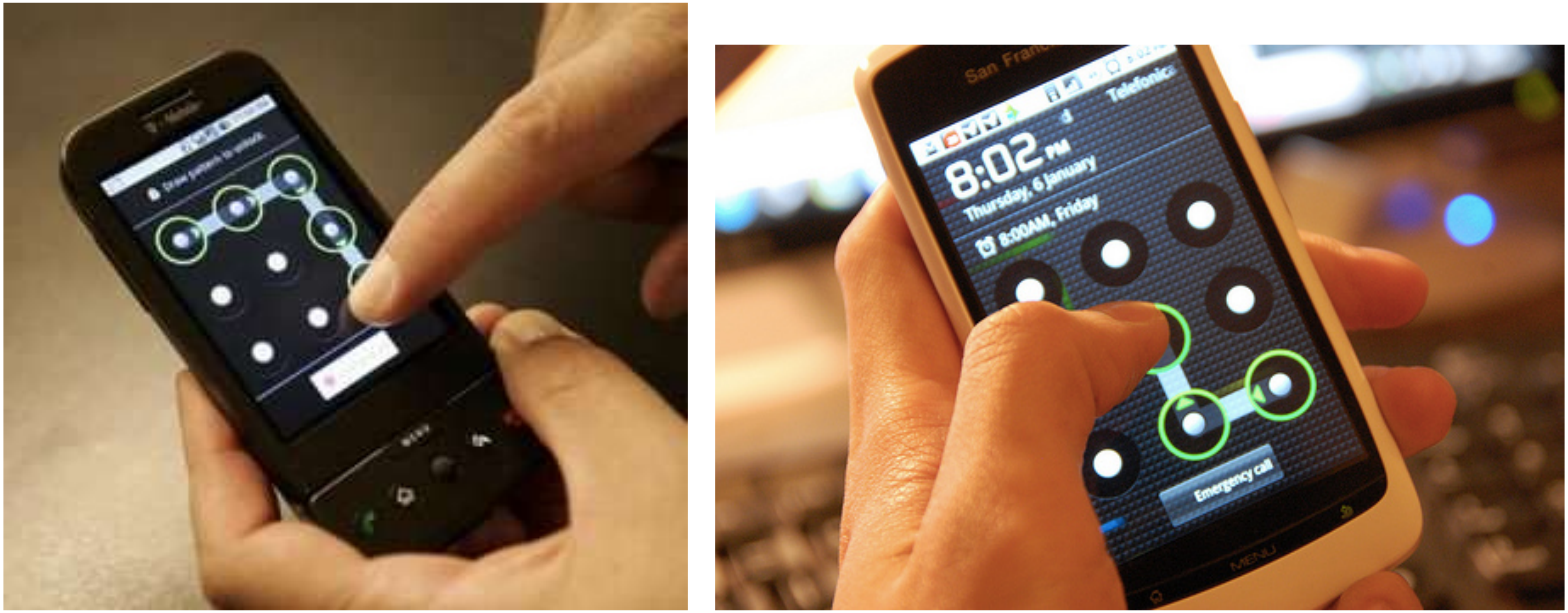
Analyzing the Impact of Collection Methods and Demographics for Android's Pattern Unlock

Adam J. Aviv Justin Maguire Jeanne Luning-Prak

Android's Pattern Unlock

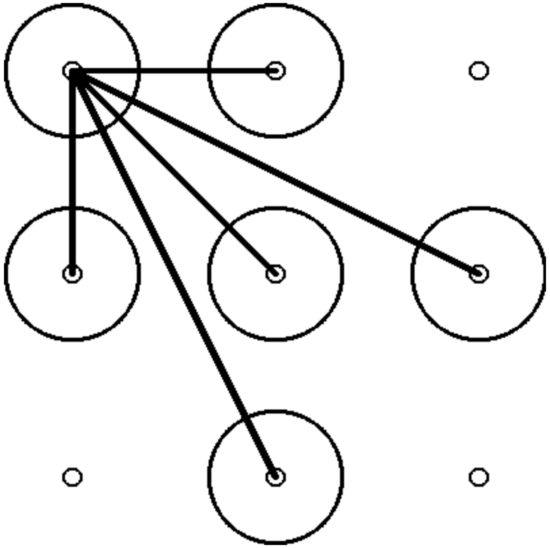


Android pattern unlock is an authentication method to lock (and unlock) Android phones

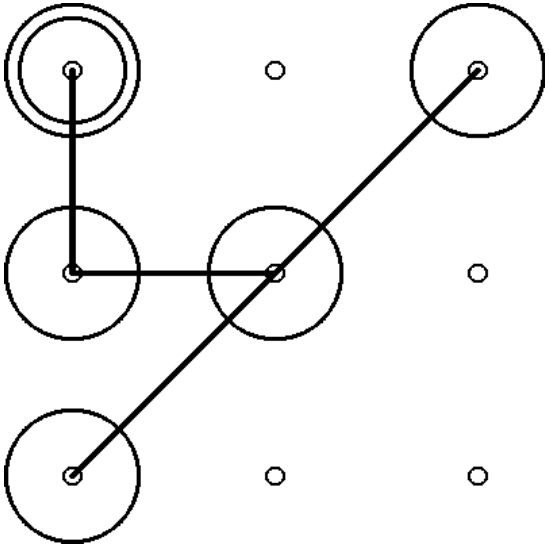
Rules of the Game

(1)
Maintain contact with the screen and connect ***4*** *points* ***without repetition***

(2)
Can only connect adjacent contact points

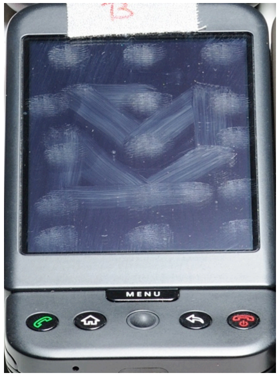


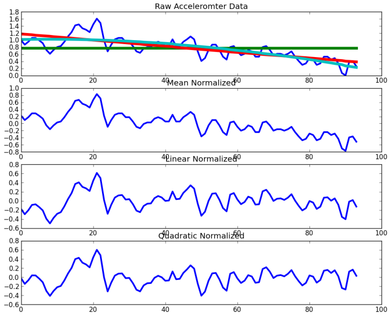
(3)
Can trace over previously contacted points



Recent Work on Android Unlock Patterns

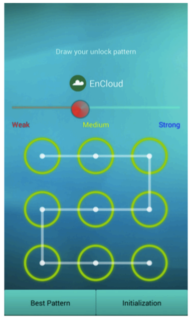
Attacks

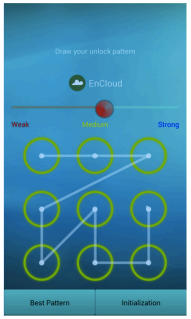




[_GMBS:WOOT’10] [_SMS:ACASC’12]

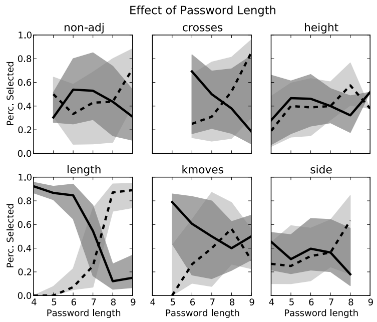
Meters





[SCOKH:CHI’15]

Perceptions



[_D:ACASC’14]

Measurement

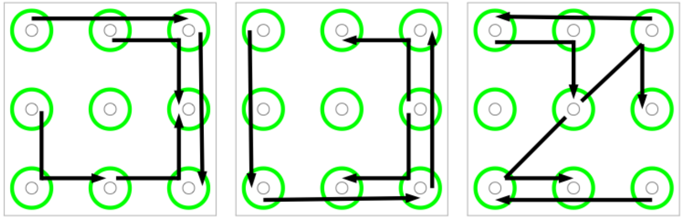
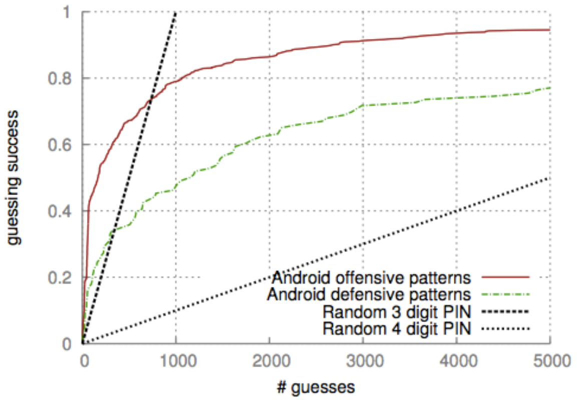
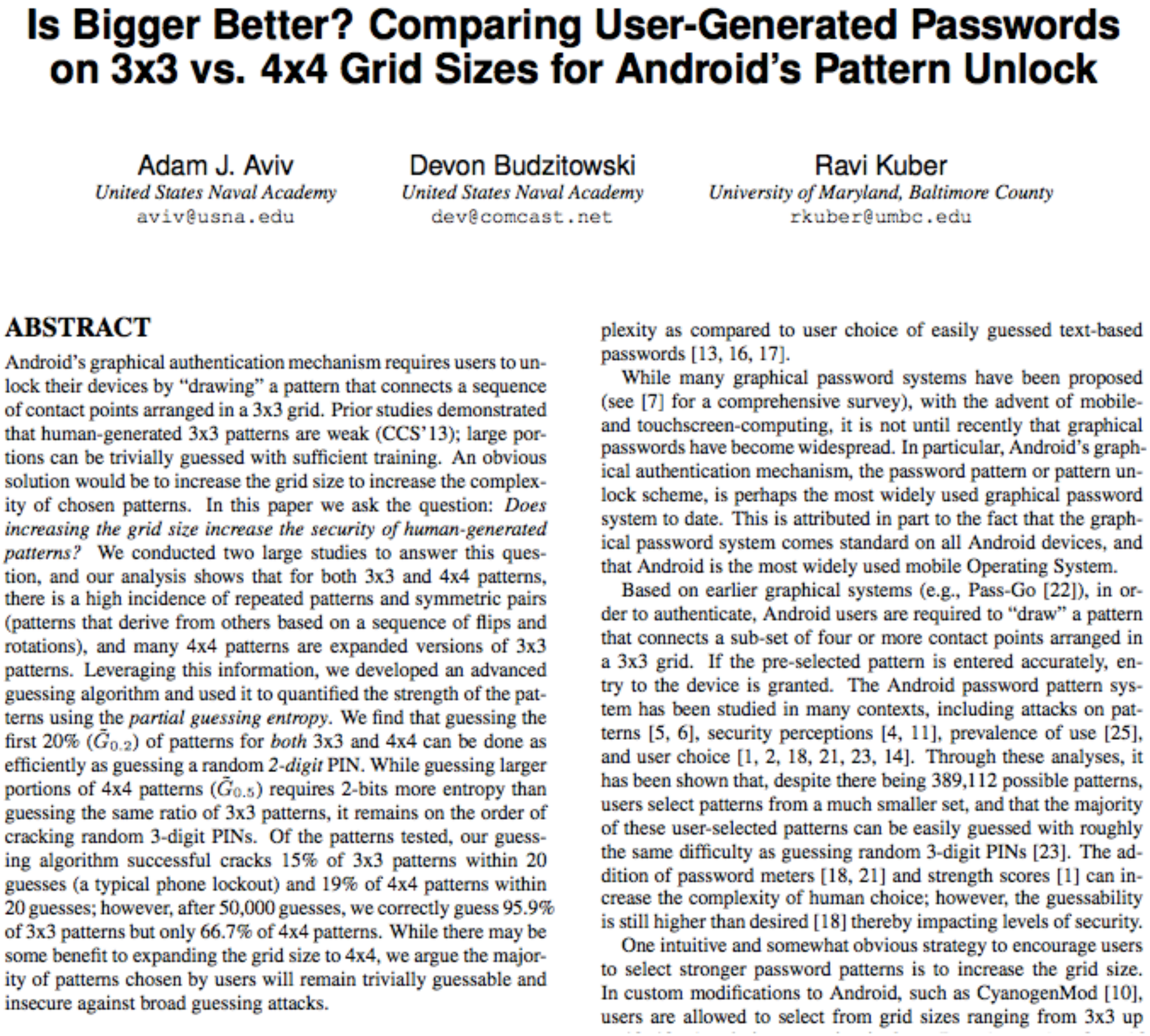


Figure 8: The most frequent 3-grams, from most frequent (left) to less frequent (right).



