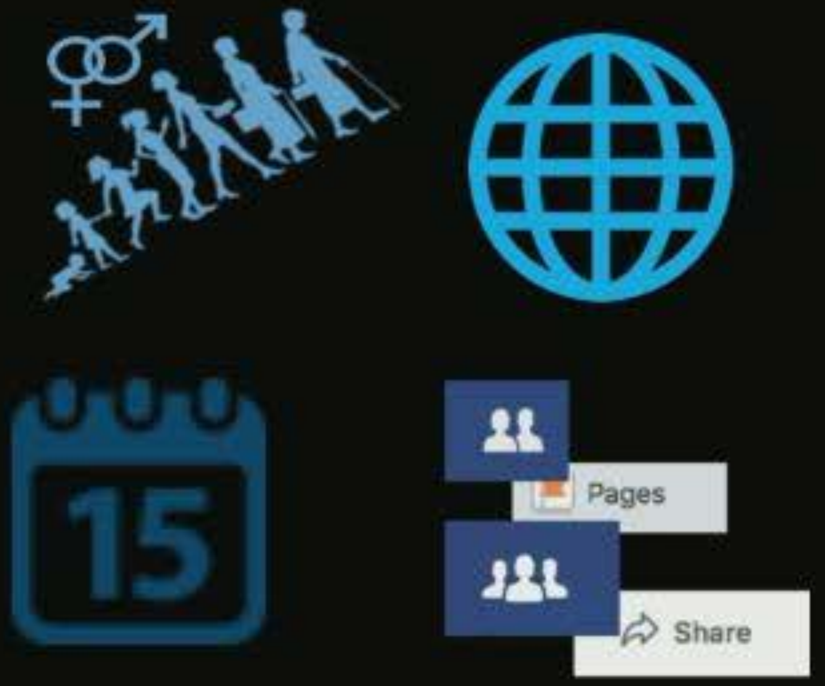
80:20 Train-Test Split



Click

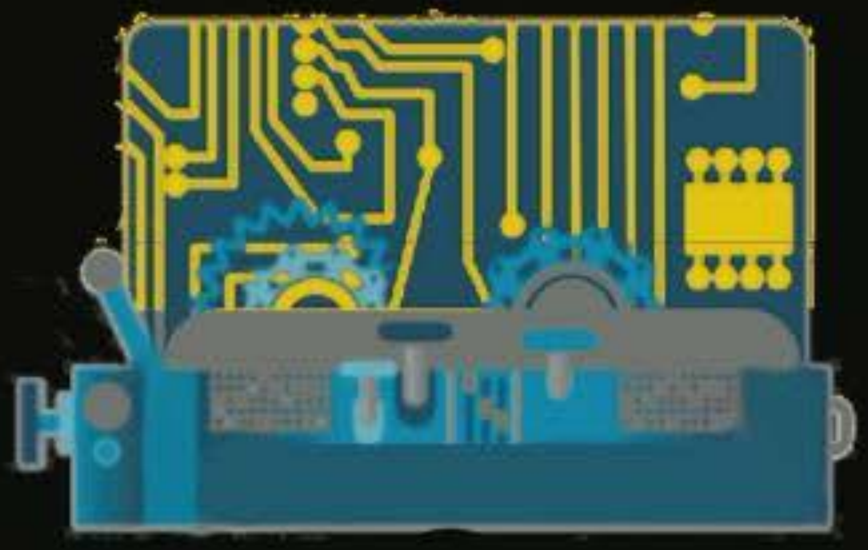
Predict whether viewer *v* clicked on piece of content c

Features



Logistic Regression

80:20 Train-Test Split



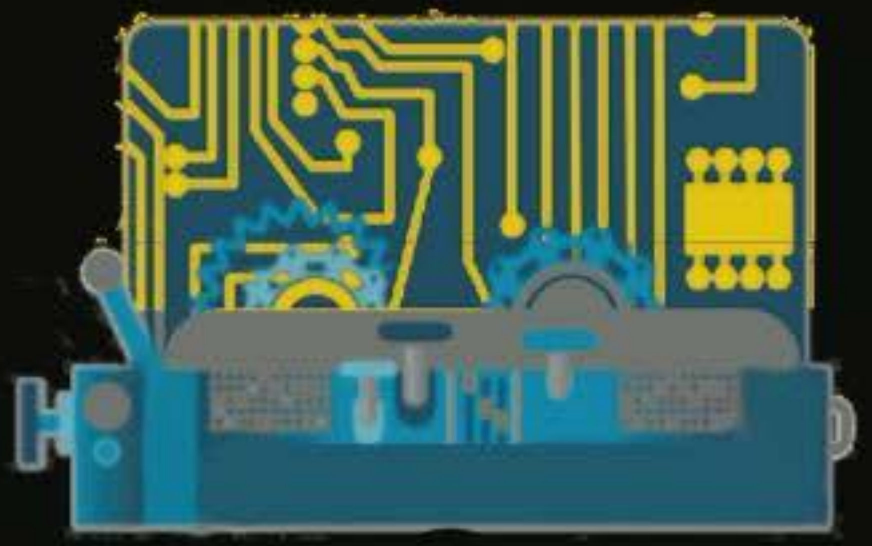
Click

Predict whether viewer *v* clicked on piece of content c

Features

Logistic Regression

80:20 Train-Test Split

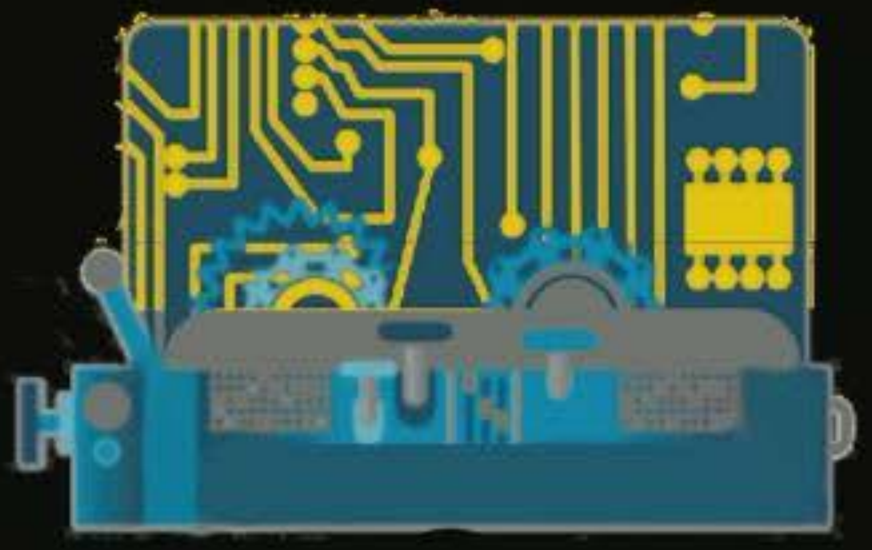


Click

Predict whether viewer *v* clicked on piece of content c

Features

Logistic Regression
80:20 Train-Test Split



Click

AUC = 0.72   AUC = 0.80

Our Model & Prior Work

“Women are more likely to click on spam”

Our Model & Prior Work

“Women are more likely to click on spam”

New research question:

Why?

Two researchers qualitatively coded 250 spam samples

Inductively defined codebook of spam types

Independently double coded content; Maximum 6% margin of error.

Our Model & Prior Work

"Women are more likely to click on spam"

New research question:

Why?

Two researchers qualitatively coded 250 spam samples

Inductively defined codebook of spam types

Independently double coded content; Maximum 6% margin of error.

Two researchers qualitatively coded 250 spam samples

Inductively defined codebook of spam types

Independently double coded content; Maximum 6% margin of error.

Shopping

38% of sample

Media

42% of sample

Interactives

18% of sample

Two researchers qualitatively coded 250 spam samples

Inductively defined codebook of spam types

Independently double coded content; Maximum 6% margin of error.

Shopping

38% of sample

Media

42% of sample

Interactives

18% of sample

Shopping spam 2x CTR vs. media

Two researchers qualitatively coded 250 spam samples

Inductively defined codebook of spam types

Independently double coded content; Maximum 6% margin of error.

Shopping

38% of sample

Women see more (66%)

Media

42% of sample

Men see more (75%)

Interactives

18% of sample

Shopping spam 2x CTR vs. media

Two researchers qualitatively coded 250 spam samples

Inductively defined codebook of spam types

Independently double coded content; Maximum 6% margin of error.

Shopping

38% of sample

Women see more (66%)

Media

42% of sample

Men see more (75%)

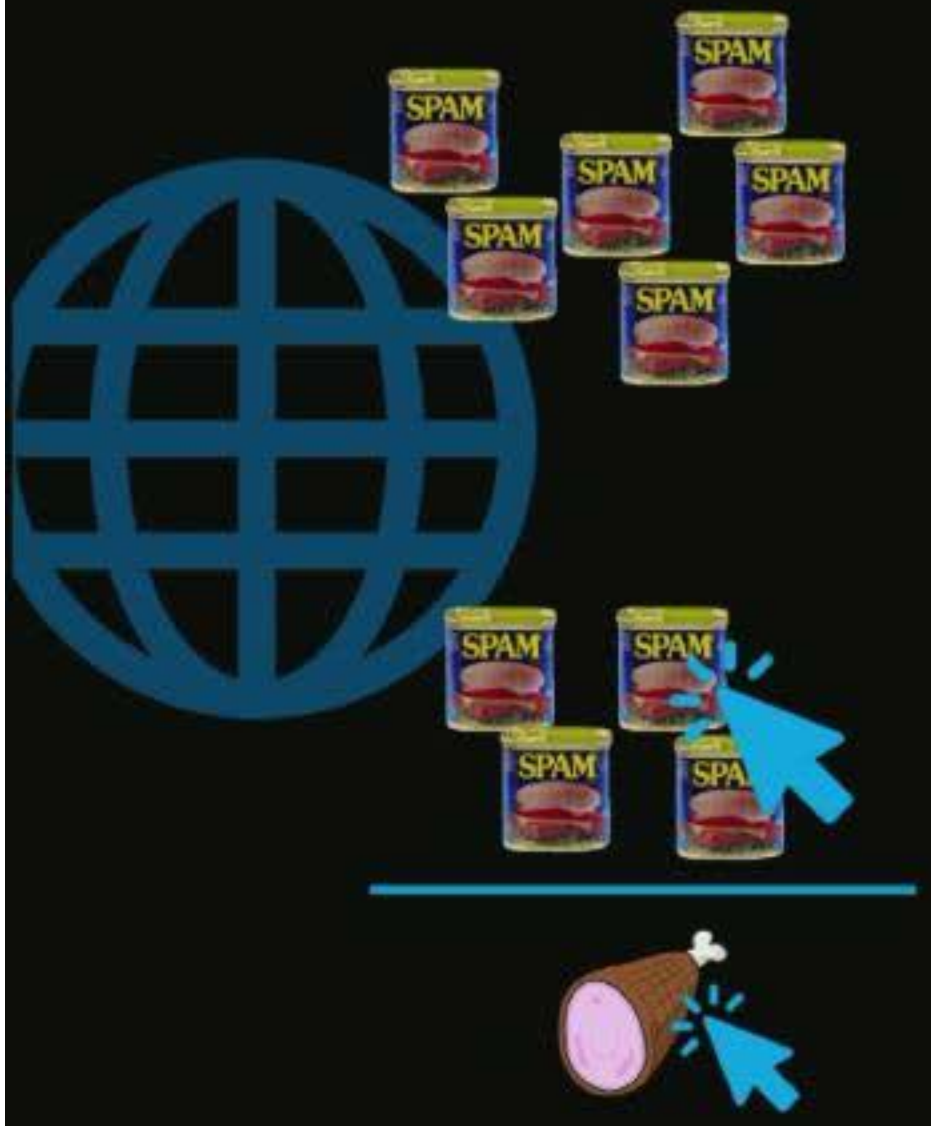
Interactives

18% of sample

Shopping spam 2x CTR vs. media

Women have a harder job to detect spam

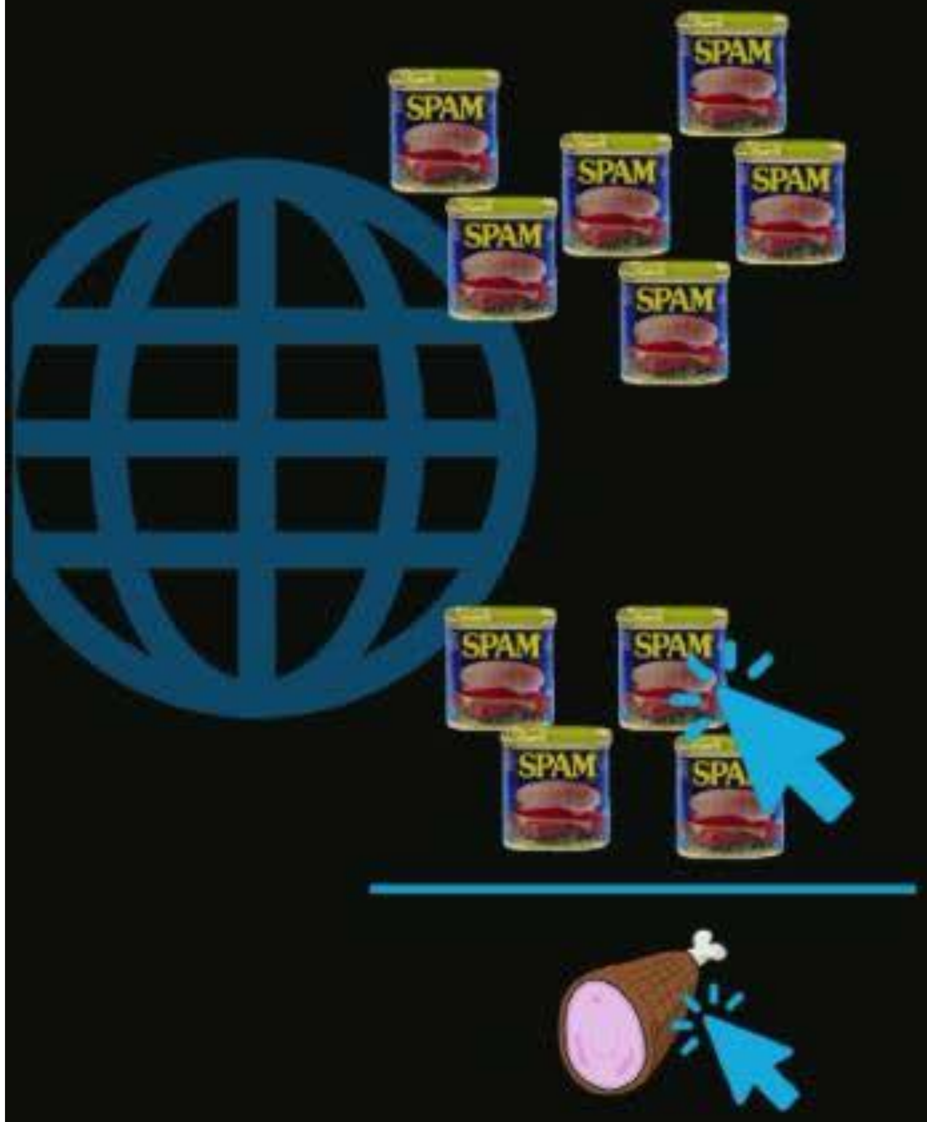
Country (network) features influence spam susceptibility



People in countries w/ high spam prevalence
59% **less likely** to click on spam

High proportion of spam to ham clicking
more likely to click on spam

Country (network) features influence spam susceptibility

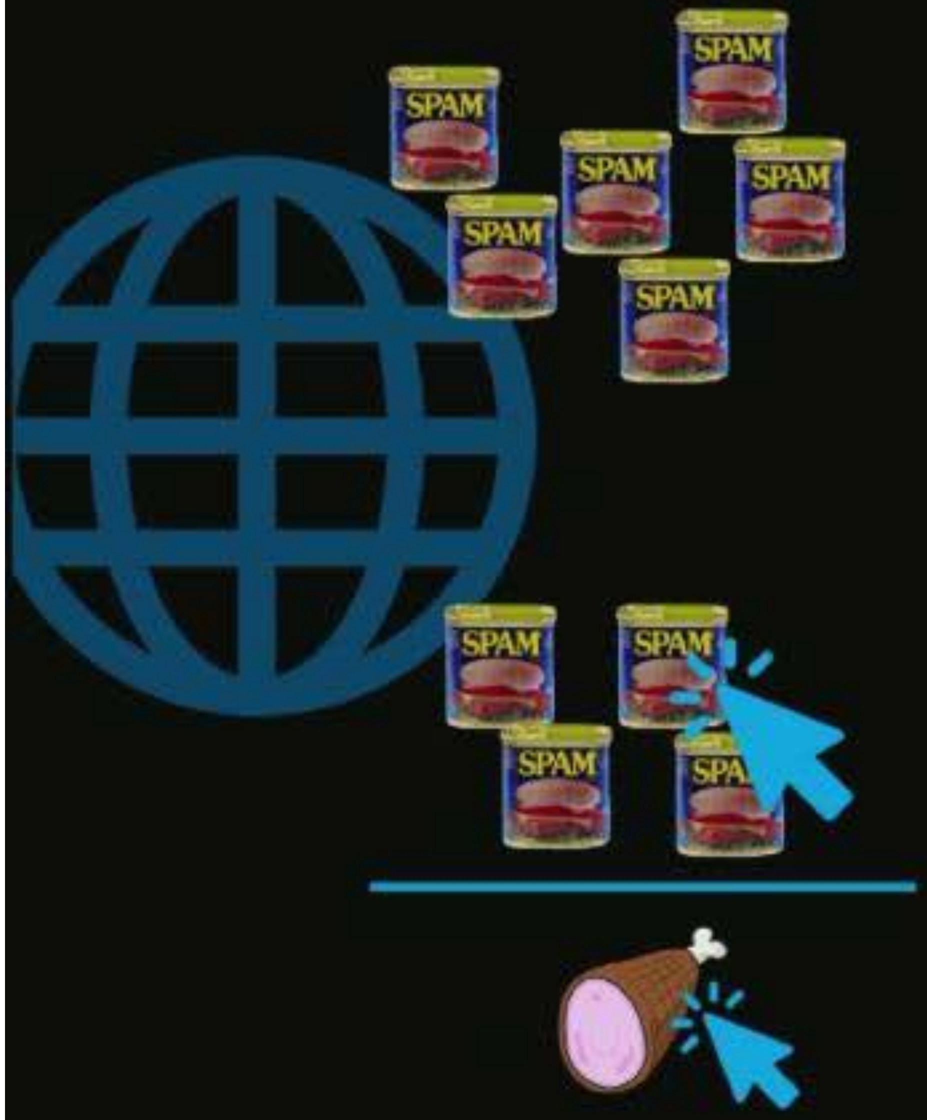


People in countries w/ high spam prevalence
59% **less likely** to click on spam

This is not just true of end users, testers are also more effective at vulnerability detection with more system experience [S&P18]

High proportion of spam to ham clicking
more likely to click on spam

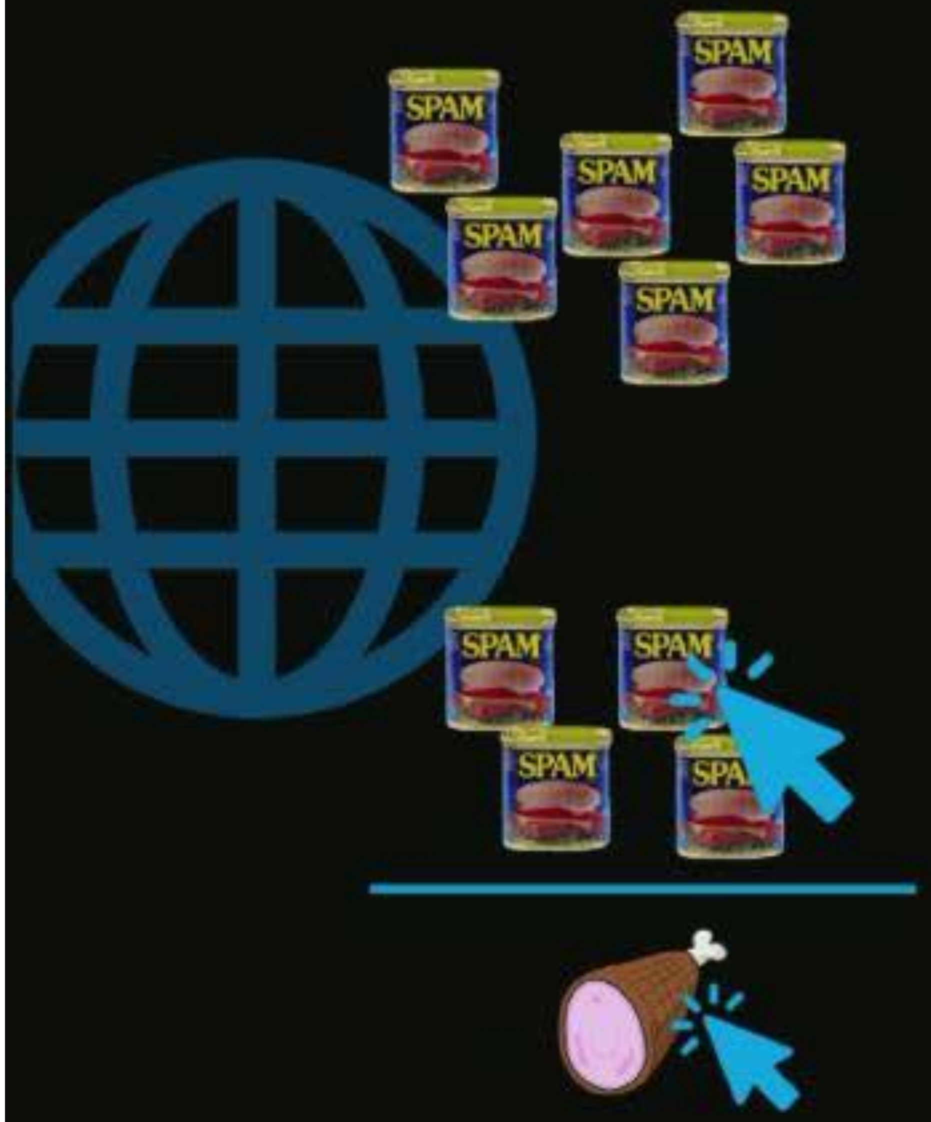
Country (network) features influence spam susceptibility



People in countries w/ high spam prevalence
59% **less likely** to click on spam

High proportion of spam to ham clicking
more likely to click on spam

Country (network) features influence spam susceptibility

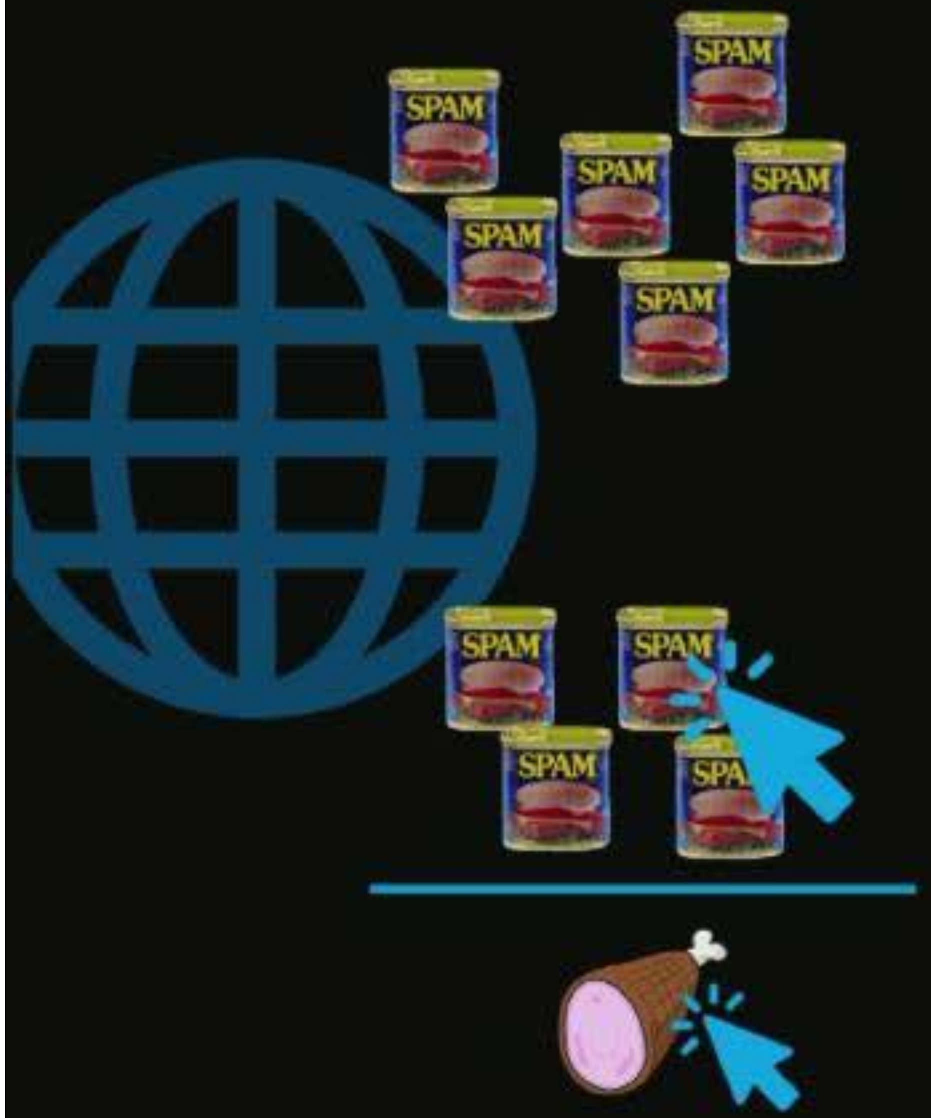


People in countries w/ high spam prevalence
59% **less likely** to click on spam

Further support that system experience matters

High proportion of spam to ham clicking
more likely to click on spam

Country (network) features influence spam susceptibility



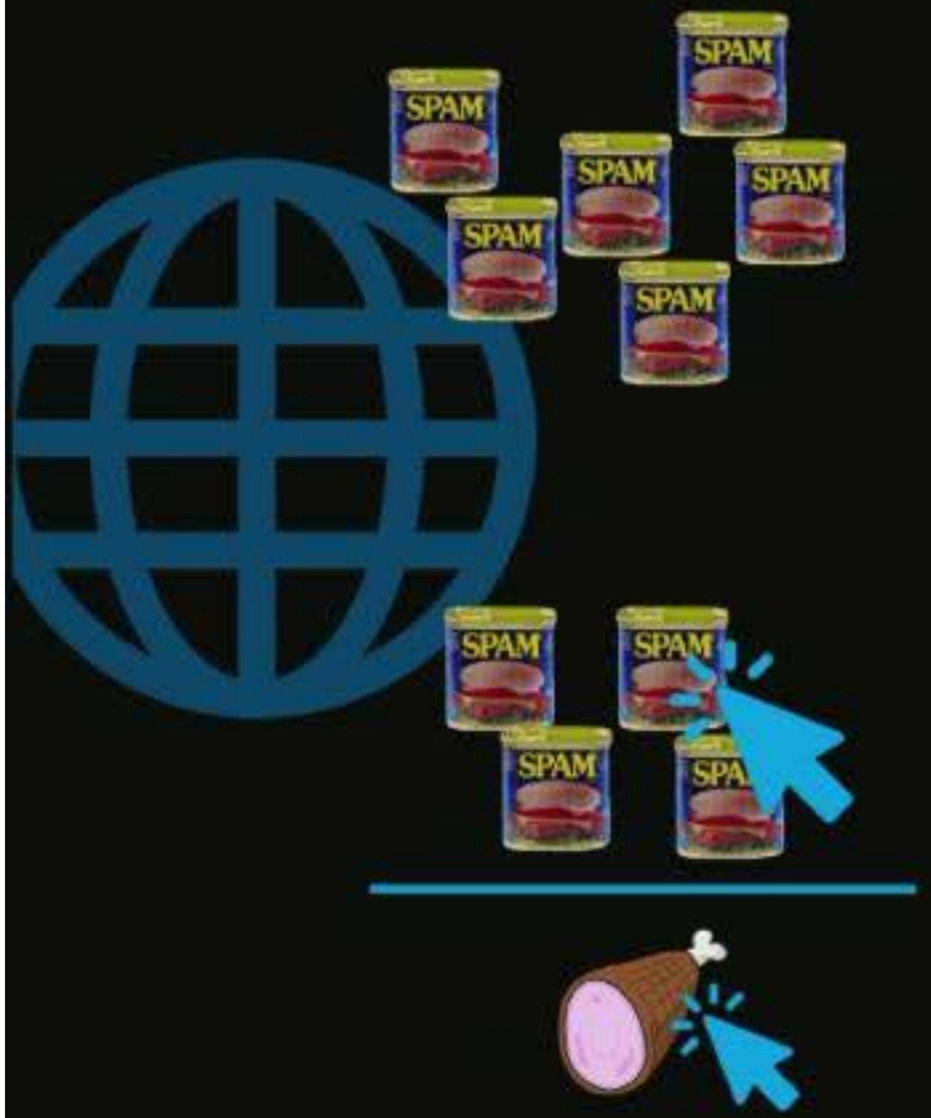
People in countries w/ high spam prevalence
59% **less likely** to click on spam