[UDWH:CCS’13]

[_BK:ACSAC'15]



Is Bigger Better? Comparing User-Generated Passwords on 3x3 vs. 4x4 Grid Sizes for Android's Pattern Unlock

Adam J. Aviv
United States Naval Academy
aviv@usna.edu

Devon Budzitowski
United States Naval Academy
dev@comcast.net

Ravi Kuber
University of Maryland, Baltimore County
rkuber@umbc.edu

ABSTRACT

Android's graphical authentication mechanism requires users to unlock their devices by "drawing" a pattern that connects a sequence of contact points arranged in a 3x3 grid. Prior studies demonstrated that human-generated 3x3 patterns are weak (CCS'13); large portions can be trivially guessed with sufficient training. An obvious solution would be to increase the grid size to increase the complexity of chosen patterns. In this paper we ask the question: *Does increasing the grid size increase the security of human-generated patterns?* We conducted two large studies to answer this question, and our analysis shows that for both 3x3 and 4x4 patterns, there is a high incidence of repeated patterns and symmetric pairs (patterns that derive from others based on a sequence of flips and rotations), and many 4x4 patterns are expanded versions of 3x3 patterns. Leveraging this information, we developed an advanced guessing algorithm and used it to quantified the strength of the patterns using the *partial guessing entropy*. We find that guessing the first 20% ($\tilde{G}_{0.2}$) of patterns for *both* 3x3 and 4x4 can be done as efficiently as guessing a random *2-digit* PIN. While guessing larger portions of 4x4 patterns ($\tilde{G}_{0.5}$) requires 2-bits more entropy than guessing the same ratio of 3x3 patterns, it remains on the order of cracking random 3-digit PINs. Of the patterns tested, our guessing algorithm successful cracks 15% of 3x3 patterns within 20 guesses (a typical phone lockout) and 19% of 4x4 patterns within 20 guesses; however, after 50,000 guesses, we correctly guess 95.9% of 3x3 patterns but only 66.7% of 4x4 patterns. While there may be some benefit to expanding the grid size to 4x4, we argue the majority of patterns chosen by users will remain trivially guessable and insecure against broad guessing attacks.

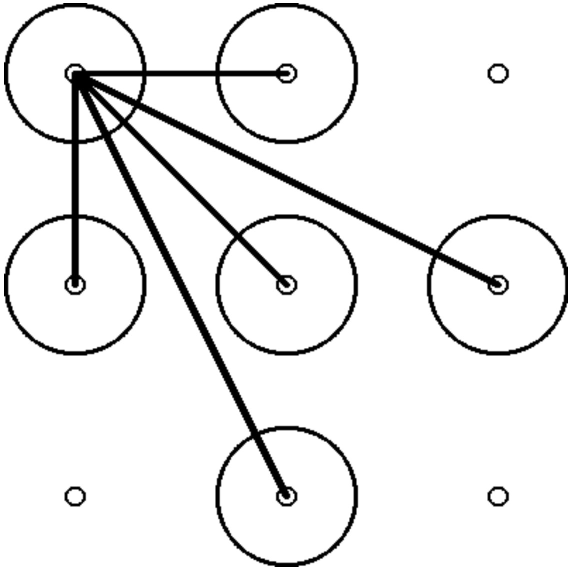
plexity as compared to user choice of easily guessed text-based passwords [13, 16, 17].

While many graphical password systems have been proposed (see [7] for a comprehensive survey), with the advent of mobile- and touchscreen-computing, it is not until recently that graphical passwords have become widespread. In particular, Android's graphical authentication mechanism, the password pattern or pattern unlock scheme, is perhaps the most widely used graphical password system to date. This is attributed in part to the fact that the graphical password system comes standard on all Android devices, and that Android is the most widely used mobile Operating System.

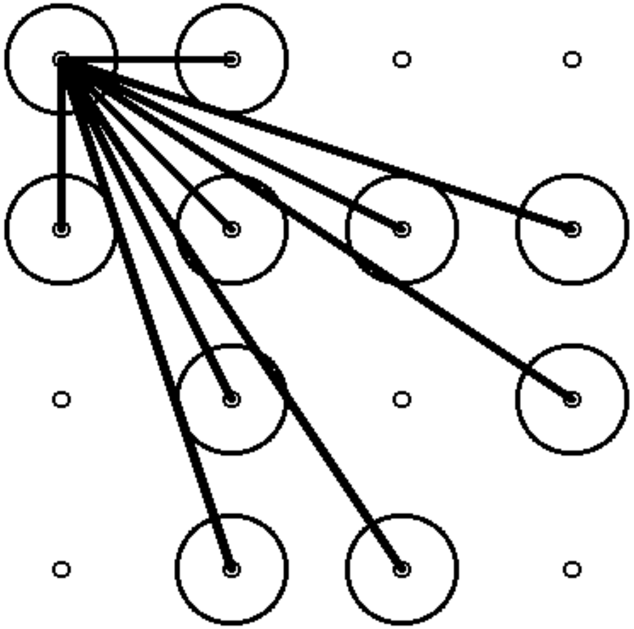
Based on earlier graphical systems (e.g., Pass-Go [22]), in order to authenticate, Android users are required to "draw" a pattern that connects a sub-set of four or more contact points arranged in a 3x3 grid. If the pre-selected pattern is entered accurately, entry to the device is granted. The Android password pattern system has been studied in many contexts, including attacks on patterns [5, 6], security perceptions [4, 11], prevalence of use [25], and user choice [1, 2, 18, 21, 23, 14]. Through these analyses, it has been shown that, despite there being 389,112 possible patterns, users select patterns from a much smaller set, and that the majority of these user-selected patterns can be easily guessed with roughly the same difficulty as guessing random 3-digit PINs [23]. The addition of password meters [18, 21] and strength scores [1] can increase the complexity of human choice; however, the guessability is still higher than desired [18] thereby impacting levels of security.

One intuitive and somewhat obvious strategy to encourage users to select stronger password patterns is to increase the grid size. In custom modifications to Android, such as CyanogenMod [10], users are allowed to select from grid sizes ranging from 3x3 up

Does increasing the grid size increase the security of human generated patterns?



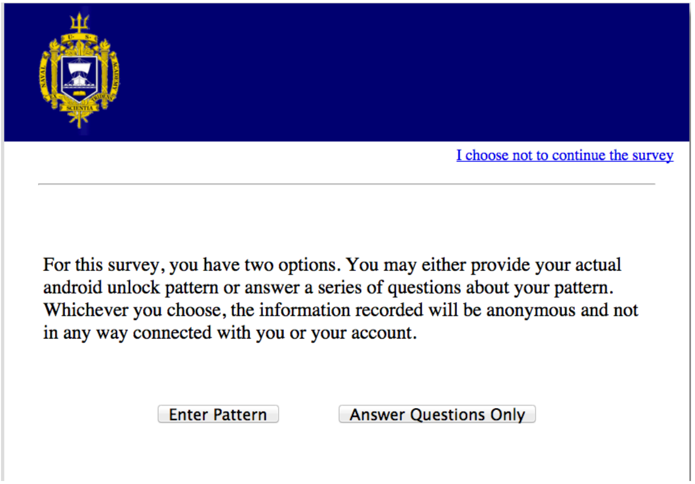
3x3 vs. 4x4



Methodology Challenges

<u>No good</u> datasets of 3x3 graphical passwords

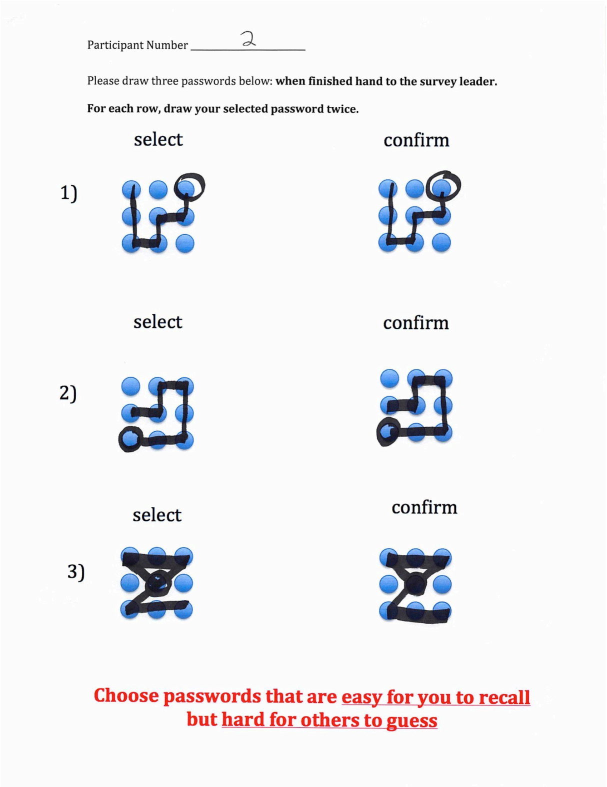
How to collect real 3x3 patterns?



For this survey, you have two options. You may either provide your actual android unlock pattern or answer a series of questions about your pattern. Whichever you choose, the information recorded will be anonymous and not in any way connected with you or your account.

4x4 patterns are <u>not</u> widely used

How to collect realistic 4x4 patterns?



Participant Number 2

Please draw three passwords below: **when finished hand to the survey leader.**

For each row, draw your selected password twice.

select
confirm
1)
select
confirm
2)
select
confirm
3)

Choose passwords that are <u>easy for you to recall</u> but <u>hard for others to guess</u>

Is 4x4 really better than 3x3?

- **NO:** The strength of *most* of the 4x4 patterns are similar to that of 3x3 patterns

- **YES:** The fraction of total guessed is less for 4x4 patterns

- **NO**: The fraction of most common 4x4 patterns are *more* guessable than the most common 3x3

- **YES:** User recall rates for 4x4 are the same for 3x3 but 4x4 patterns are less easily naïvely guessed

Demographic and Collection Methods

What are the statistical differences in the patterns based on the collection methods?

Are there differences in patterns based on demographics?

Can we leverage demographics to attack patterns?

Talk Outline

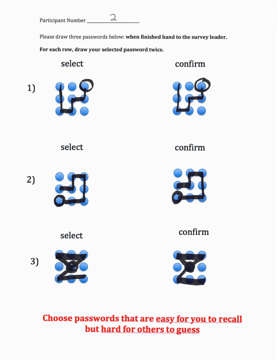
Research Questions

What are the statistical differences in the patterns based on the collection methods?

Are there differences in patterns based on demographics?

Can we leverage demographics to attack patterns?

Methodologies and Data Characteristics



Participant Number 2

Please draw three passwords below: **when finished hand to the survey leader.**

For each row, draw your selected password twice.

select confirm

1)

select confirm

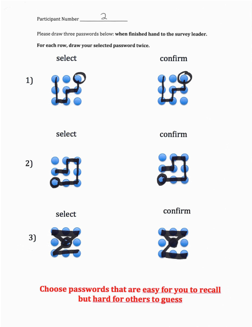
2)

select confirm

3)

Choose passwords that are easy for you to recall but hard for others to guess

Comparisons of Collection Methodologies and Demographics



Participant Number 2

Please draw three passwords below: **when finished hand to the survey leader.**

For each row, draw your selected password twice.

select confirm

1)

select confirm

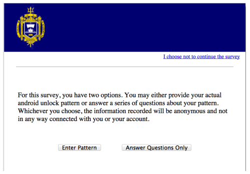
2)

select confirm

3)

Choose passwords that are easy for you to recall but hard for others to guess

Vs.



I choose not to continue the survey

For this survey, you have two options. You may either provide your actual android unlock pattern or answer a series of questions about your pattern. Whichever you choose, the information recorded will be anonymous and not in any way connected with you or your account.