Further support that system experience matters

This is not just true of end users, testers are also more effective at vulnerability detection with more system experience [S&P18]

High proportion of spam to ham clicking
more likely to click on spam

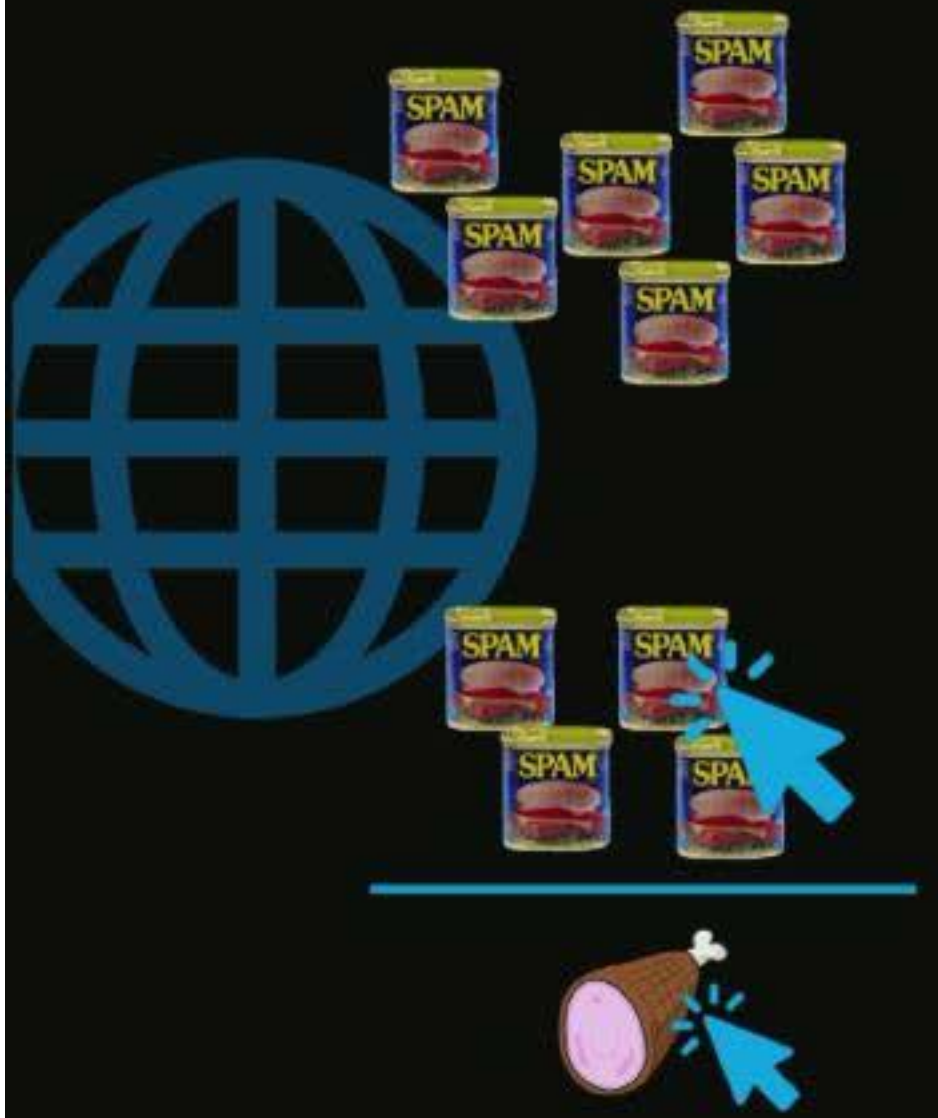
Country (network) features influence spam susceptibility



People in countries w/ high spam prevalence
59% **less likely** to click on spam

High proportion of spam to ham clicking
more likely to click on spam

Country (network) features influence spam susceptibility



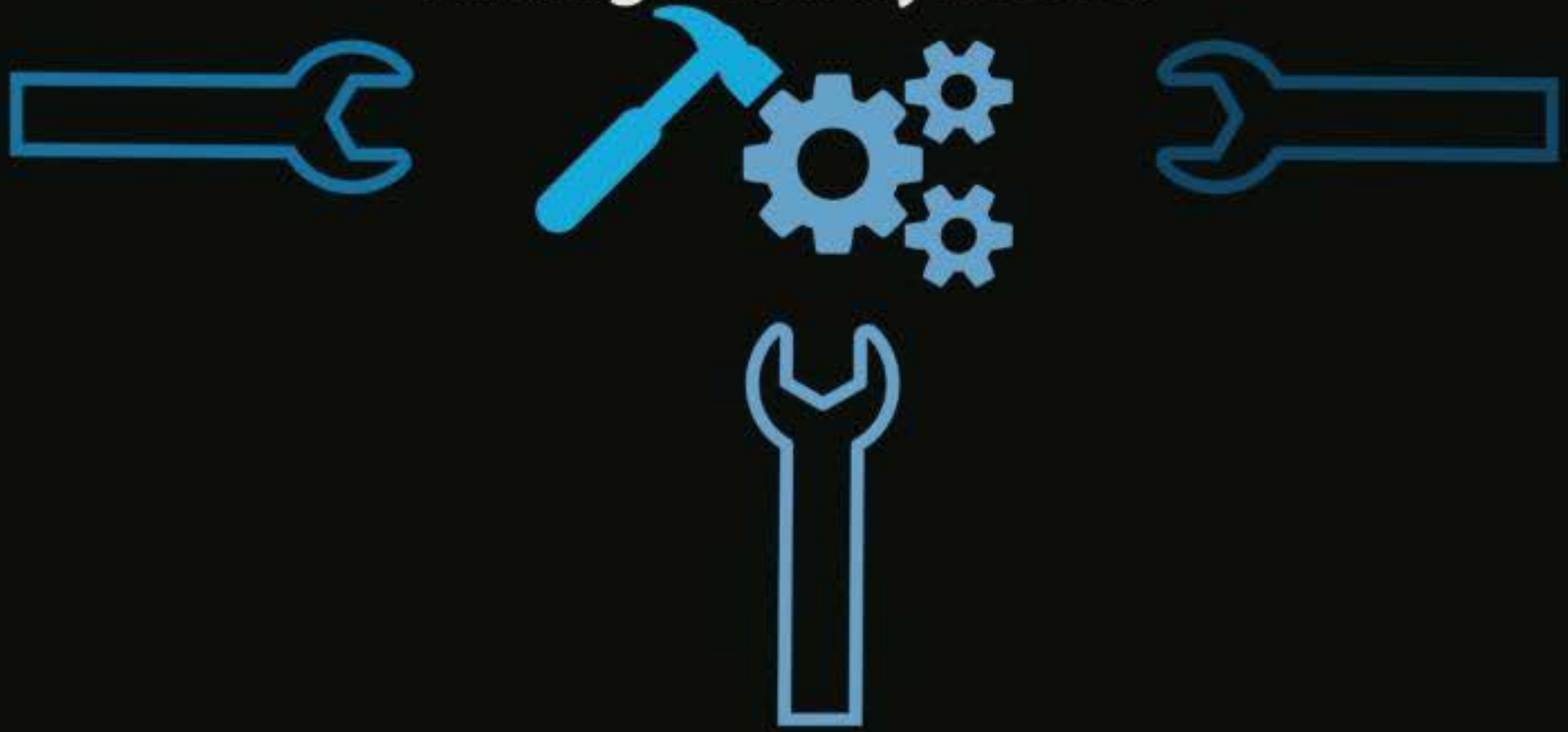
People in countries w/ high spam prevalence
59% **less likely** to click on spam

High proportion of spam to ham clicking
more likely to click on spam

Social norms may provide feedback re: insecure behavior

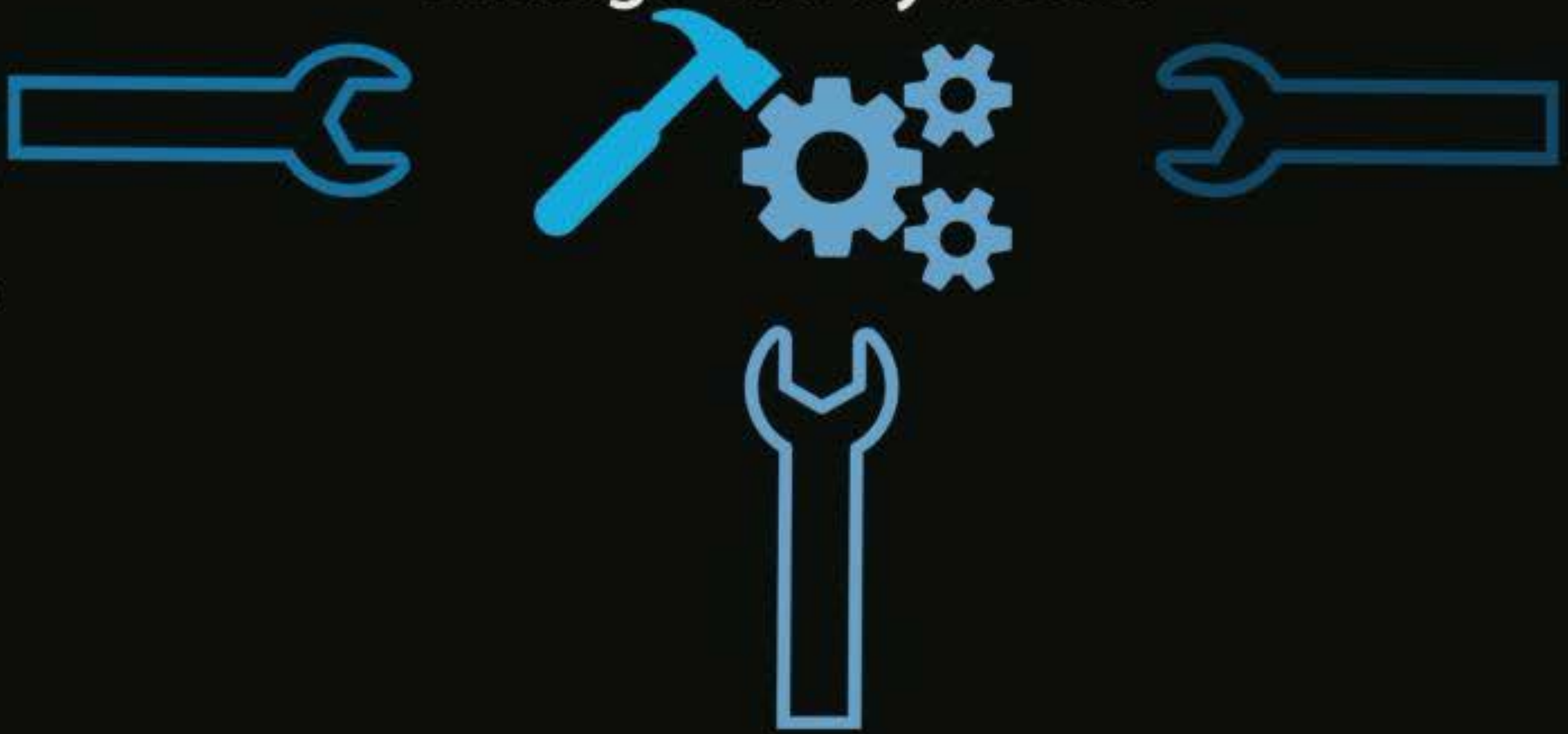
System changes that improve equity can increase security

Change the systems



System changes that improve equity can increase security

Change the systems

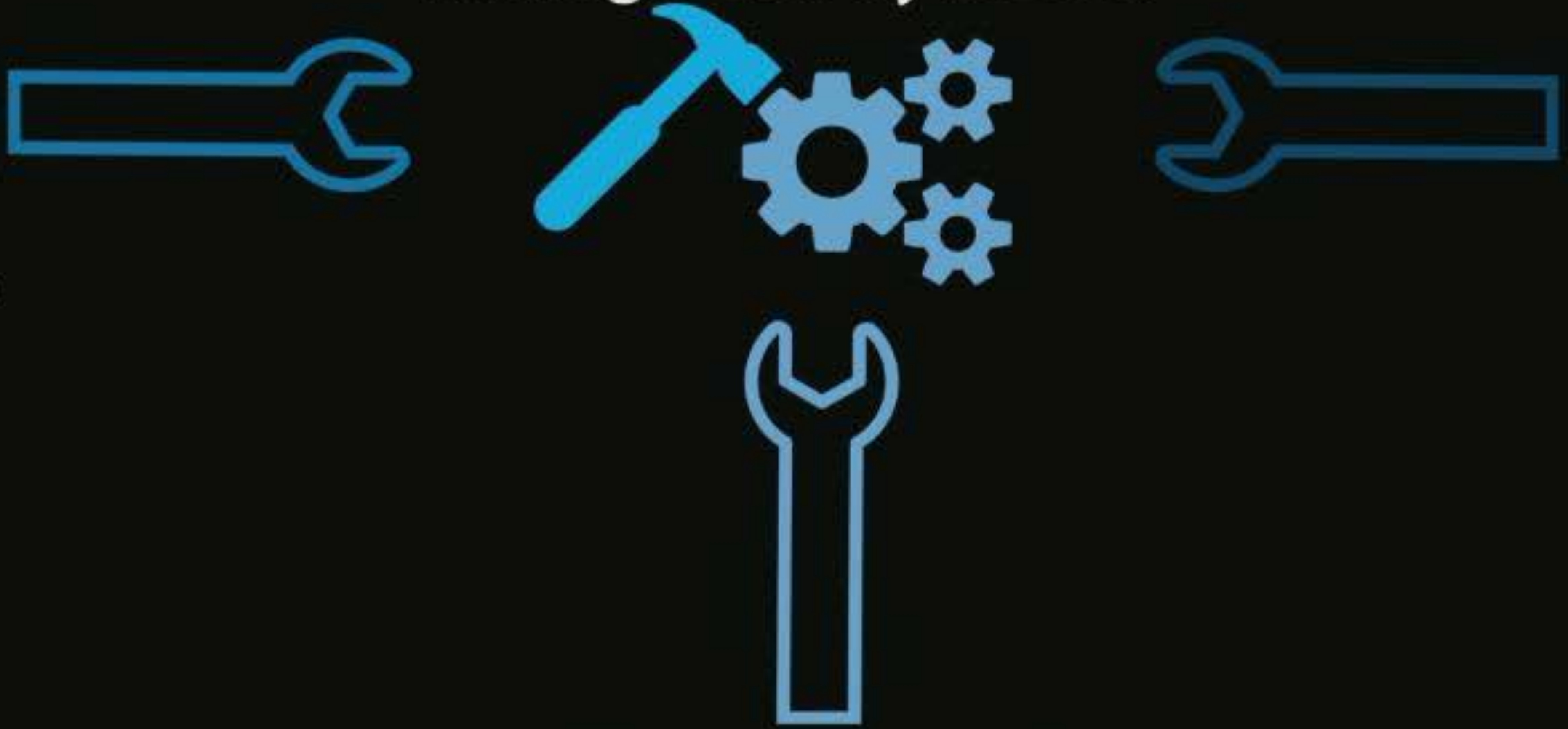


Social Influence

Surface number of friends who have reported a piece of content

System changes that improve equity can increase security

Change the systems



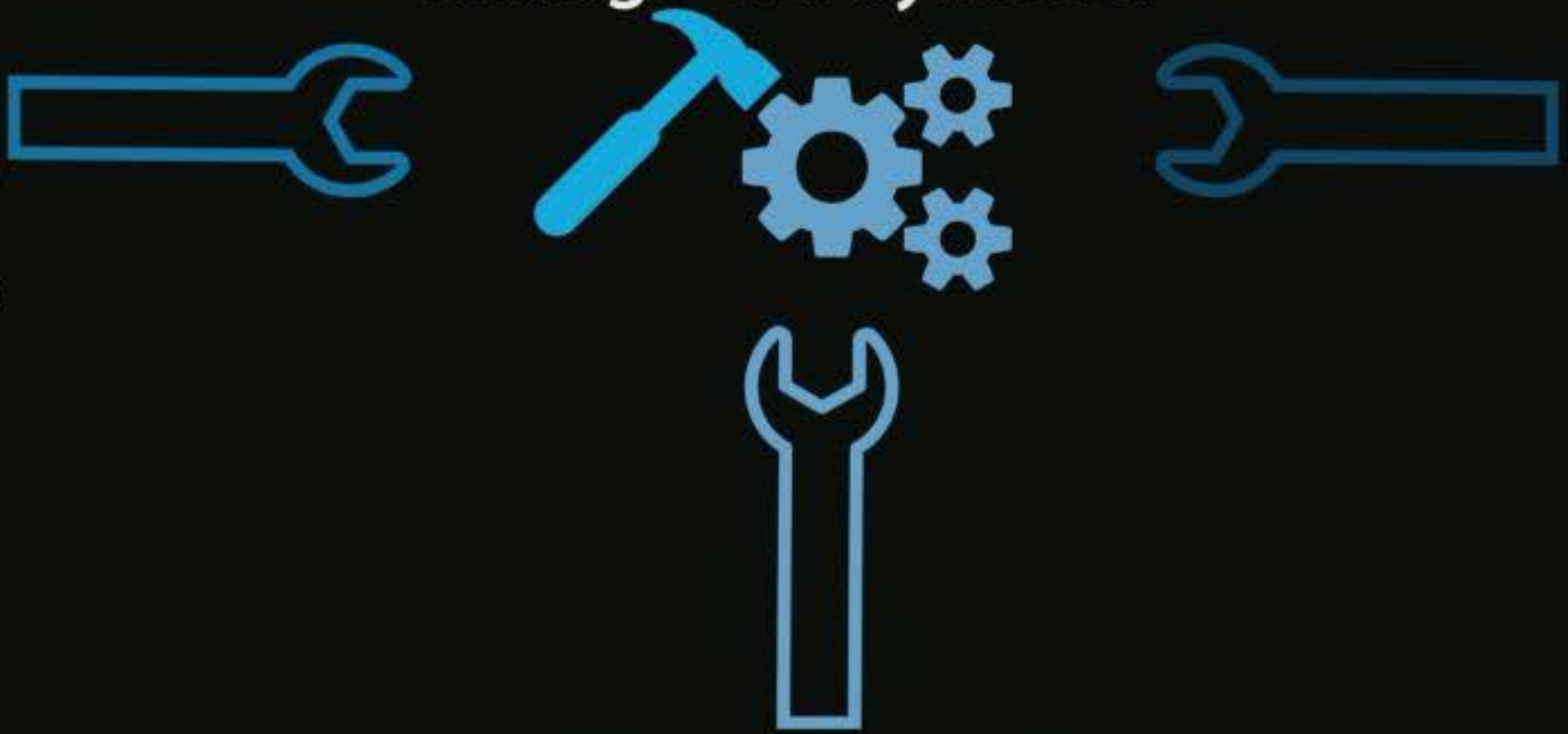
Social Influence

Surface number of friends who have reported a piece of content

Content Heuristics

Authenticity indicators
Downrank promotions

System changes that improve equity can increase security



Social Influence
Surface number of friends who have reported a piece of content

Classification
Use CTR as a feature
Separate classifiers

Content Heuristics
Authenticity indicators
Downrank promotions

System changes that improve equity can increase security

Multiple changes to real facebook systems

Change the systems

Social Influence
Surface number of friends who have reported a piece of content

Classification
Use CTR as a feature
Separate classifiers

Content Heuristics
Authenticity indicators
Downrank promotions

Finding broad inequities through survey methods

Economic
Behavioral Econ
Mechanism Design

Security Measurement
Large-scale Log Analysis

Social Scientific
Survey Methods
Interview Studies

Scientifically understand insecure behavior

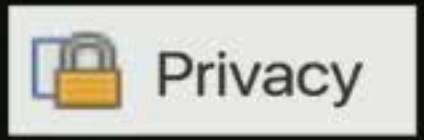
Identified multiple policy-relevant, general inequities using a fully representative survey dataset (n=3,000)

- Survey data on general security & privacy collected by Data&Society
- Probabilistic random digit dial (RDD) survey (n=3,000) in the U.S.
- Statistically raked (weighted) to generalize to the entire U.S. within 2.7%

One of many inequity-related findings: inequities can be inherited

$

Higher income parents are 66% more likely to help their children with

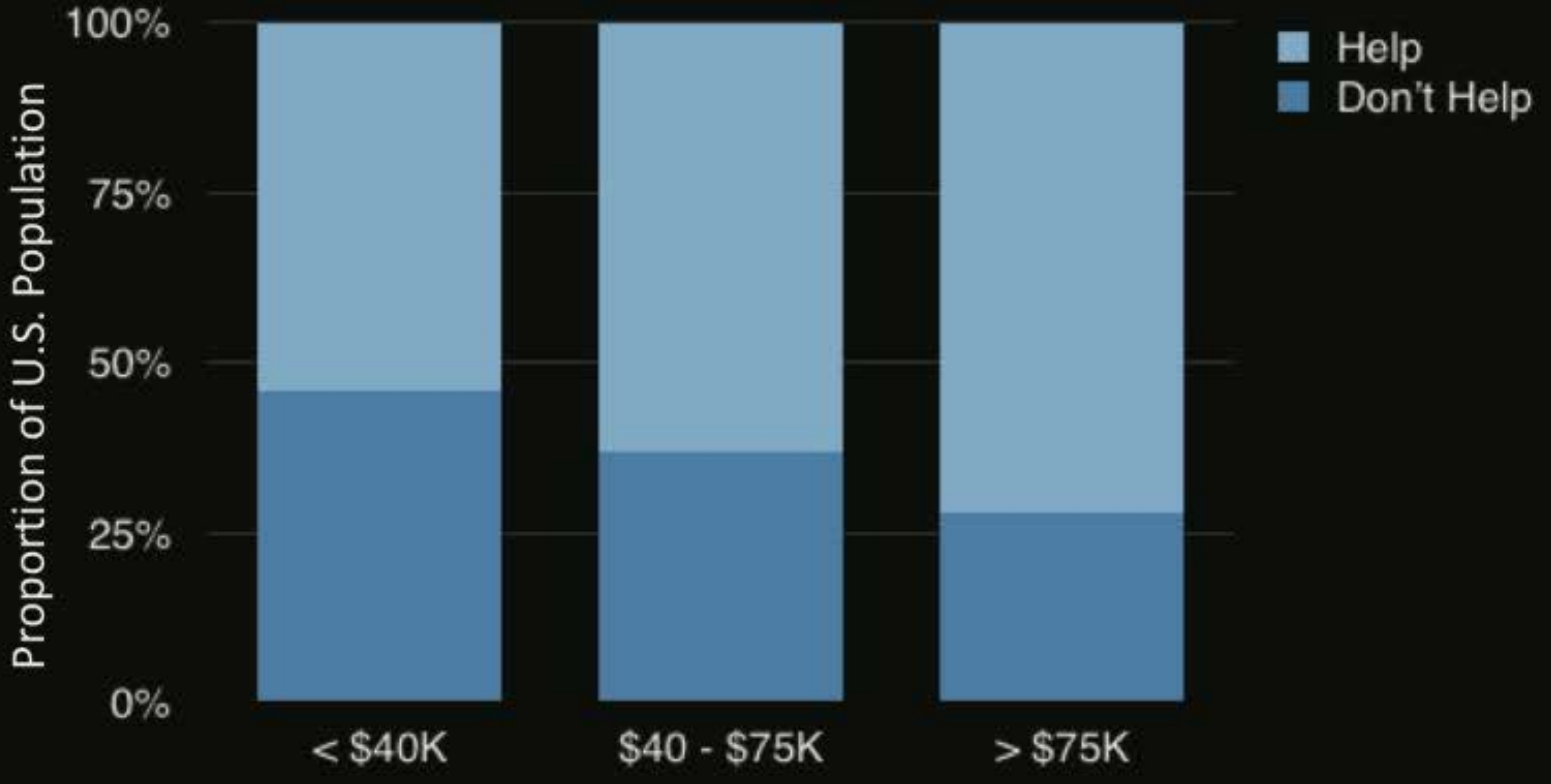


Redmiles, E.M. *Net Benefits: Digital Inequities in Social Capital, Privacy Preservation, and Digital Parenting Practices of U.S. Social Media Users.* AAAI **International Conference on Web and Social Media (ICWSM)** 2018.

One of many inequity-related findings: inequities can be inherited

$

Higher income parents are 66% more likely to help their children with Privacy



Redmiles, E.M. *Net Benefits: Digital Inequities in Social Capital, Privacy Preservation, and Digital Parenting Practices of U.S. Social Media Users.* AAAI **International Conference on Web and Social Media (ICWSM)** 2018.

One of many inequity-related findings: inequities can be inherited

Higher income parents are 66% more likely to help their children with Privacy

Parents with some college education are 3.2x more likely to help children with Privacy

Redmiles, E.M. *Net Benefits: Digital Inequities in Social Capital, Privacy Preservation, and Digital Parenting Practices of U.S. Social Media Users.* AAAI **International Conference on Web and Social Media (ICWSM)** 2018.

Today's Agenda: finding a model of best fit for security behavior & balancing structural inequities in security

Model of best fit for security behavior

Balancing structural inequities in real systems

Epistemology of methods

A science of behavioral security: comparing method & sample validity; building scalable experimentation tools

A science of behavioral security: comparing method & sample validity; building scalable experimentation tools

CCS2018 When to use observational log data vs. survey data

A science of behavioral security: comparing method & sample validity; building scalable experimentation tools

CCS2018 When to use observational log data vs. survey data

n = 517,932

Host records

response to update prompts

n = 2,092

Survey carefully constructed to match intended behavior response to same prompts

A science of behavioral security: comparing method & sample validity; building scalable experimentation tools

CCS2018 When to use observational log data vs. survey data

A science of behavioral security: comparing method & sample validity; building scalable experimentation tools

CCS2018 When to use observational log data vs. survey data

S&P2019 generalizability of Mturk & webpanels vs. probabilistic samples

A science of behavioral security: comparing method & sample validity; building scalable experimentation tools

CCS2018 When to use observational log data vs. survey data

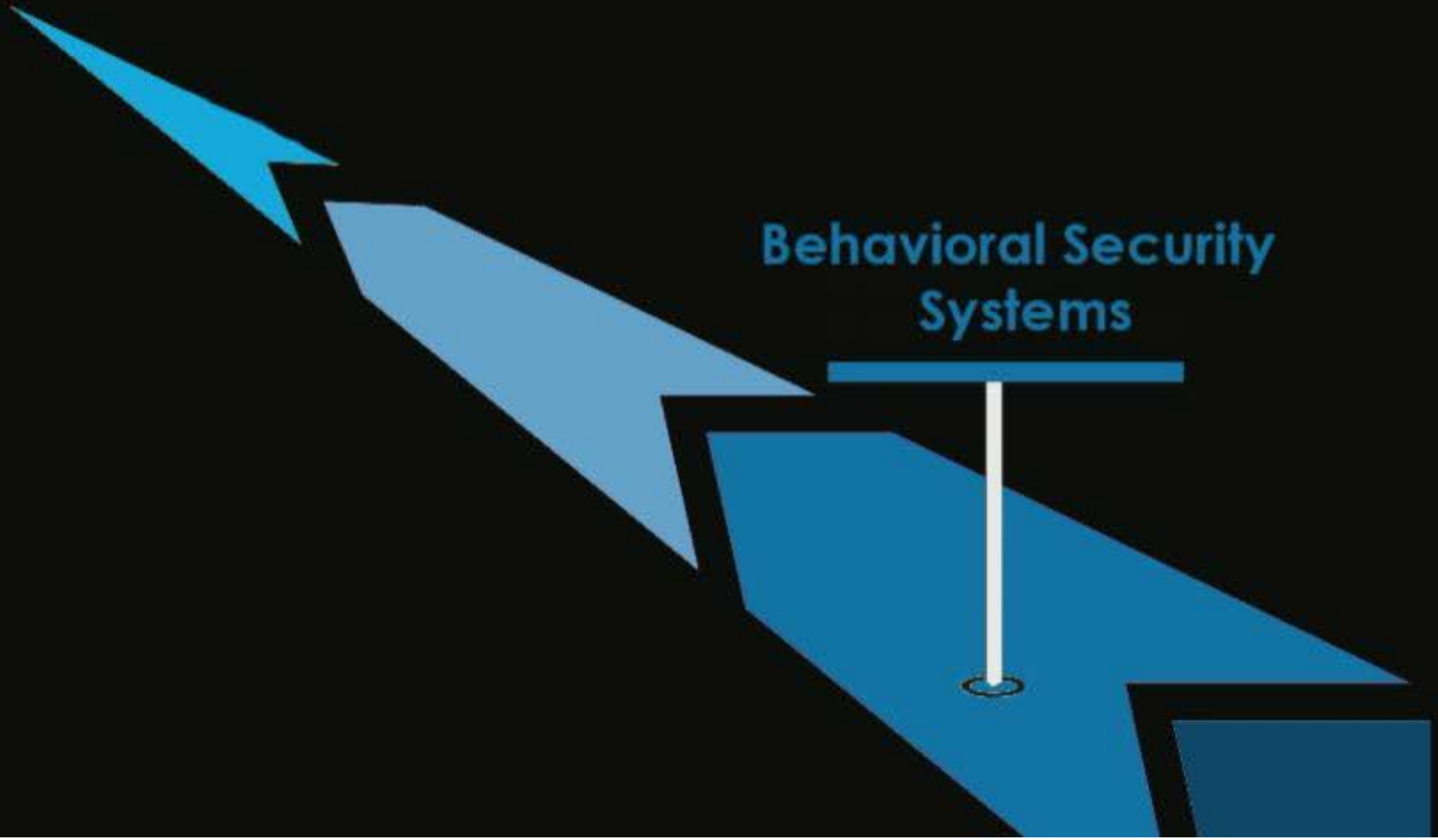
S&P2019 generalizability of Mturk & webpanels vs. probabilistic samples