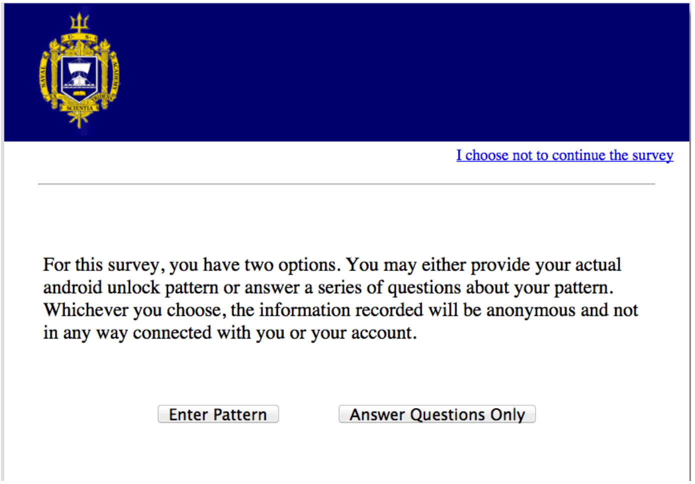
Demographics and Pattern Strength

Distribution	α =0.1	α =0.2	α =0.5	α =0.7
Pen-and-Paper	6.59	6.99	8.92	10.12
Self-Reporting	6.62	6.95	9.49	10.74
Men (SR)	5.86	7.00	9.59	10.60
Women (SR)	5.57	7.33	9.80	10.64
Urban (SR)	6.09	7.22	9.95	10.69
Suburban (SR)	6.58	8.70	9.58	10.60
Rural (SR)	6.08	8.22	9.60	10.42
Random 4-Digit PIN	13.28	13.28	13.28	13.28
Random 3-Digit PIN	9.97	9.97	9.97	9.97
Random 3x3 Pattern	18.57	18.57	18.57	18.57

METHODOLOGY

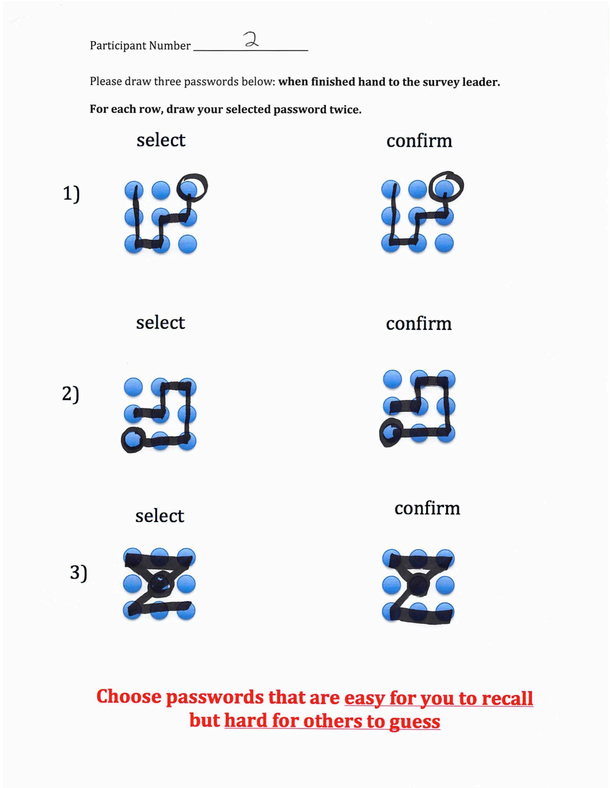
Methodologies

Self-Reported



For this survey, you have two options. You may either provide your actual android unlock pattern or answer a series of questions about your pattern. Whichever you choose, the information recorded will be anonymous and not in any way connected with you or your account.

Pen-and-Paper



Participant Number 2

Please draw three passwords below: **when finished hand to the survey leader.**

For each row, draw your selected password twice.

1) select confirm

2) select confirm

3) select confirm

Choose passwords that are <u>easy for you to recall</u> but <u>hard for others to guess</u>

All protocols were reviewed by the USNA and UMBC Institutional Review Board

Online Self-Report Survey

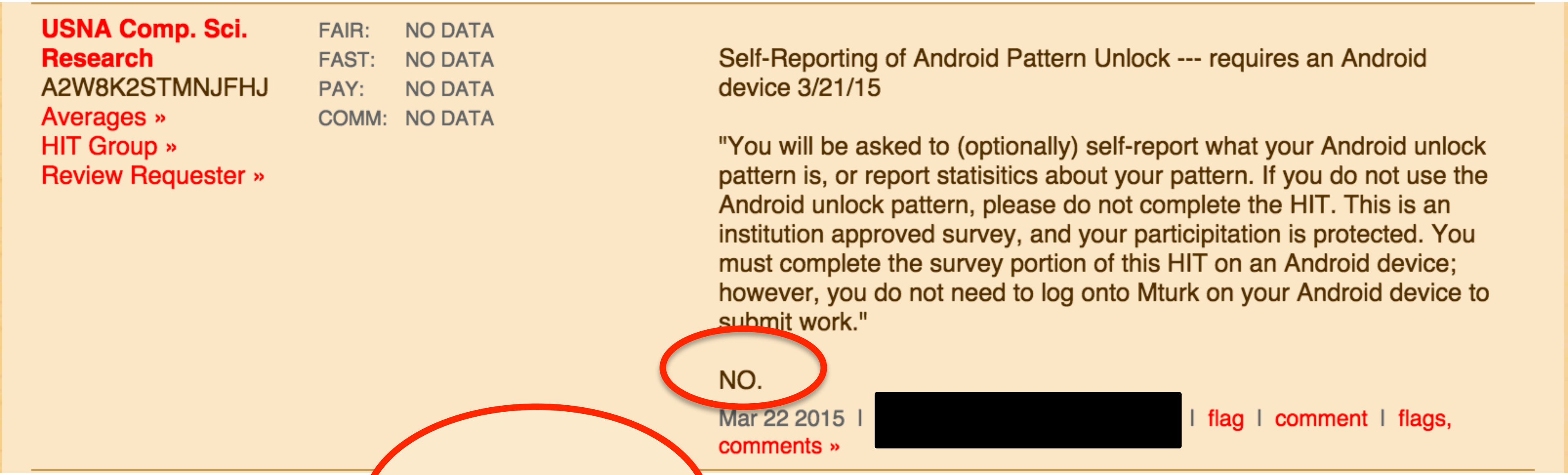
- *Pay people to self-report their pattern or* provide statistics about their pattern

- Amazon Mechanical Turk
  - Paid participants $0.50 or $0.75 (two runs)
  - 750 respondents data was included

- Must complete the survey on their mobile devices

Not for Everyone

Turk Opticom



USNA Comp. Sci. Research
A2W8K2STMNJFHJ
Averages »
HIT Group »
Review Requester »

FAIR:	NO DATA
FAST:	NO DATA
PAY:	NO DATA
COMM:	NO DATA

Self-Reporting of Android Pattern Unlock --- requires an Android device 3/21/15

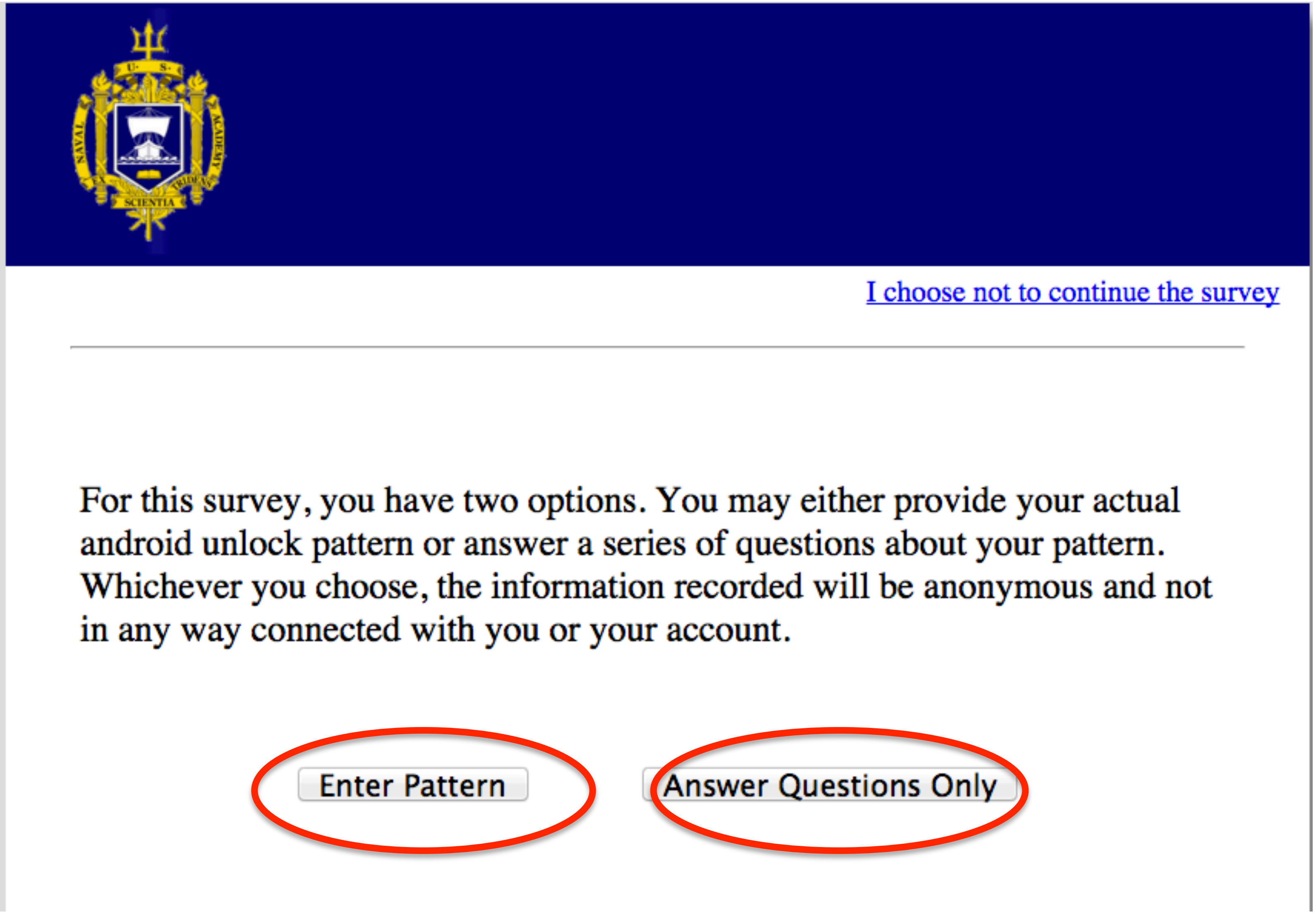
"You will be asked to (optionally) self-report what your Android unlock pattern is, or report statisitics about your pattern. If you do not use the Android unlock pattern, please do not complete the HIT. This is an institution approved survey, and your participitation is protected. You must complete the survey portion of this HIT on an Android device; however, you do not need to log onto Mturk on your Android device to submit work."

NO.

Mar 22 2015 | [redacted] | flag | comment | flags, comments »

NO.

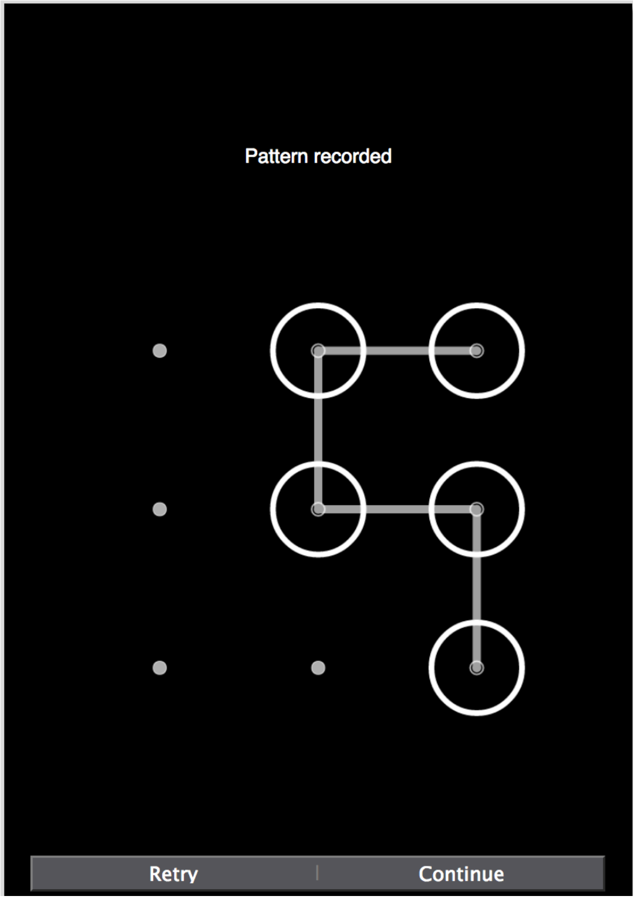
Procedure



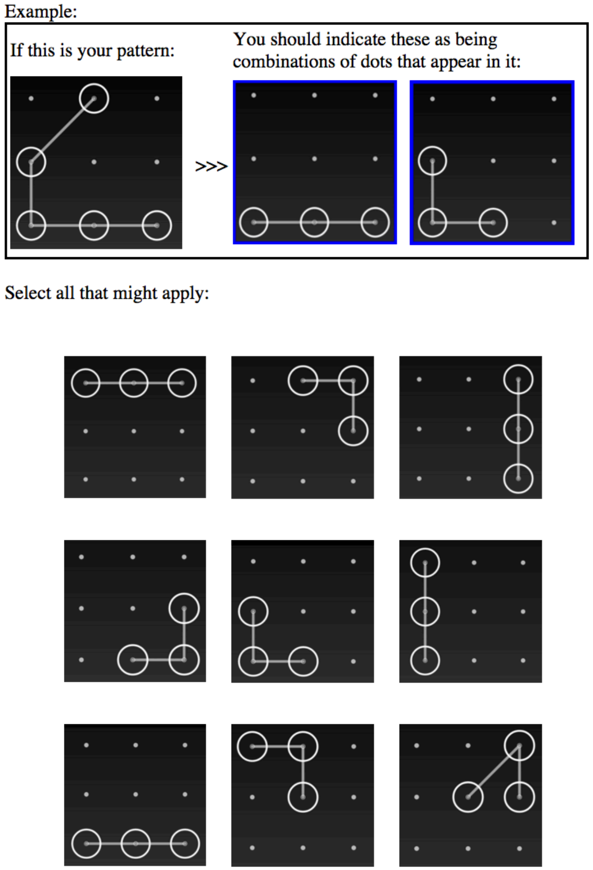
For this survey, you have two options. You may either provide your actual android unlock pattern or answer a series of questions about your pattern. Whichever you choose, the information recorded will be anonymous and not in any way connected with you or your account.