EC2018 Open-source, scalable platform for behavioral security experiments

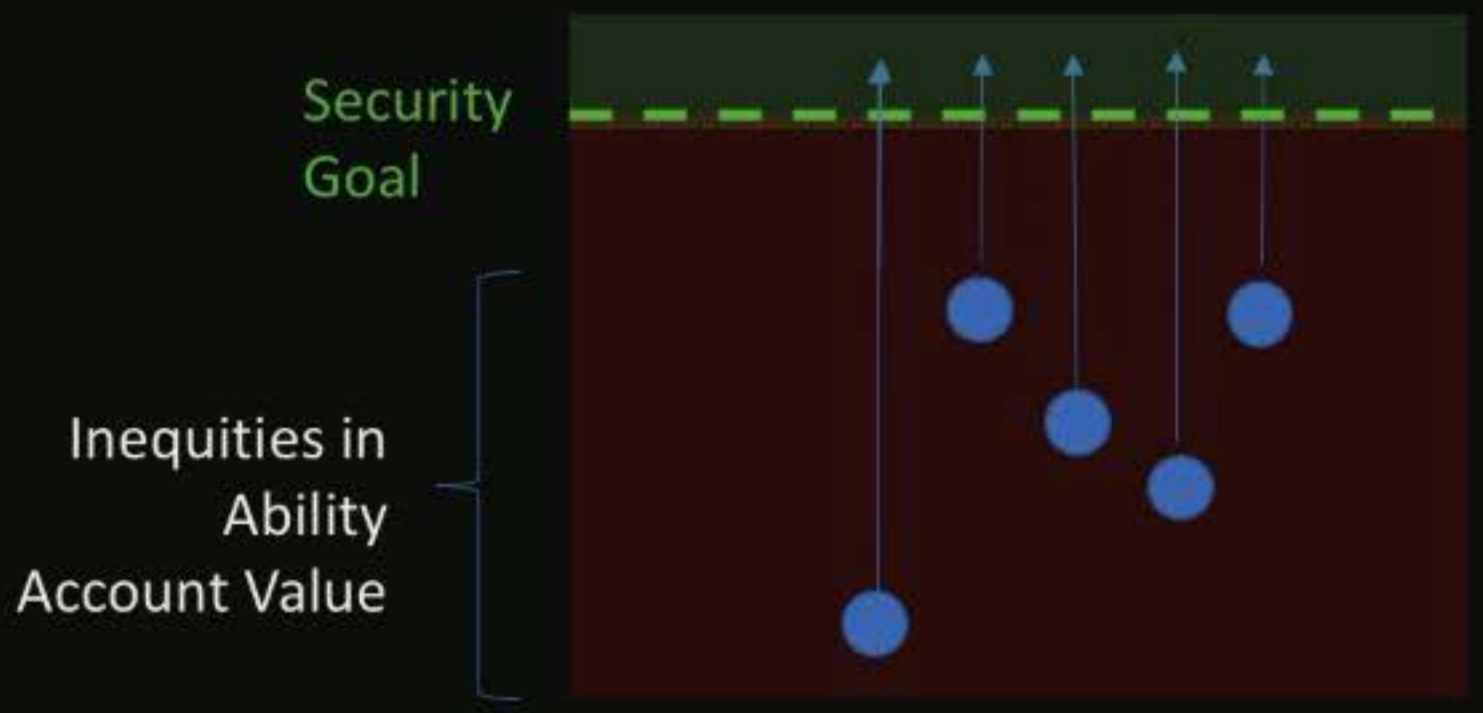
Future Work

What's Next?

Moving from understanding to behavioral security systems



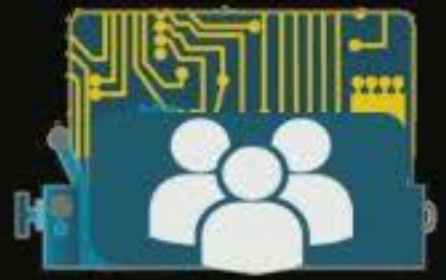
Incorporate human understanding in security systems

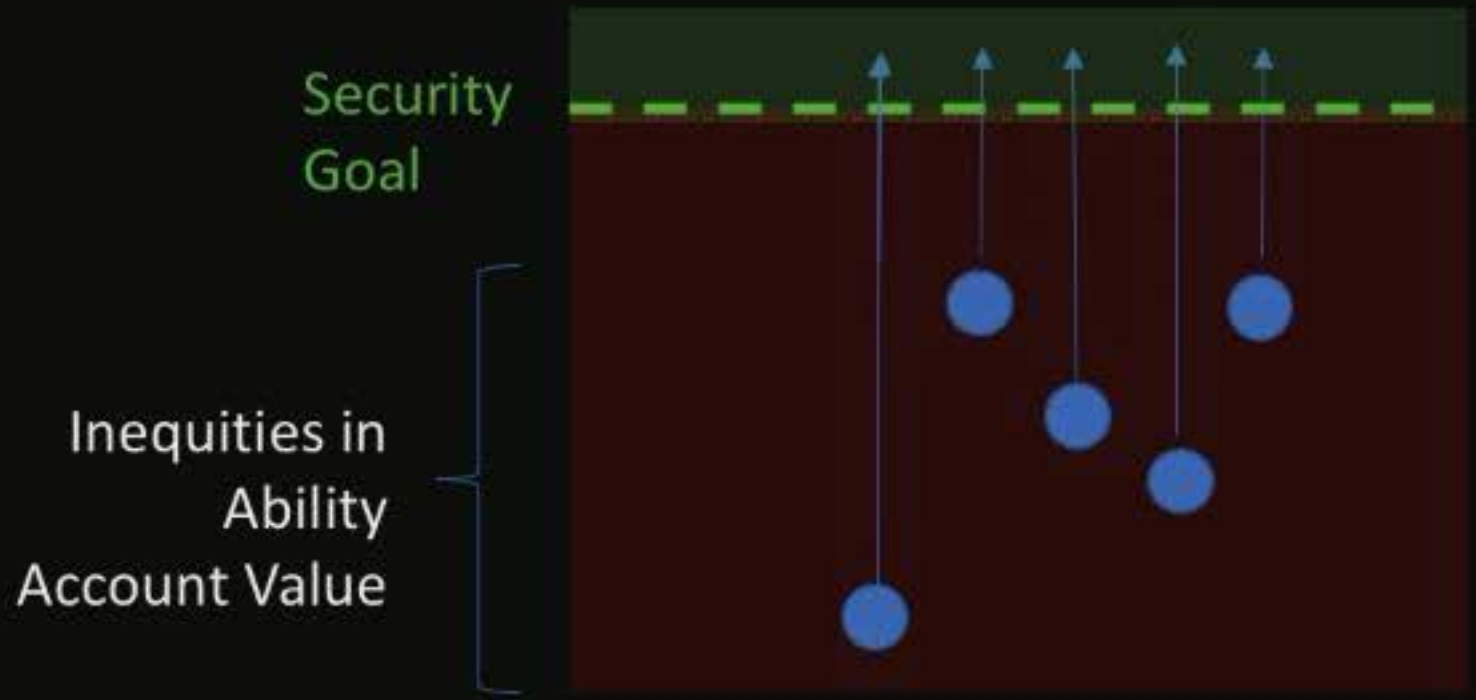


Mechanism design to optimize equitable security policies

Machine teaching security skills (e.g., password creation)

Incorporate human understanding in security systems

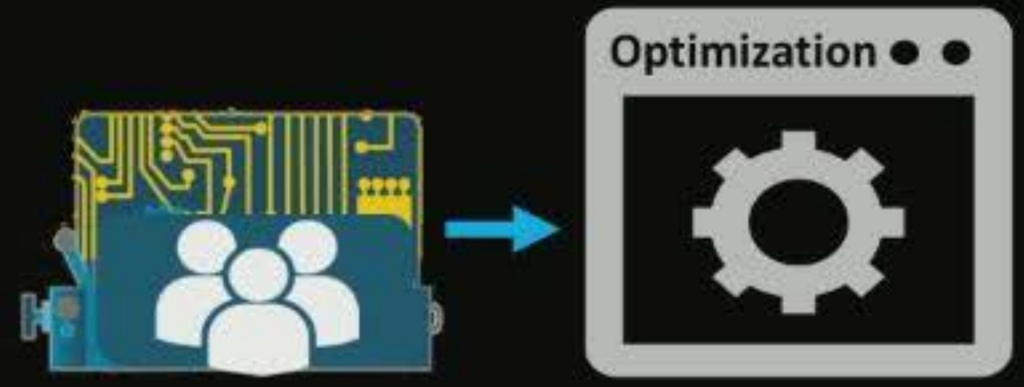


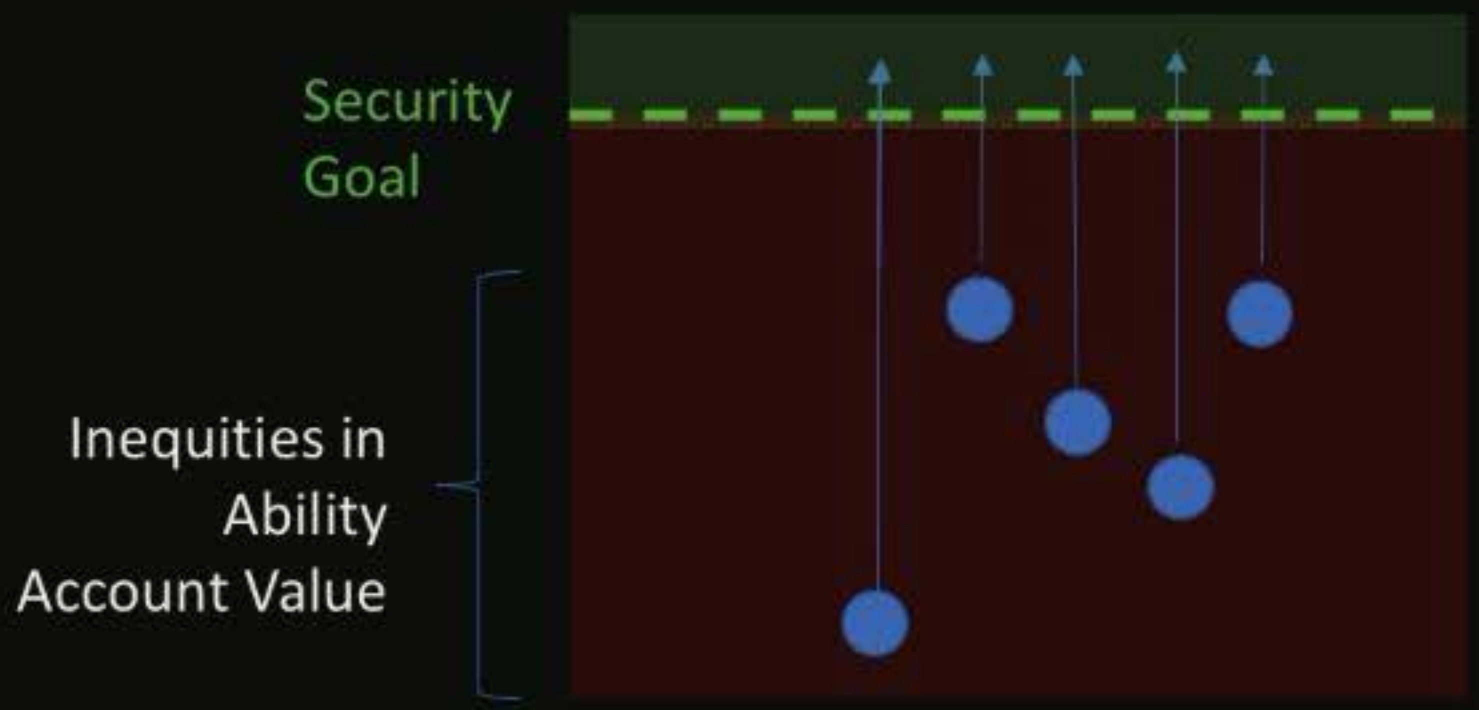


Mechanism design to optimize equitable security policies

Machine teaching security skills (e.g., password creation)

Incorporate human understanding in security systems

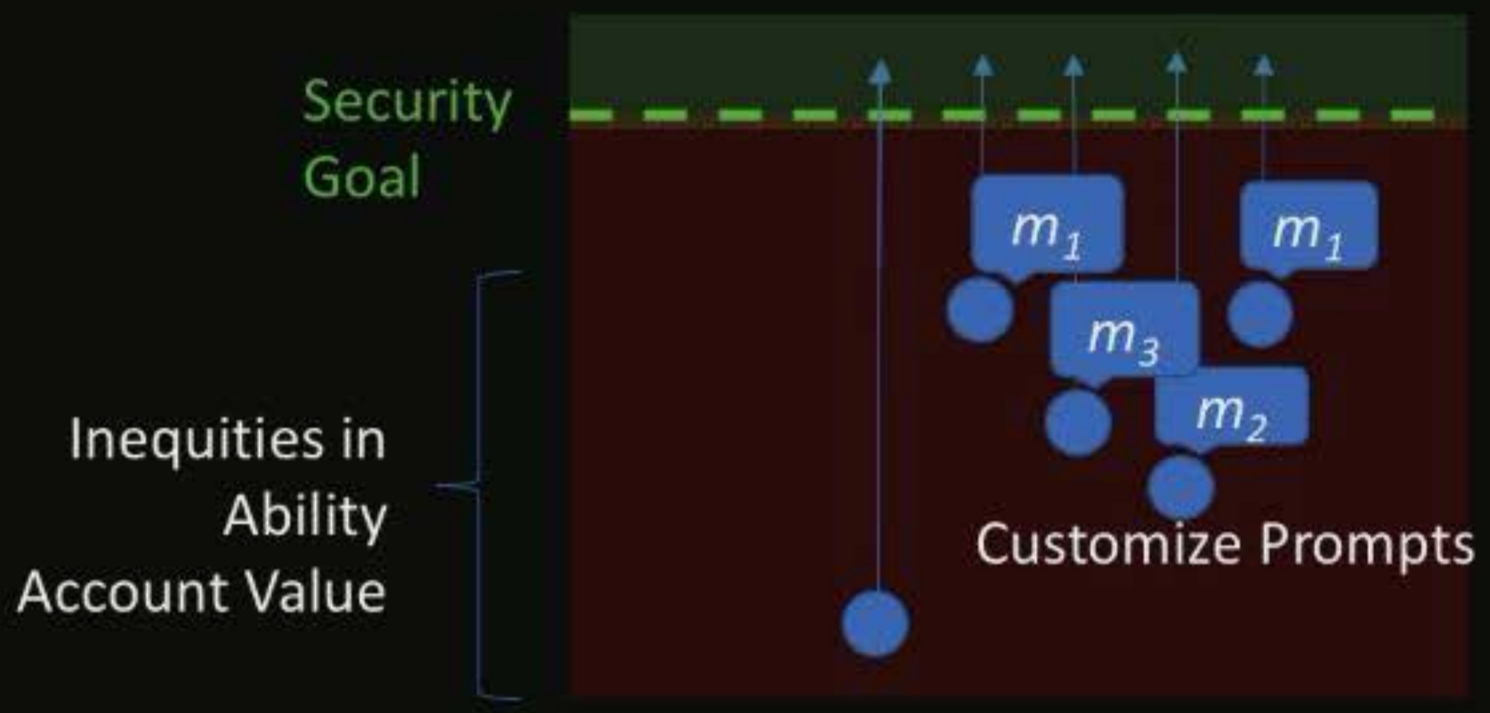




Mechanism design to optimize equitable security policies

Machine teaching security skills (e.g., password creation)