Report Pattern or Stats

Pattern Entry on Device


Pattern recorded

Select Features of Pattern


Example:

If this is your pattern:

>>>

You should indicate these as being combinations of dots that appear in it:

Select all that might apply:

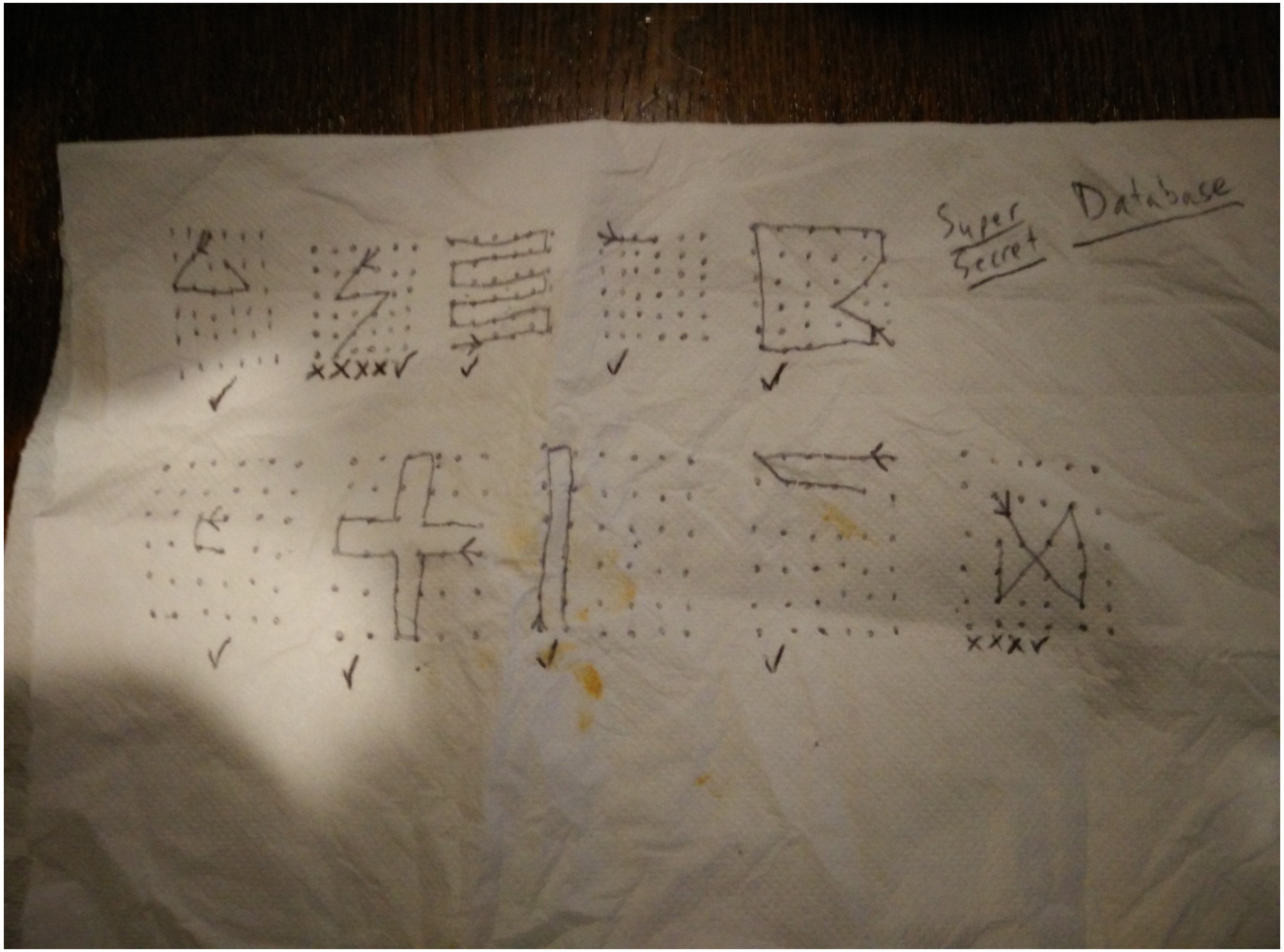
Attention Tests

- Device Completion
  - Elaborate code/token system
  - Check user agents (*yes, I know that can be forged*)

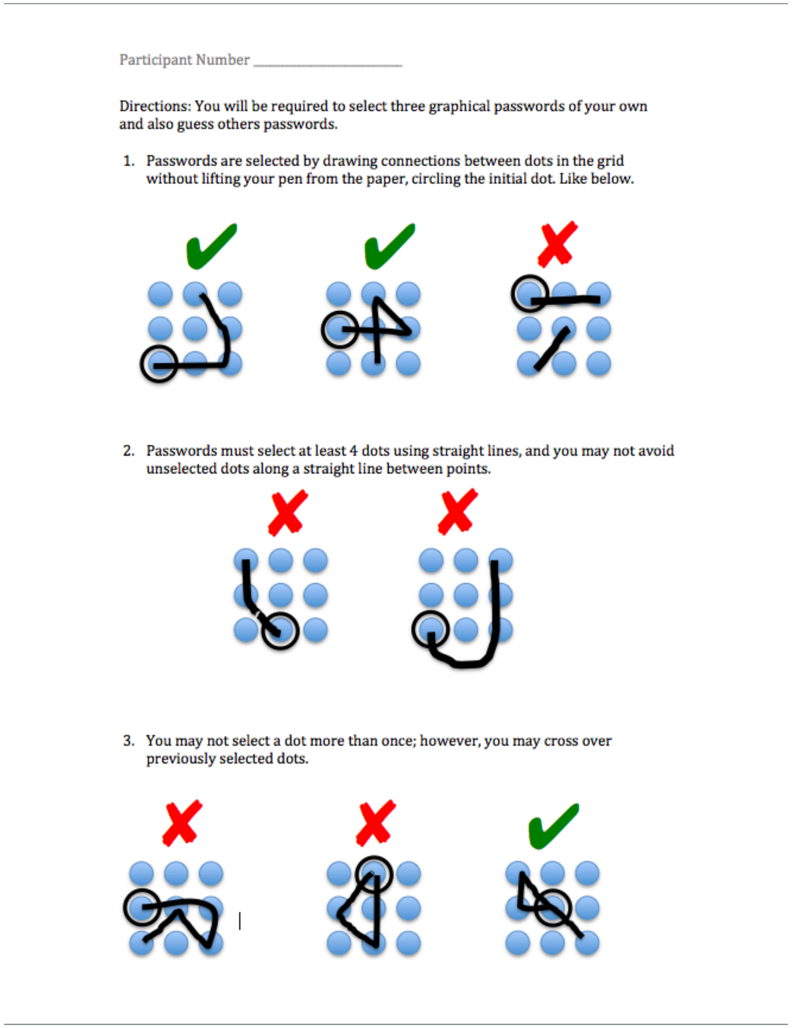
- Must enter in results twice
  - Report pattern, then survey, then report again
  - If results don't match, throw data out

- Rejections
  - We do not reject people within Mturk
  - Just don't include their data

Pen and Paper Survey



Adversarial Model



Participant Number ____________________

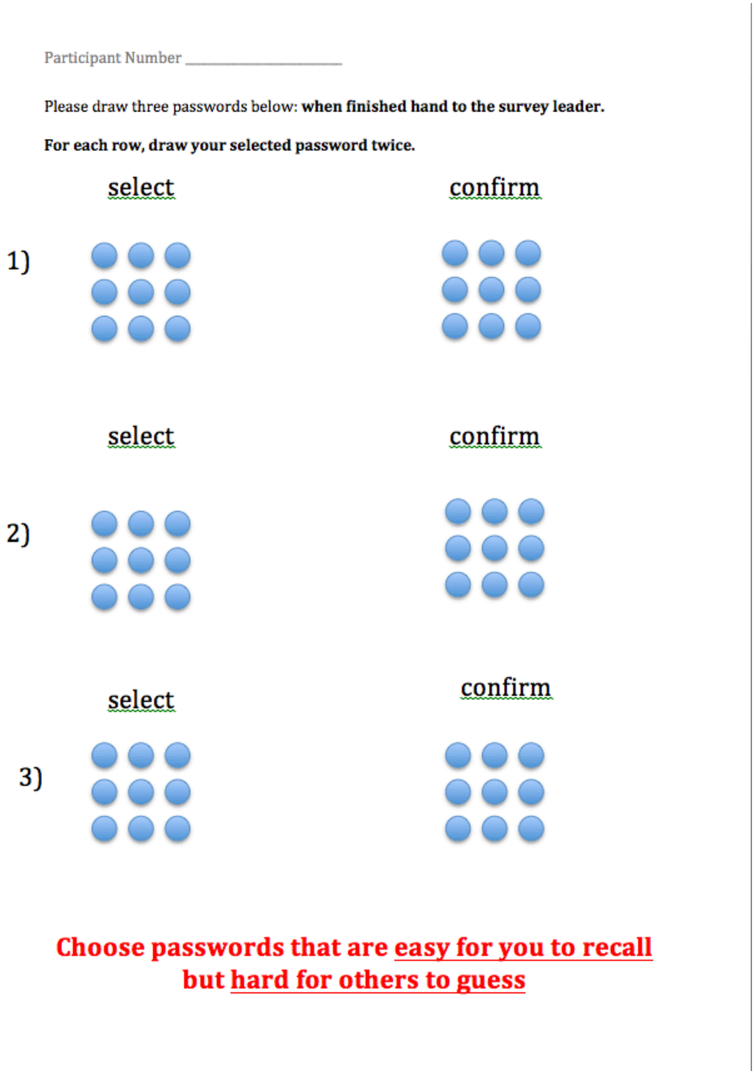
Directions: You will be required to select three graphical passwords of your own and also guess others passwords.

1. Passwords are selected by drawing connections between dots in the grid without lifting your pen from the paper, circling the initial dot. Like below.

2. Passwords must select at least 4 dots using straight lines, and you may not avoid unselected dots along a straight line between points.

3. You may not select a dot more than once; however, you may cross over previously selected dots.

Defensive Selection



Participant Number ____________________

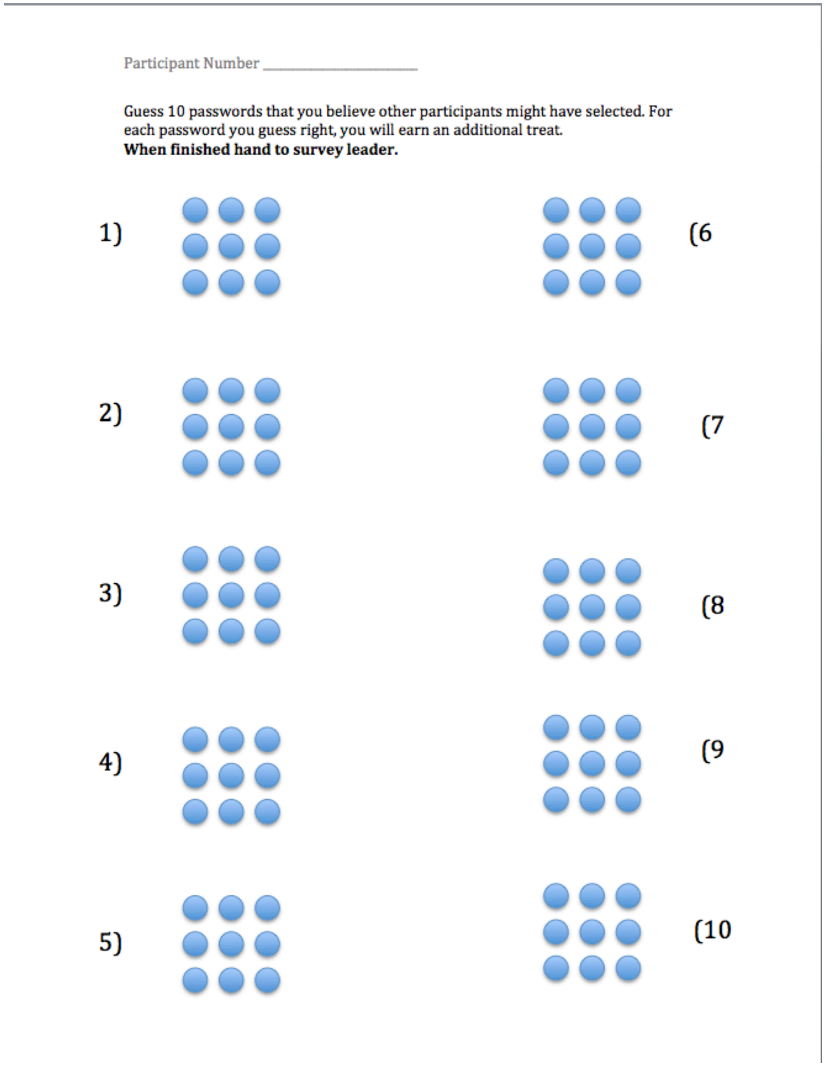
Please draw three passwords below: **when finished hand to the survey leader.**

For each row, draw your selected password twice.

	select	confirm
1)		
2)		
3)		

Choose passwords that are easy for you to recall but hard for others to guess

Offensive Selection



Participant Number ____________________

Guess 10 passwords that you believe other participants might have selected. For each password you guess right, you will earn an additional treat.
When finished hand to survey leader.

1)

2)

3)

4)

5)

(6

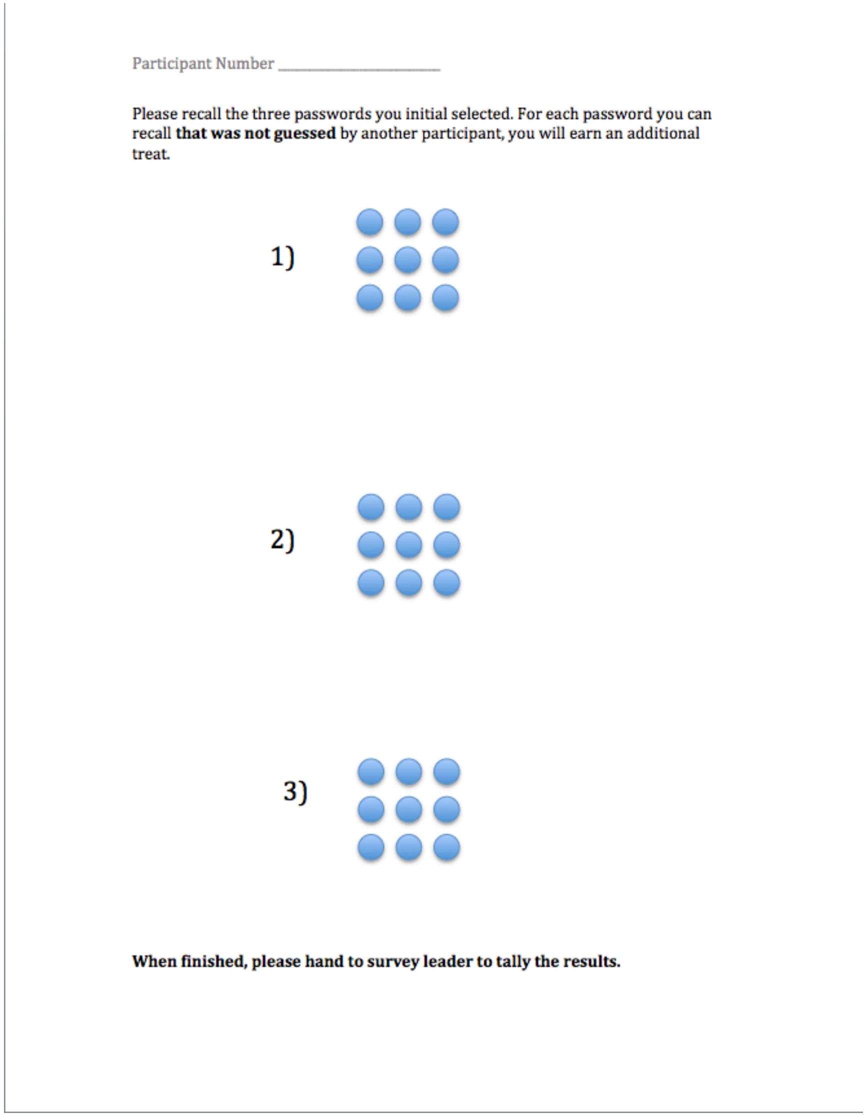
(7

(8

(9

(10

Recall



Participant Number ____________________

Please recall the three passwords you initial selected. For each password you can recall **that was not guessed** by another participant, you will earn an additional treat.

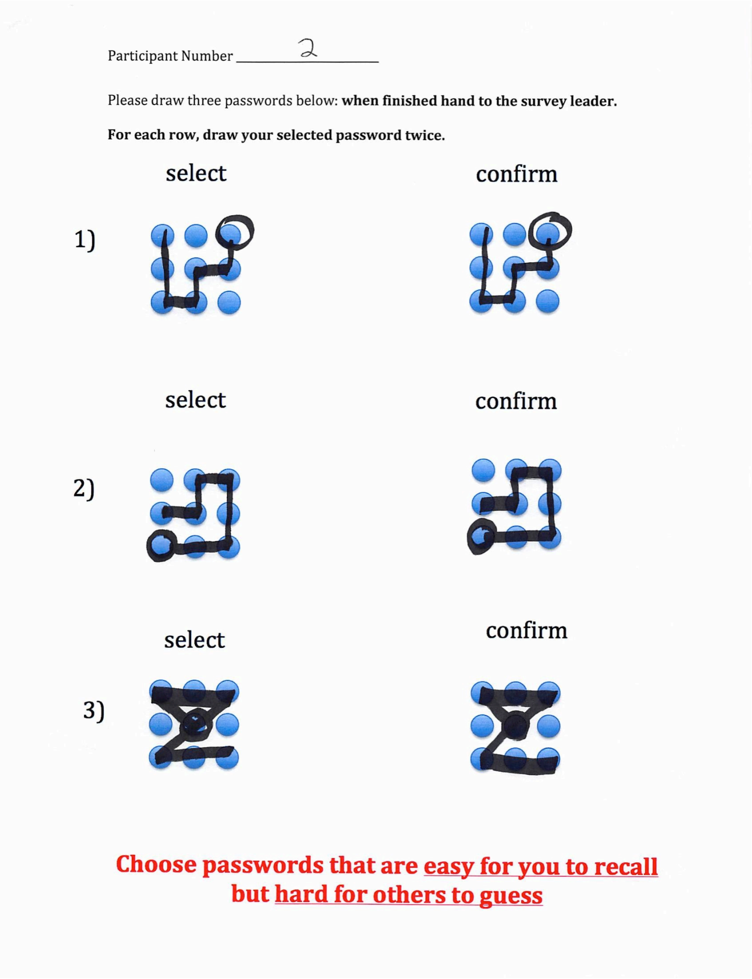
1)

2)

3)

When finished, please hand to survey leader to tally the results.

Sample Data



Participant Number 2

Please draw three passwords below: **when finished hand to the survey leader.**

For each row, draw your selected password twice.

select confirm

1)

select confirm

2)

select confirm

3)

Choose passwords that are easy for you to recall but hard for others to guess

Demographics and Collection

Self Report

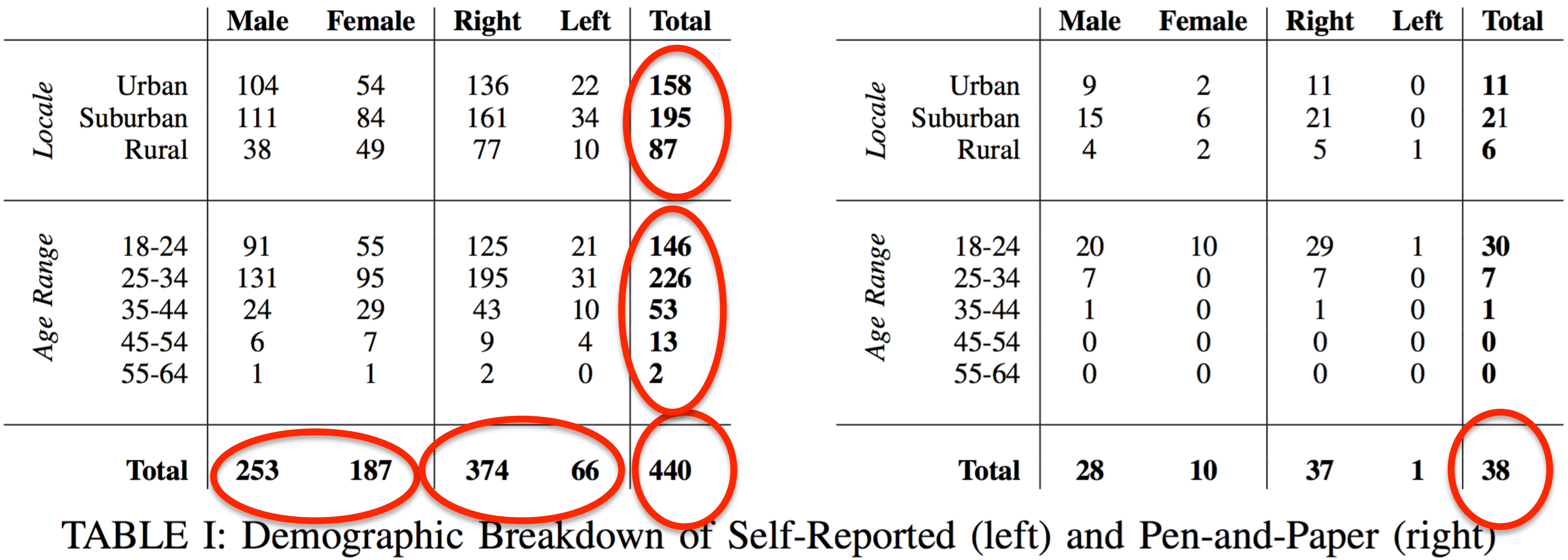
- 750 Respondents
- 440 Self-Reported their password, remaining provided statistics
- 251 Males, 189 Females
- Age range: 18 – 55+
- Location: USA

Pen and Paper

- 78 Participants
  - 8 Focus Groups over 6 weeks
  - 7 to 22 members per group
- 491 3x3 Patterns
  - 378 offensive
  - 113 defensive
- 501 4x4 Patterns
  - 382 Offensive
  - 119 Defensive
- 55males, 23 Females
- Age range: 18-40
- Location: @USNA and @UMBC

Demographics of 3x3 Data

Self-Reported **Pen-and-Paper**



		Male	Female	Right	Left	Total
Locale	Urban	104	54	136	22	**158**
	Suburban	111	84	161	34	**195**
	Rural	38	49	77	10	**87**
Age Range	18-24	91	55	125	21	**146**
	25-34	131	95	195	31	**226**
	35-44	24	29	43	10	**53**
	45-54	6	7	9	4	**13**
	55-64	1	1	2	0	**2**
	Total	**253**	**187**	**374**	**66**	**440**