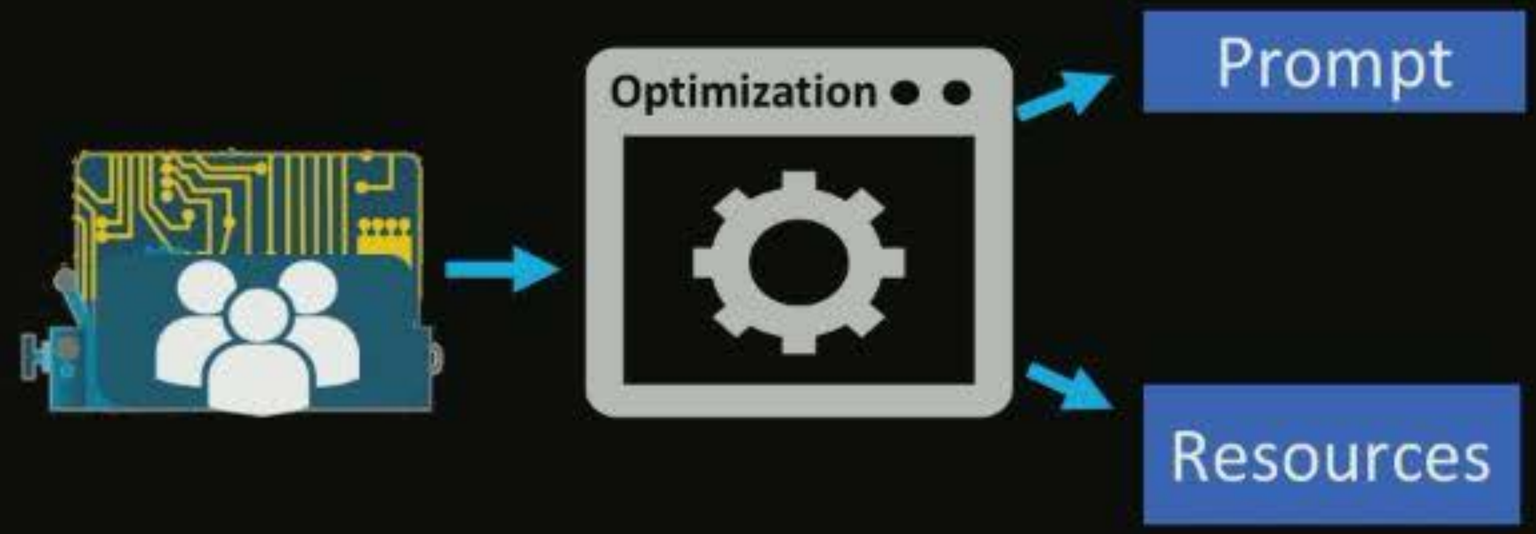
Incorporate human understanding in security systems

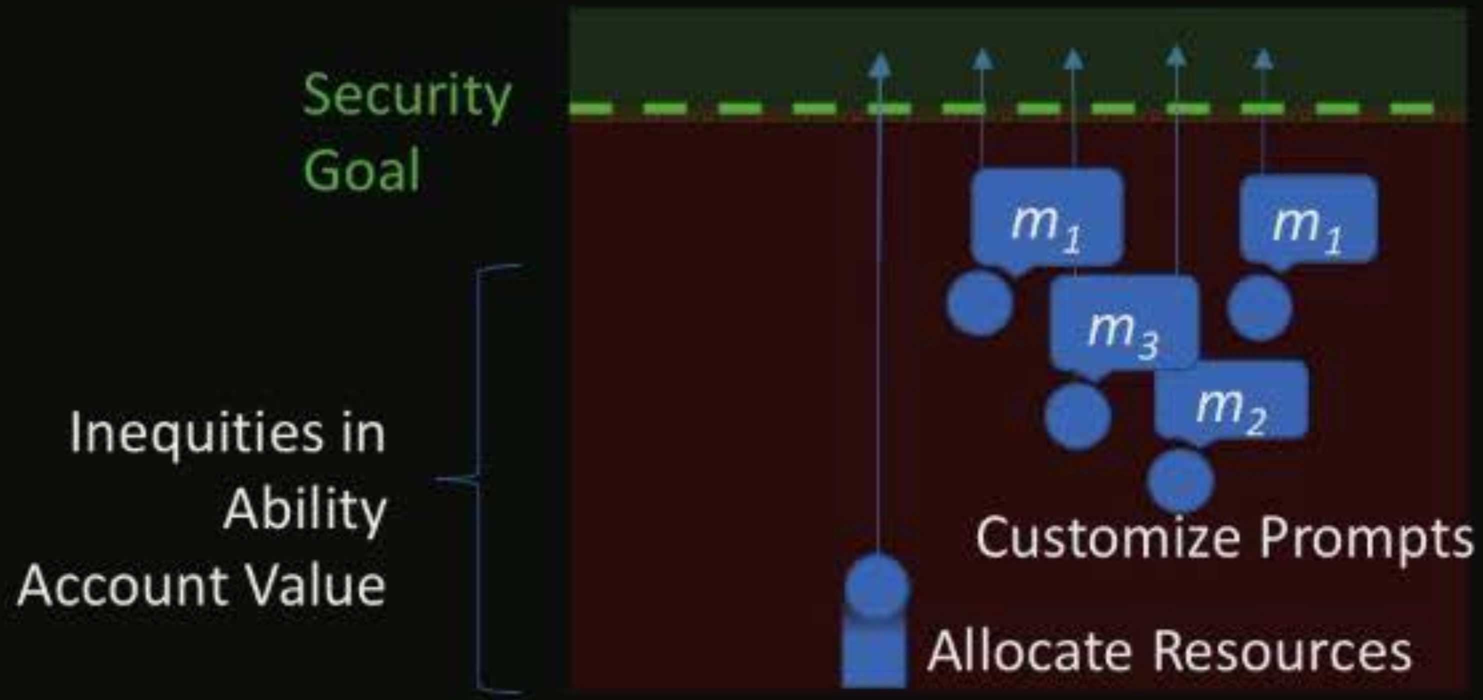


Mechanism design to optimize equitable security policies

Machine teaching security skills (e.g., password creation)

Incorporate human understanding in security systems

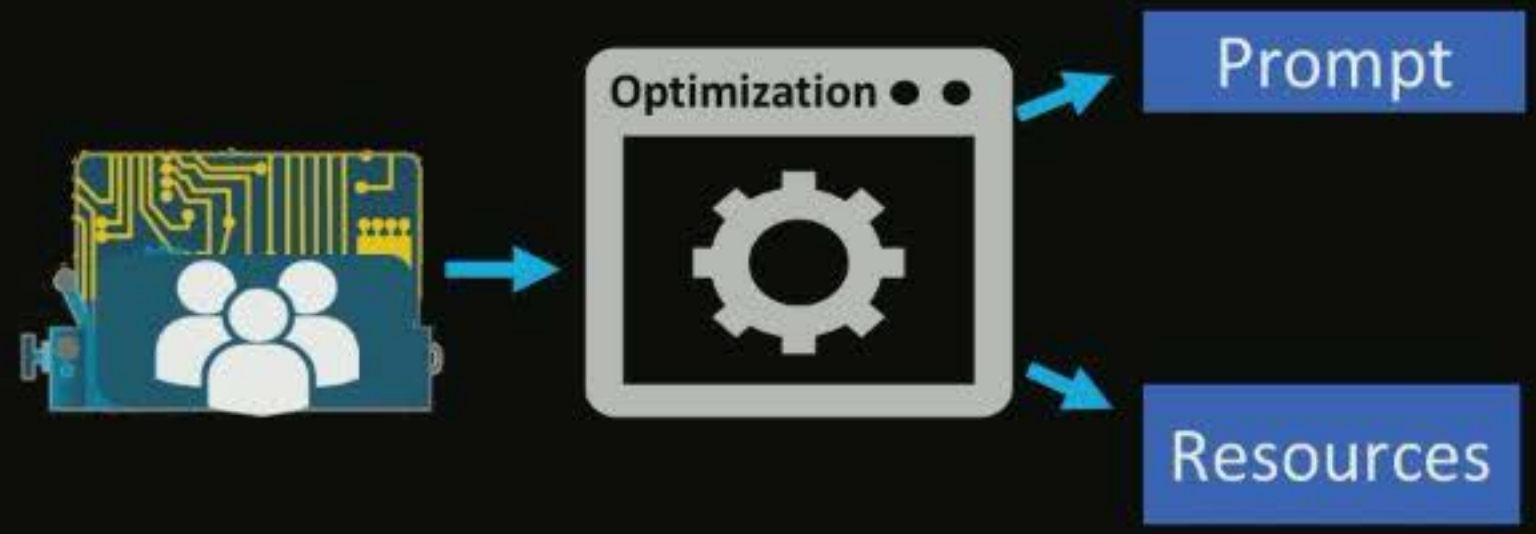


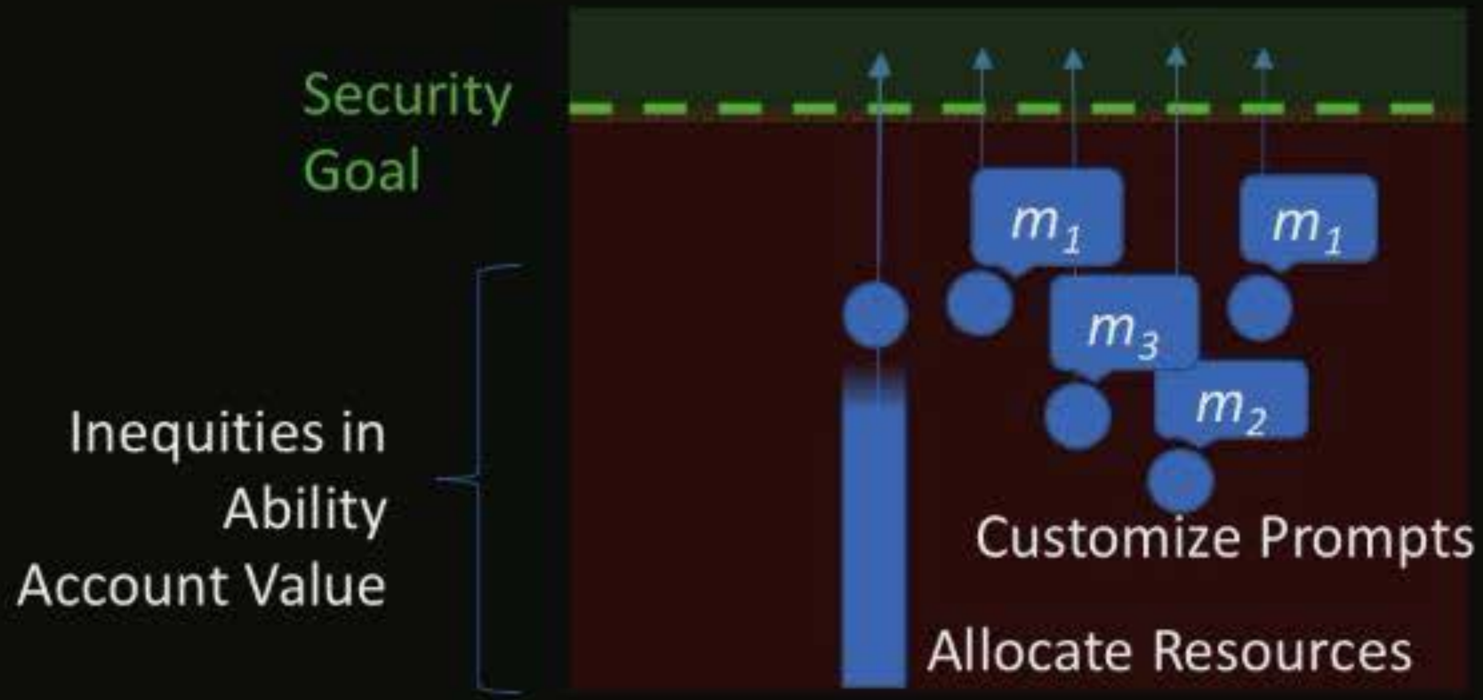


Mechanism design to optimize equitable security policies

Machine teaching security skills (e.g., password creation)