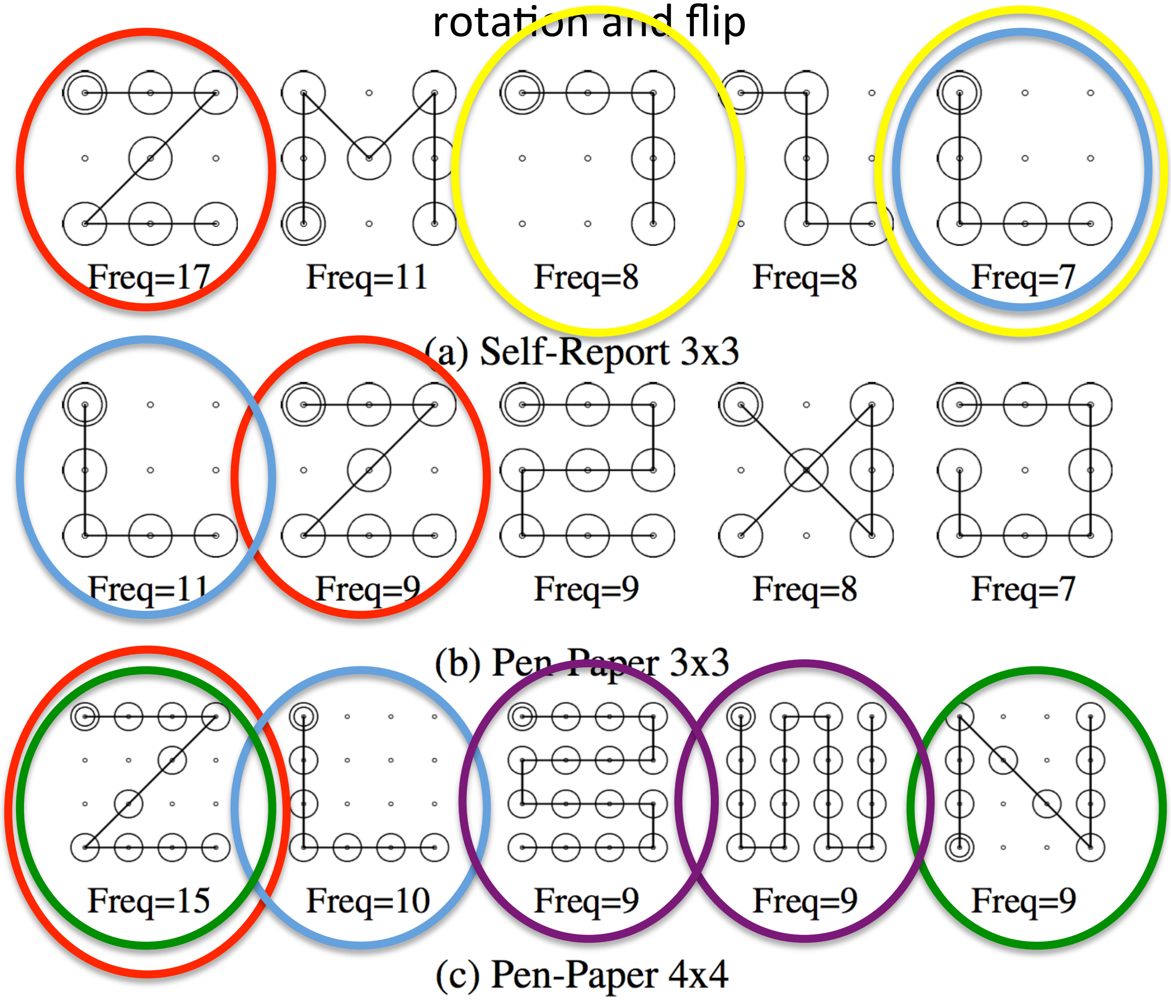
		Male	Female	Right	Left	Total
Locale	Urban	9	2	11	0	**11**
	Suburban	15	6	21	0	**21**
	Rural	4	2	5	1	**6**
Age Range	18-24	20	10	29	1	**30**
	25-34	7	0	7	0	**7**
	35-44	1	0	1	0	**1**
	45-54	0	0	0	0	**0**
	55-64	0	0	0	0	**0**
	Total	**28**	**10**	**37**	**1**	**38**

TABLE I: Demographic Breakdown of Self-Reported (left) and Pen-and-Paper (right)

DATA CHARACTERIZATION

[_BK:ACSAC'15]

Most Common Patterns



rotation and flip

Flip

(a) Self-Report 3x3

Embedding and Repetition

(b) Pen-Paper 3x3

Embedding and Repetition

Rotations

Flip

(c) Pen-Paper 4x4

Figure 6: Top 5 Most Frequently Occurring Patterns

[_BK:ACSAC'15]

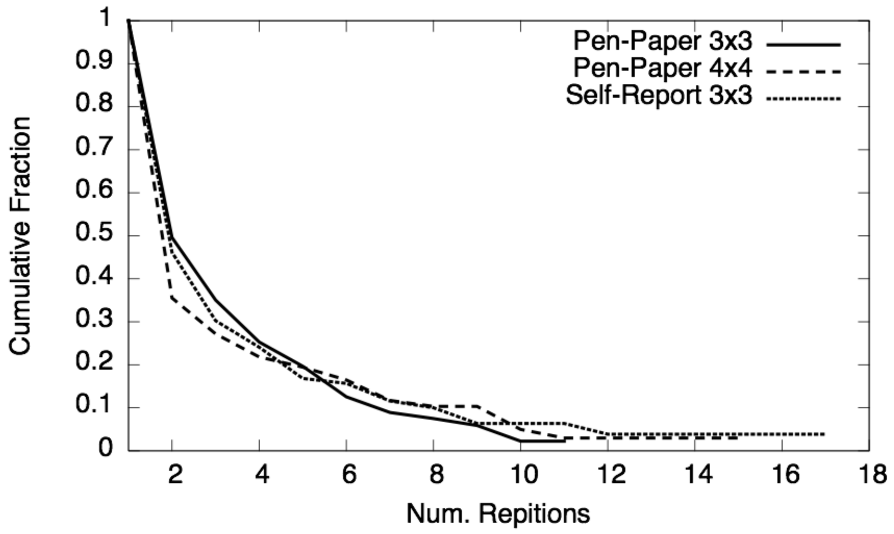


Figure 7: Cumulative fraction of patterns that repeat

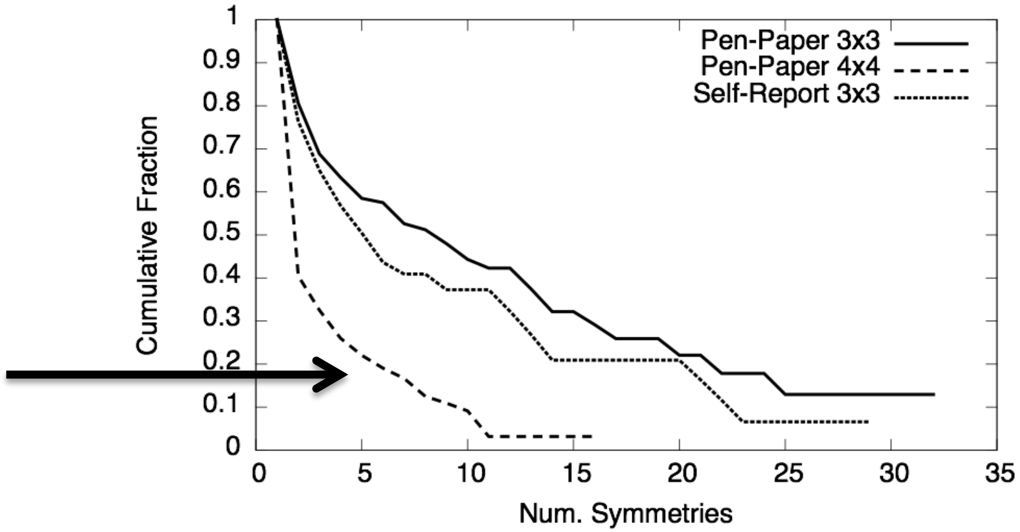
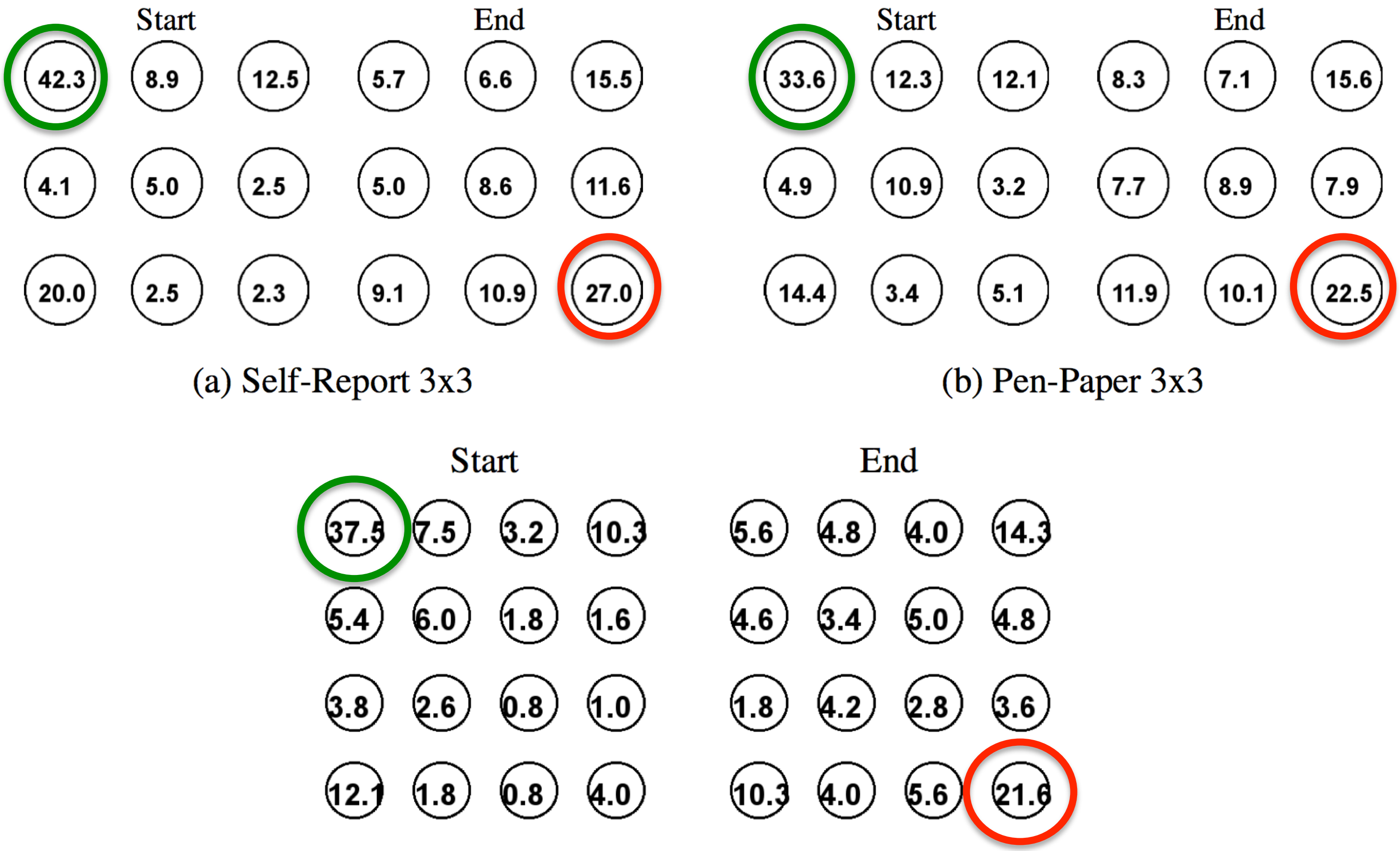


Figure 8: Cumulative fraction of patterns that have symmetries

Repetitions are consistent between 3x3 and 4x4, but 4x4 patterns have less symmetries

[_BK:ACSAC'15]

Start and End Conditions



Start			End		
42.3	8.9	12.5	5.7	6.6	15.5
4.1	5.0	2.5	5.0	8.6	11.6
20.0	2.5	2.3	9.1	10.9	27.0

(a) Self-Report 3x3

Start			End		
33.6	12.3	12.1	8.3	7.1	15.6
4.9	10.9	3.2	7.7	8.9	7.9
14.4	3.4	5.1	11.9	10.1	22.5

(b) Pen-Paper 3x3

Start				End			
37.5	7.5	3.2	10.3	5.6	4.8	4.0	14.3
5.4	6.0	1.8	1.6	4.6	3.4	5.0	4.8
3.8	2.6	0.8	1.0	1.8	4.2	2.8	3.6
12.1	1.8	0.8	4.0	10.3	4.0	5.6	21.6

(c) Pen-Paper 4x4

[_BK:ACSAC'15]

Length and Stroke Length

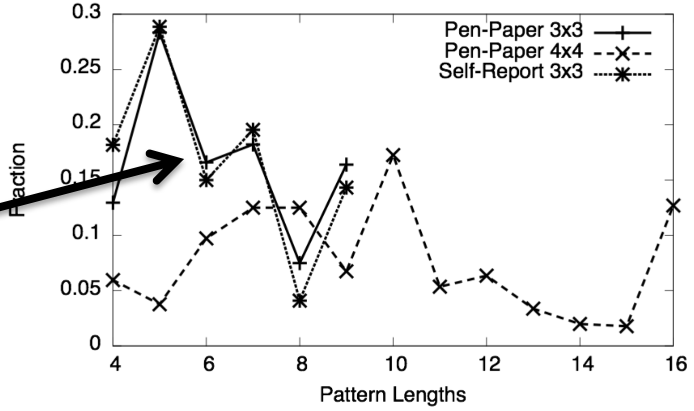


Figure 2: The distribution of length in the data sets.

Similar Distribution

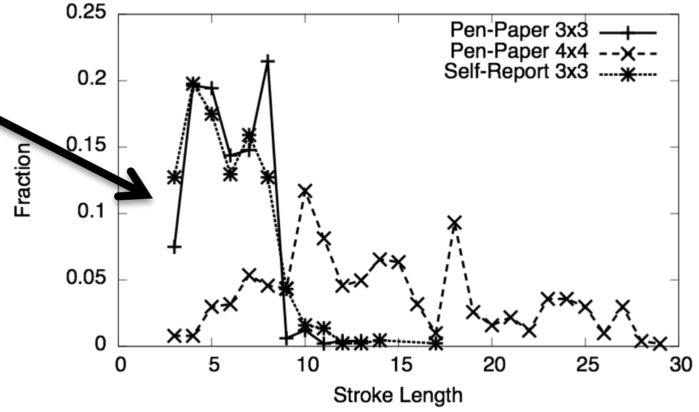


Figure 3: The distribution of stroke-lengths in the data set

Suggests that the pen-and-paper and self-report survey have similarly shaped patterns

[_BK:ACSAC'15]

STATISTICAL COMPARISONS

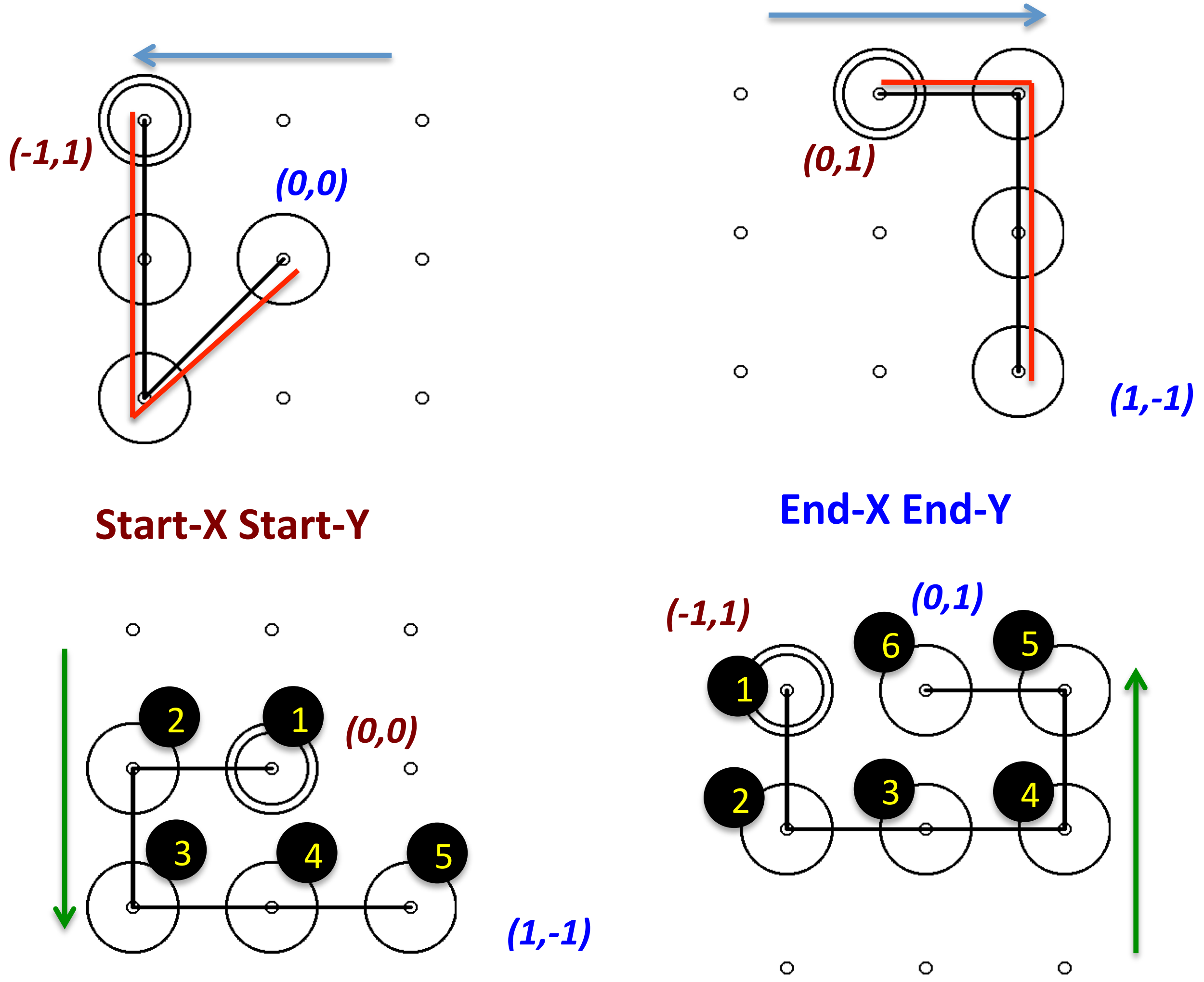
Features

- Ordinal Features
  - Height and Side
  - Length
  - Stroke Length
  - Start-X, Start-Y
  - End-X, End-Y

- Non-Ordinal Features
  - Knight Moves
  - Crosses
  - Exes
  - Non-Adjacency

[ATOY: WiSec'13] [UDWH: CCS'13] [_D:ACASC'14]

Ordinal Features



Side

Stroke Length

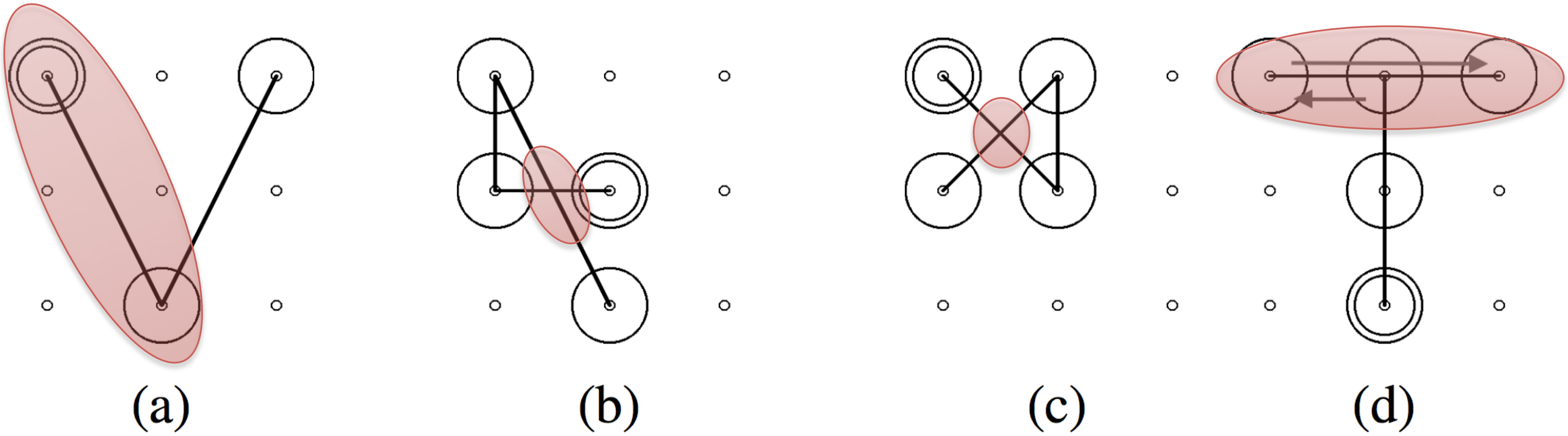
Start-X Start-Y

End-X End-Y

Height

Length

Non-Ordinal Features



Knight Move (kmove)

Cross

Ex

Non-Adjacent

Analysis Considerations

- Pen and Paper (PP)
  - *Average Pattern Stats*
    - For Pen and Paper: participants created 13 patterns (3 defensive and 10 offensive)
    - Used the average of the individual features across each of the 13 patterns
  - Example: *Length:* average length of all patterns generated for a given user

- Demographic Analysis
  - Only consider Self Reported (SR)
  - Samples for Pen and Paper were too small in some fields

Statistical Methods

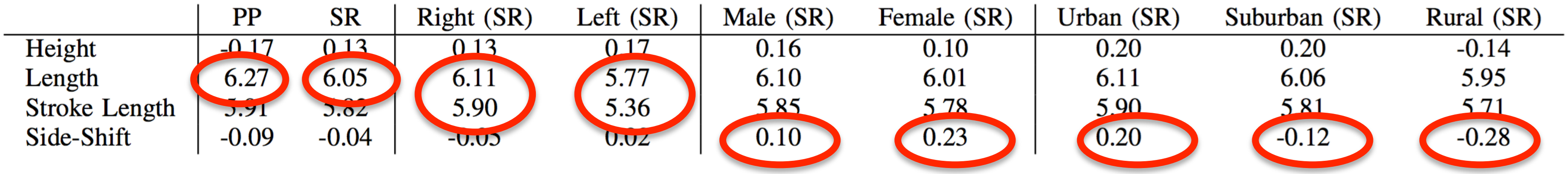
- Ordinal Measurements
  - Normality Tests: Anderson-Darling
  - Student's t-Test for normal data
  - Mann-Whitney U-Test for non-normal

- Non-Ordinal Measurements
  - χ^2-Test : presence or absence

- Multi Group Analysis
  - ANOVA testing

- P Corrections
  - Bonferonni

Ordinal Measurements

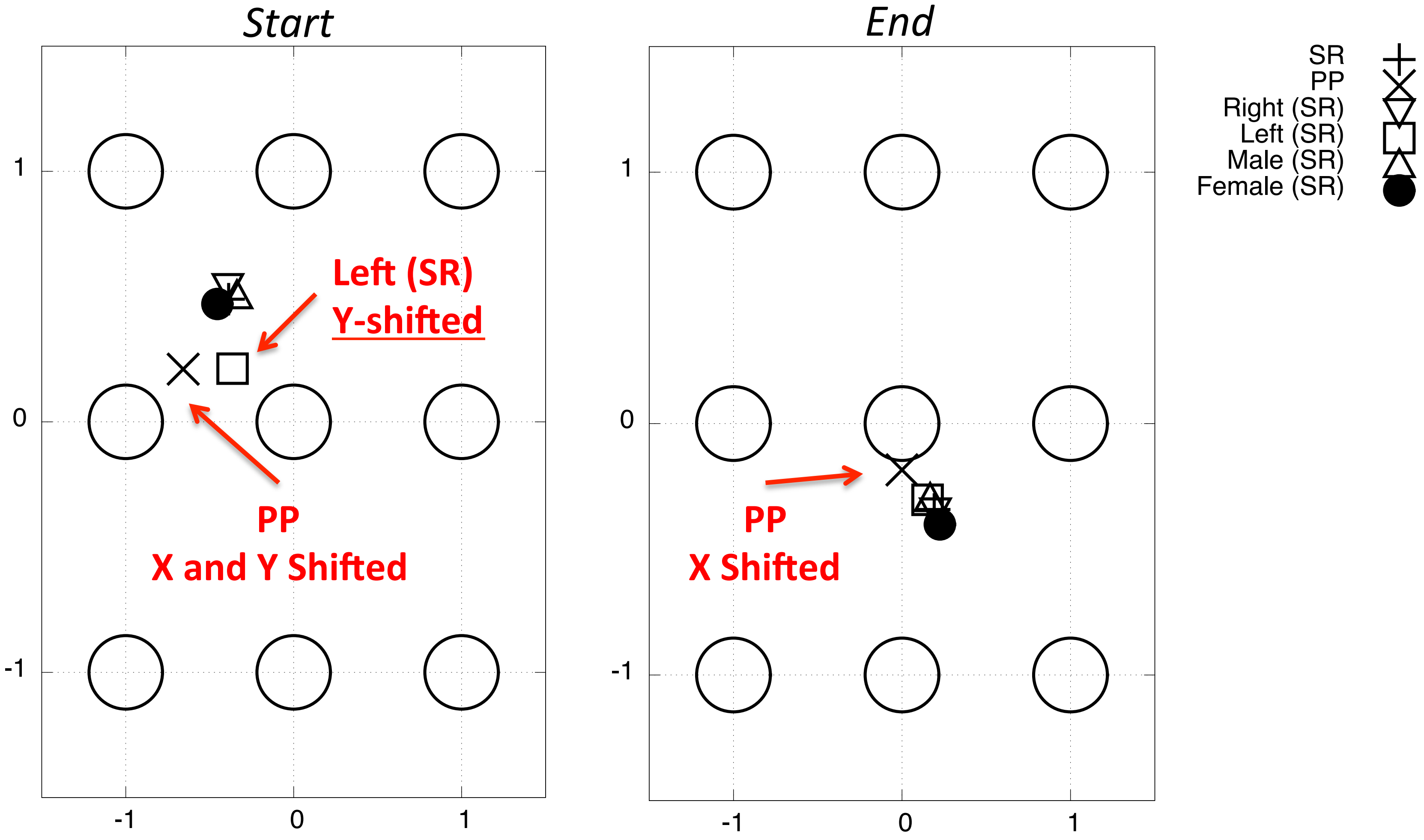


	PP	SR	Right (SR)	Left (SR)	Male (SR)	Female (SR)	Urban (SR)	Suburban (SR)	Rural (SR)
Height	-0.17	0.13	0.13	0.17	0.16	0.10	0.20	0.20	-0.14
Length	6.27	6.05	6.11	5.77	6.10	6.01	6.11	6.06	5.95
Stroke Length	5.91	5.82	5.90	5.36	5.85	5.78	5.90	5.81	5.71
Side-Shift	-0.09	-0.04	-0.05	0.02	0.10	0.23	0.20	-0.12	-0.28