Incorporate human understanding in security systems





Mechanism design to optimize equitable security policies

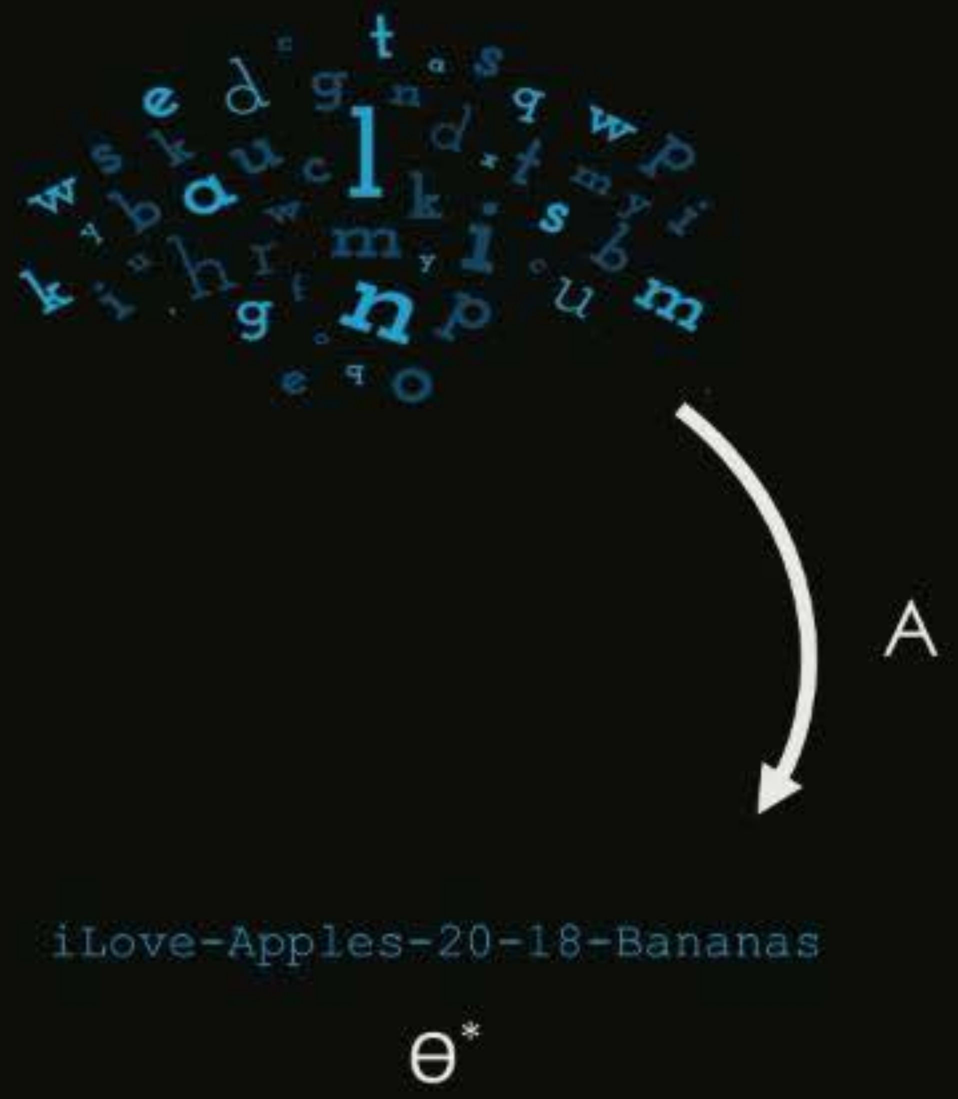
Machine teaching security skills (e.g., password creation)

Incorporate human understanding in security systems

Mechanism design to optimize equitable security policies

Machine teaching security skills (e.g., password creation)

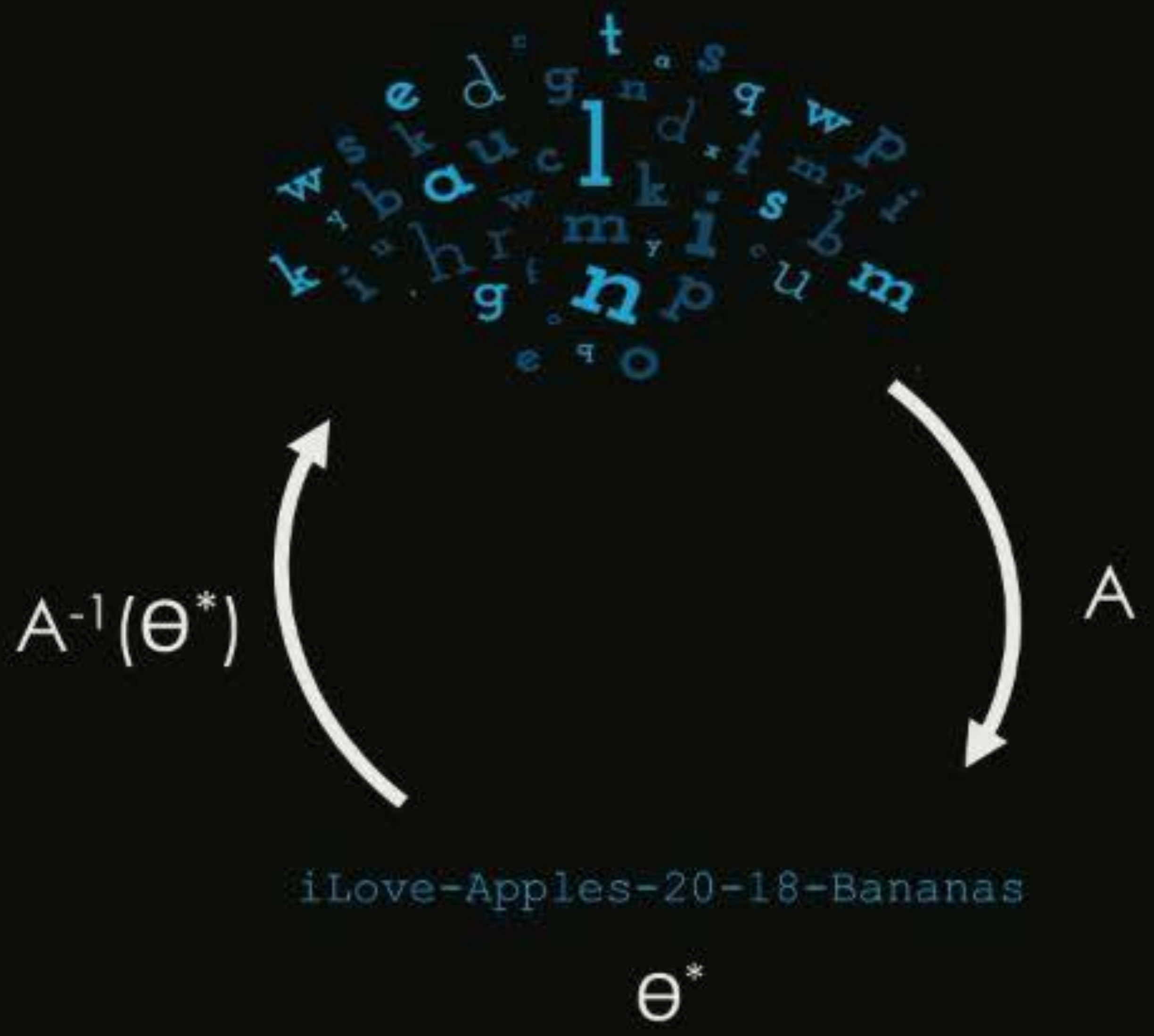
Incorporate human understanding in security systems



Mechanism design to optimize equitable security policies

Machine teaching security skills (e.g., password creation)

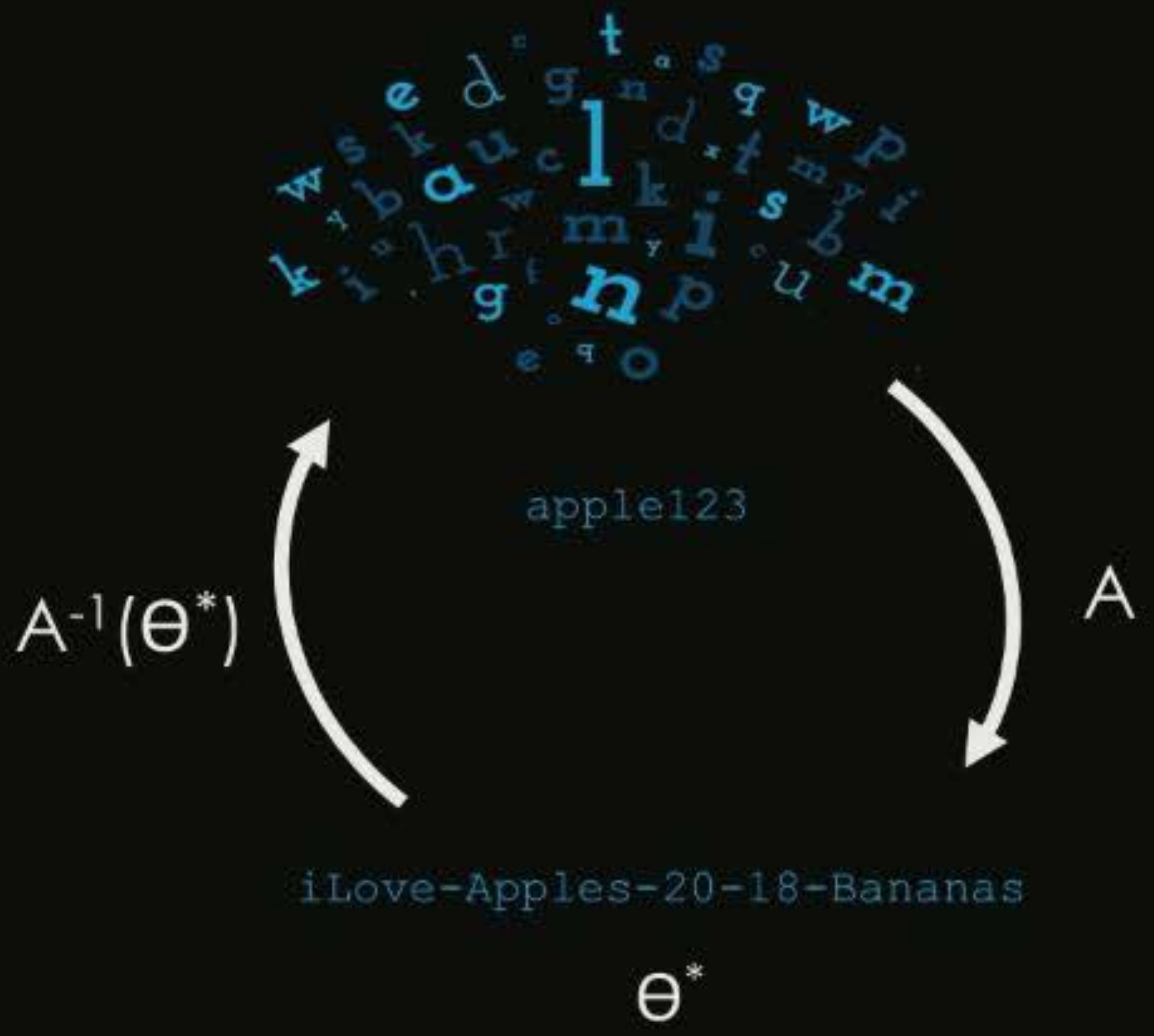
Incorporate human understanding in security systems



Mechanism design to optimize equitable security policies

Machine teaching security skills (e.g., password creation)

Incorporate human understanding in security systems



Mechanism design to optimize equitable security policies

Machine teaching security skills (e.g., password creation)

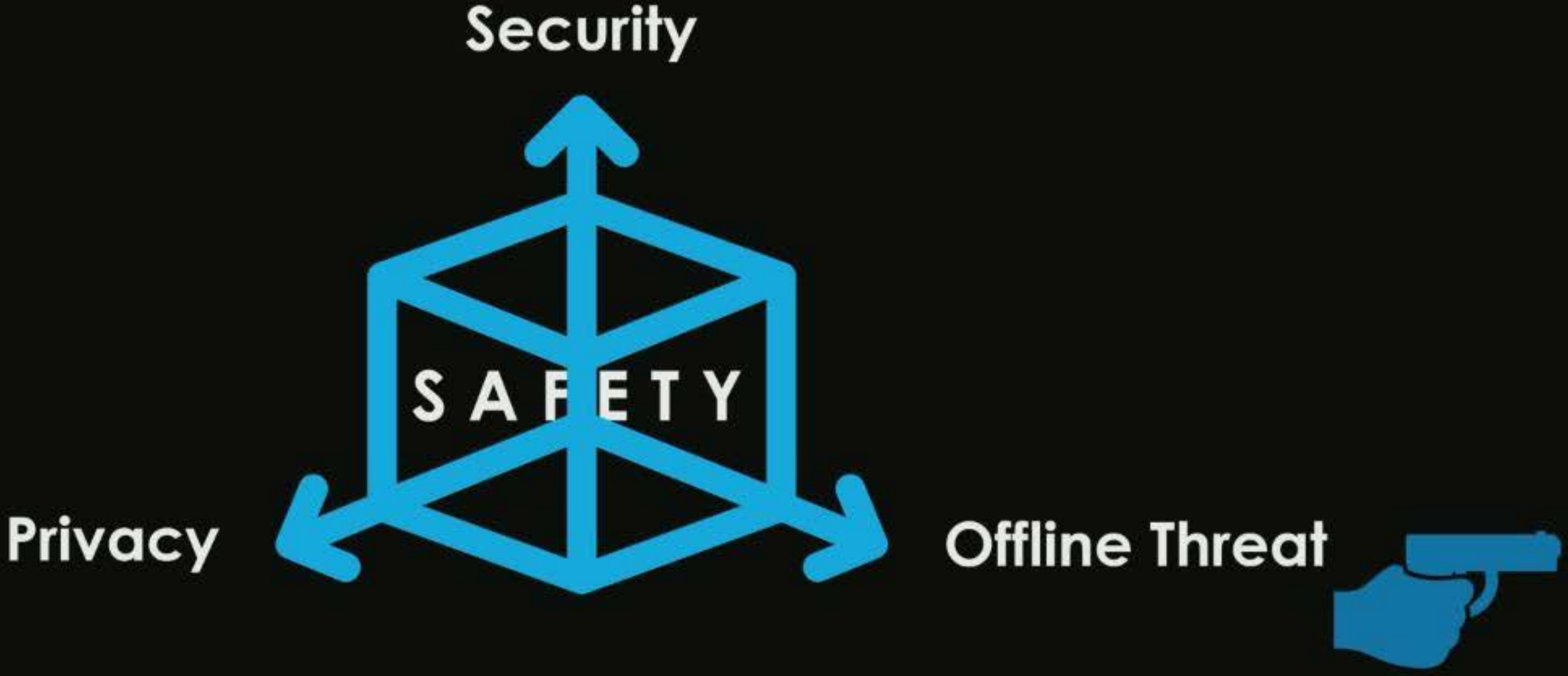
Expand modeling & inequity quantification beyond security

Users view online safety as a combination of security, privacy, and blurred offline / online threat