$p < 0.05$ $p < 0.03$ $p < 0.02$ $p < 0.02$

- PP vs. SR
  - Pen and Paper patterns tend to be a bit longer
- Handedness
  - Right handed patterns a bit longer and with more stroke length
- Side Shifting
  - Gender: Woman respondents more right shifted than Men
  - Locale: Urban respondents more right shifted than Urban and Rural respondents

Start and End Locations

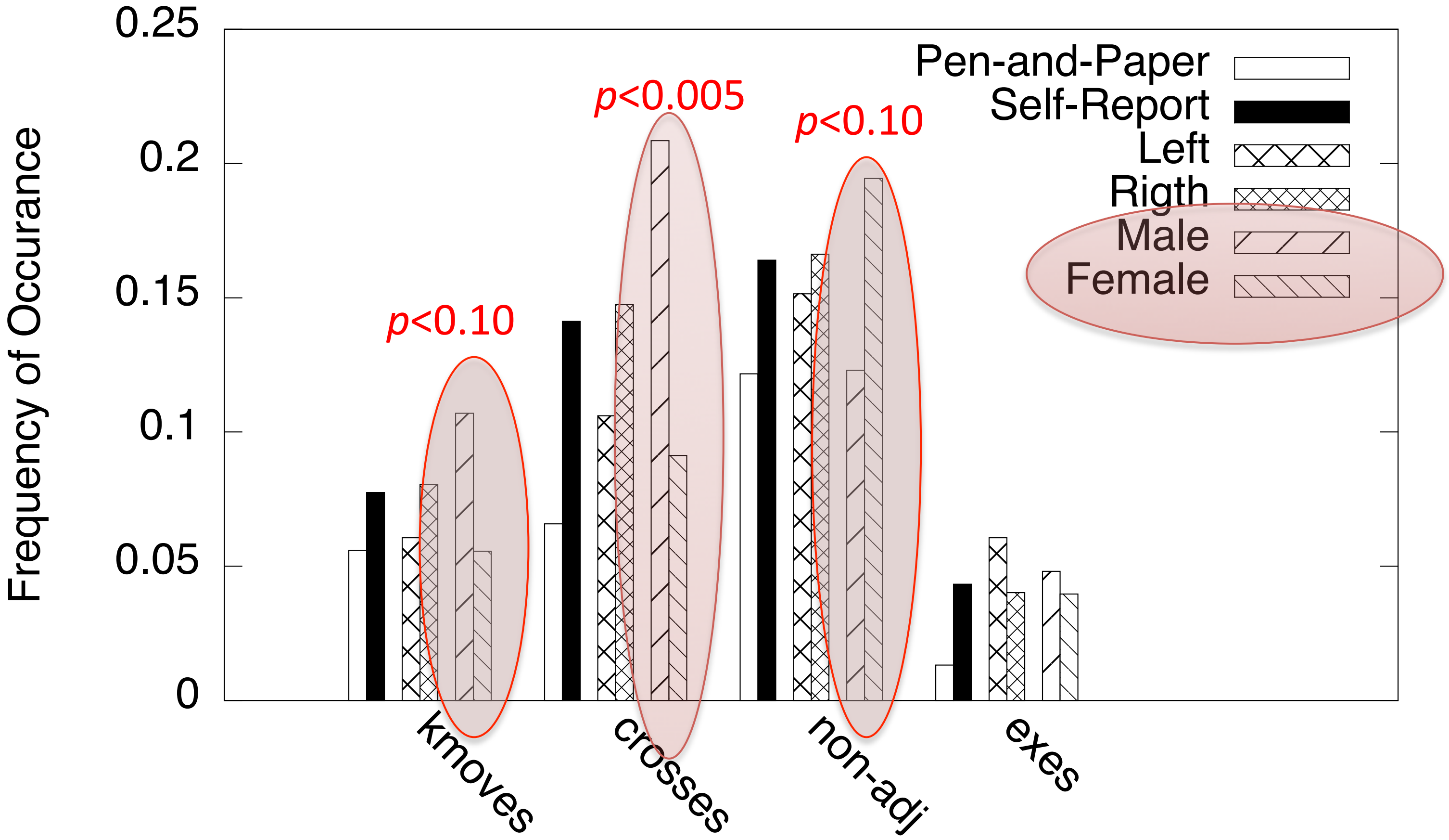


Left (SR)
Y-shifted

PP
X and Y Shifted

PP
X Shifted

Non-Ordinal Features



STRENGTH METRICS

Guessability

- *How many guesses does it take for an attacker to guess a given password?*

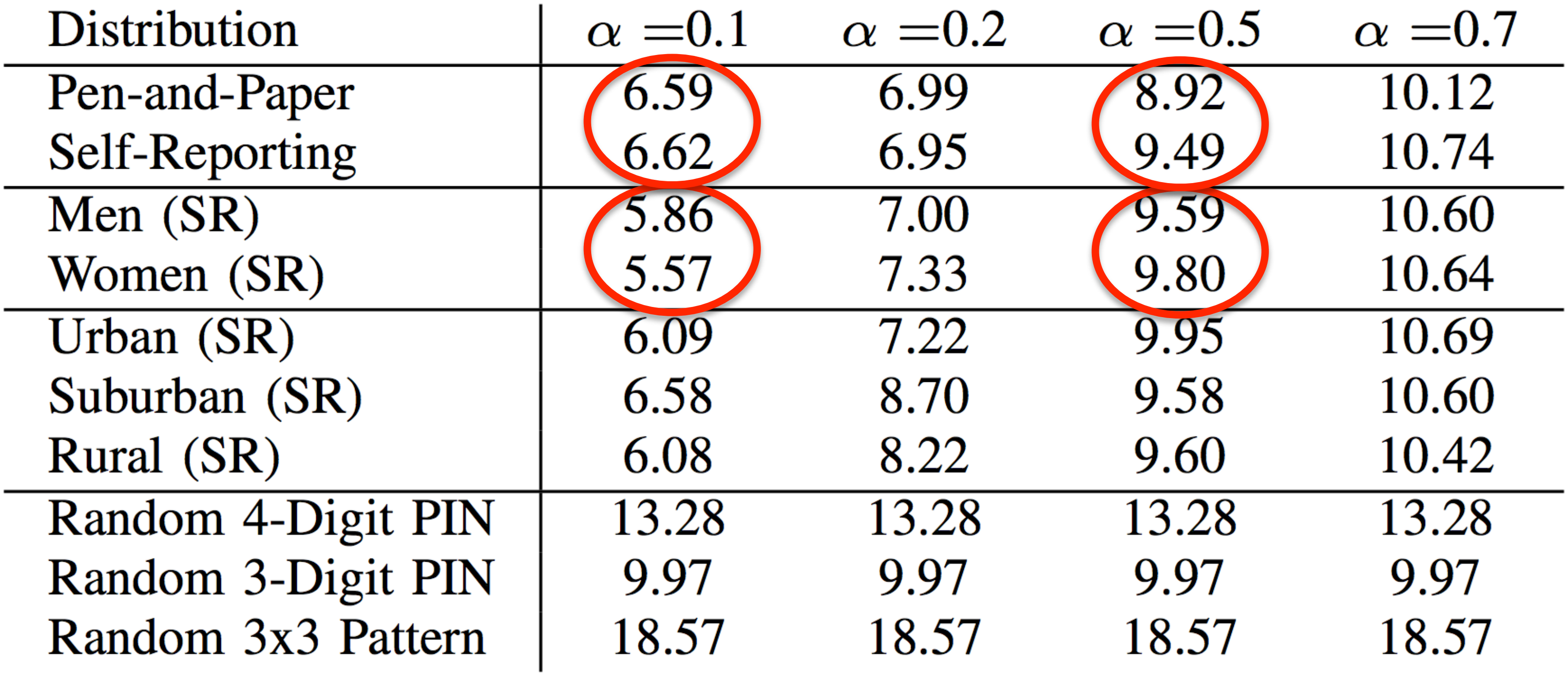
- **PARTIAL GUESSABILITY** (alpha-guesswork)
  - *How many guesses does it take to guess a fraction of the dataset?*
  - Measured in bits of information

- **Offline Attack:**
  - Assumes attack can crack passwords without having to engage the authentication method (e.g., cracking hashes)
  - No lockouts (traditionally, 20 guesses on Android)

Guessing Algorithm

- Input: Training Set, Guessing Set
- Train Likelihood Measure (Markov model)
  - Use training set and symmetries of training set with different weights
- Guess Order
  1. All patterns in training set order based on frequency with ties broken by likelihood measure
  2. All rotations/symmetries of training set ordered based on likelihood measure
  3. Set of generated patterns using the Markov Model ordered by likelihood measure

Guessability Strength



Distribution	α =0.1	α =0.2	α =0.5	α =0.7
Pen-and-Paper	6.59	6.99	8.92	10.12
Self-Reporting	6.62	6.95	9.49	10.74
Men (SR)	5.86	7.00	9.59	10.60
Women (SR)	5.57	7.33	9.80	10.64
Urban (SR)	6.09	7.22	9.95	10.69
Suburban (SR)	6.58	8.70	9.58	10.60
Rural (SR)	6.08	8.22	9.60	10.42
Random 4-Digit PIN	13.28	13.28	13.28	13.28
Random 3-Digit PIN	9.97	9.97	9.97	9.97
Random 3x3 Pattern	18.57	18.57	18.57	18.57

Summary

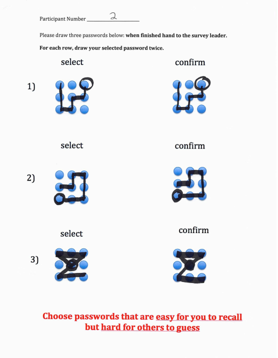
Research Questions

What are the statistical differences in the patterns based on the collection methods?

Are there differences in patterns based on demographics?

Can we leverage demographics to attack patterns?

Methodologies and Data Characteristics



Participant Number 2

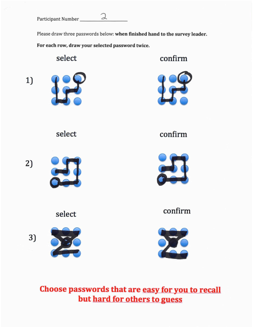
Please draw three passwords below: **when finished hand to the survey leader.**

For each row, draw your selected password twice.

select
confirm
1)
select
confirm
2)
select
confirm
3)

Choose passwords that are easy for you to recall but hard for others to guess

Comparisons of Collection Methodologies and Demographics



Participant Number 2

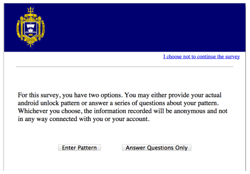
Please draw three passwords below: **when finished hand to the survey leader.**

For each row, draw your selected password twice.

Choose passwords that are easy for you to recall but hard for others to guess

select
confirm
1)
select
confirm
2)
select
confirm
3)

Vs.



For this survey, you have two options. You may either provide your actual android unlock pattern or answer a series of questions about your pattern. Whichever you choose, the information recorded will be anonymous and not in any way connected with you or your account.

Demographics and Pattern Strength

Distribution	α =0.1	α =0.2	α =0.5	α =0.7
Pen-and-Paper	6.59	6.99	8.92	10.12
Self-Reporting	6.62	6.95	9.49	10.74
Men (SR)	5.86	7.00	9.59	10.60
Women (SR)	5.57	7.33	9.80	10.64
Urban (SR)	6.09	7.22	9.95	10.69
Suburban (SR)	6.58	8.70	9.58	10.60
Rural (SR)	6.08	8.22	9.60	10.42
Random 4-Digit PIN	13.28	13.28	13.28	13.28
Random 3-Digit PIN	9.97	9.97	9.97	9.97
Random 3x3 Pattern	18.57	18.57	18.57	18.57

Conclusions

- **Pen-and-Paper is a *good* proximate for patterns as they are self reported**
  - Minor difference in length (.2 contact points)
  - Minimal difference in stroke length

- Demographics:
  - Gender differences in side shifting: **hand size?**
  - **Handedness differences in length?**
  - Locale difference in side shifting?

- Guessability/Pattern Strength
  - **Some knowledge of demographics can help**
  - Advantage **fades quickly** when attack more of the patterns

Adam J. Aviv Justin Maguire Jeanne Luning-Prak

THANKS AND QUESTIONS?