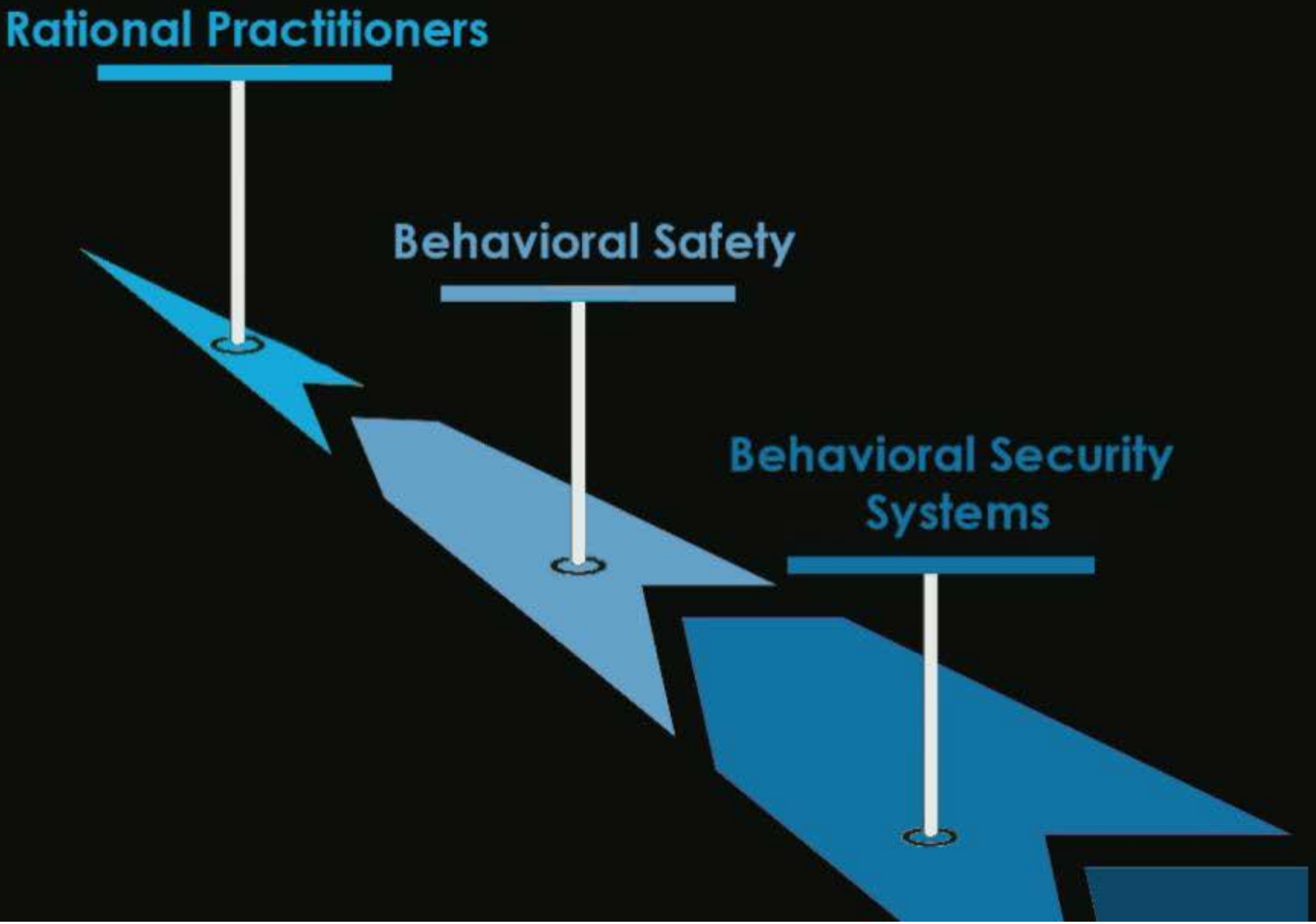
Redmiles, E.M., Bodford, J., and Blackwell, L. *"I just want to feel safe": A Diary Study of Safety Perceptions on Social Media.*
AAAI **International Conference on Web and Social Media (ICWSM)** 2019.

Users view online safety as a combination of security, privacy, and blurred offline / online threat

Users view online safety as a combination of security, privacy, and blurred offline / online threat



More rational security decisions by practitioners help users



Quantifying user harm & preference can help practitioners make more rational tradeoffs

Which Security Requirements to Set?

Quantifying user harm & preference can help practitioners make more rational tradeoffs

Which Security Requirements to Set?

Quantify impact of personalized job ads on income & job apps

Collaboration: Facebook

Quantifying user harm & preference can help practitioners make more rational tradeoffs

Which Security Requirements to Set?

Quantify impact of personalized job ads on income & job apps

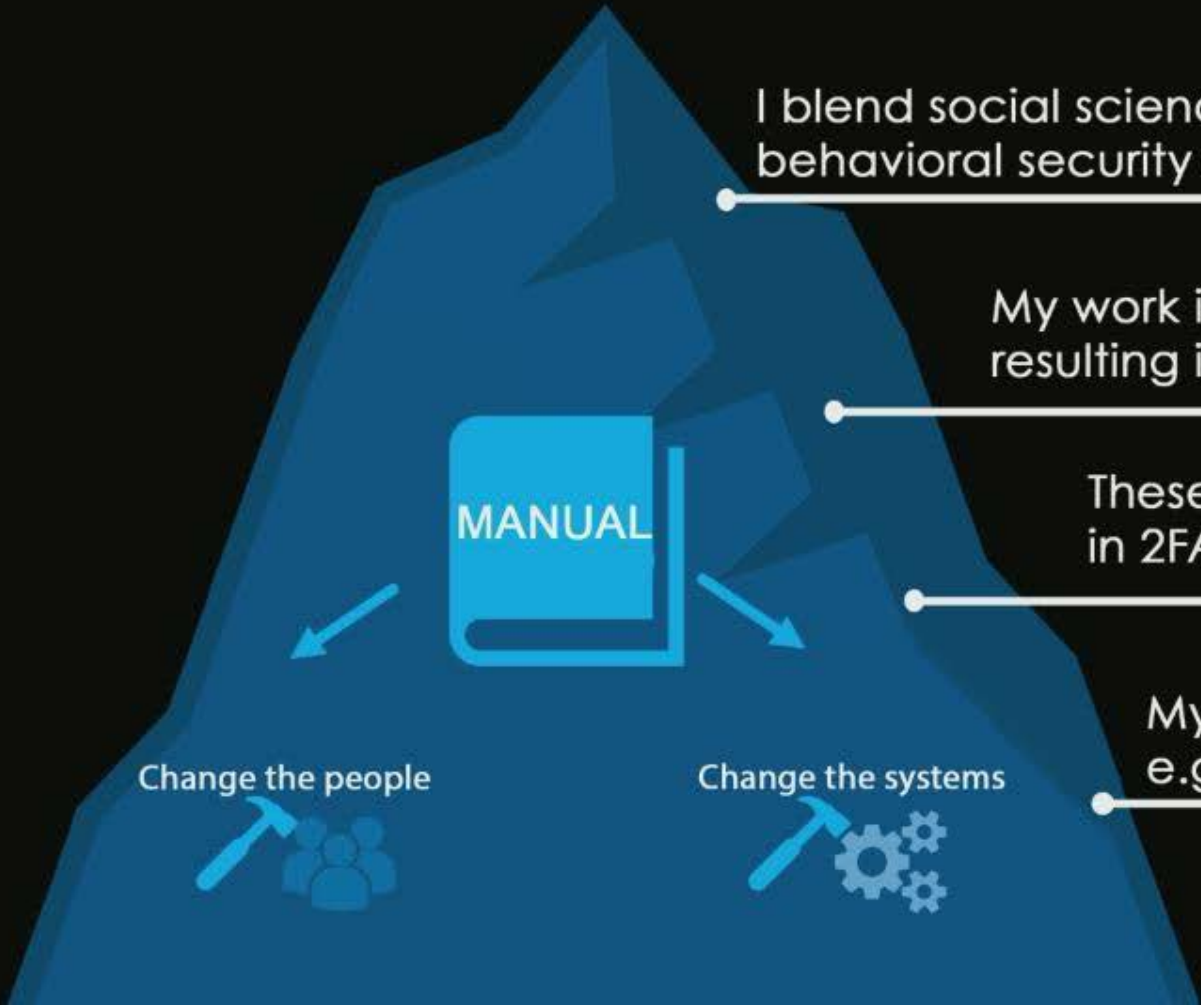
Collaboration: Facebook

$$\Pr[A(D_1) \in S] \leq e^{\varepsilon} \times \Pr[A(D_2) \in S]$$

Inform more computationally efficient ε based on people's information revealing behavior

Collaboration: Georgia Institute of Technology

My work modeling structural inequities enables the design of systems that are secure for all users



I blend social science, economics & ML methods to construct behavioral security models & examine structural security inequities

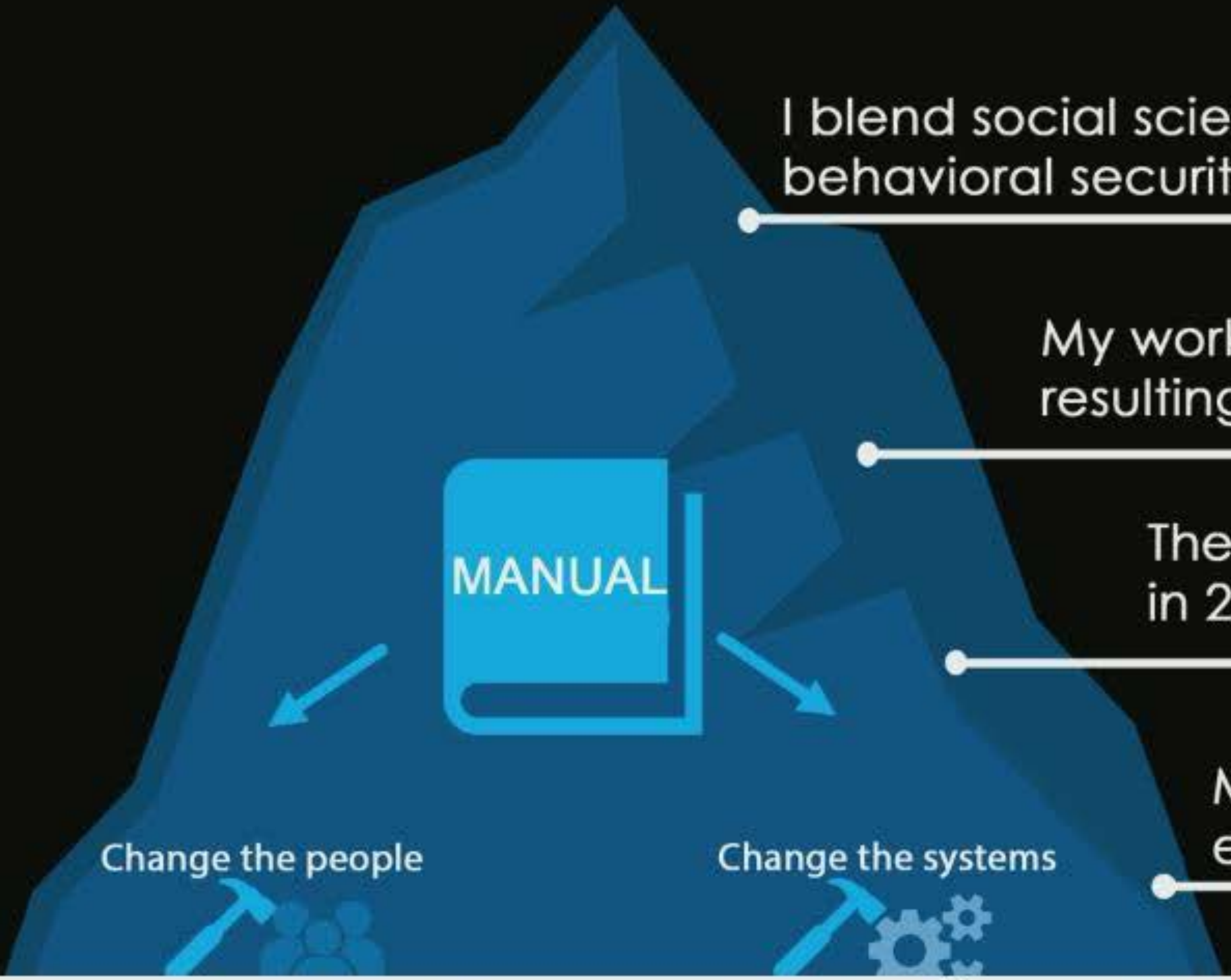
My work identified early evidence of security inequity resulting in policy discussion with the FTC, US CERT & NSF

These models have also driven real-world changes in 2FA, suspicious login & spam systems at

My modeling approaches apply beyond security e.g., to improve fair feature selection (WWW18)

My work modeling structural inequities enables the design of systems that are secure for all users

Elissa M. Redmiles
eredmiles@cs.umd.edu



I blend social science, economics & ML methods to construct behavioral security models & examine structural security inequities

My work identified early evidence of security inequity resulting in policy discussion with the FTC, US CERT & NSF

These models have also driven real-world changes in 2FA, suspicious login & spam systems at facebook

My modeling approaches apply beyond security e.g., to improve fair feature selection (WWW18)

University of Zurich UZH | Graduate Research Fellowship Program | HCIL | University of Maryland | MARYLAND CYBERSECURITY CENTER | facebook | MAX PLANCK INSTITUTE FOR SOFTWARE SYSTEMS

Requiring security can be costly: 2FA code fees + engagement losses

Value of accounts to users

Approach	User Costs	2FA Benefit	Loss/Gain	Market Impact 500K MTurk Users
2FA Required	$275 per 1000 MTurkers	$148 per 1000 MTurkers	(-) $126 per 1000 MTurkers	(-) $63,606
Perfect Rationality	$32 per 1000 MTurkers	$128 per 1000 MTurkers	(+) $96 per 1000 MTurkers	(+) $47,865
No 2FA Offered	$266 per 1000 MTurkers	$0 per 1000 MTurkers	(-) $266 per 1000 MTurkers	(-) $133,000

Redmiles, E.M., Mazurek, M.L., and Dickerson, J.P. *Dancing Pigs or Externalities? Measuring the Rationality of Security Decisions.* ACM **EC** 2018.

CCS18: When to use survey vs. log data

Research Question

How well do survey and log data align for questions regarding user security behavior?

Methods

Compare log (n=517,932) and survey (n=2,092) data about software updating

Findings

Surveys approximate general not detailed constructs

Take Aways

Use surveys for perceptions & broad reactions
Try filtering non-sensical responses
Use observation for assessing detailed variations

Redmiles, E.M., Zhu, Z., Kross, S., Kuchhal, D., Dumitras, T.., and Mazurek M.L.. *Asking for a Friend: Evaluating Response Biases in Security User Studies.* ACM **CCS** 2018.

CCS18: Carefully designed survey & selected test cases

Imagine that you see this message appear on your computer.
Would you install the update?

- Yes, the first time I saw this message.
- Yes, within a week of seeing this message.
- Yes, within a few weeks of seeing this message.
- Yes, within a few months of seeing this message.
- No.
- I don't know.

Detailed	Application
	Update Cost
	Security-Only
	Message Length
General	Update Risk
	Tendency to Update

Redmiles, E.M., Zhu, Z., Kross, S., Kuchhal, D., Dumitras, T.., and Mazurek M.L.. *Asking for a Friend: Evaluating Response Biases in Security User Studies.* ACM **CCS** 2018.