

An Empirical Investigation of Security Fatigue

The Case of Password Choice after solving a CAPTCHA

Kovila P.L. Coopamootoo

Thomas Gross

M. Faizal R. Pratama

Copyright © 2017, Kovila Coopamootoo, Newcastle University



Outline

- **Background**
- **Previous Research**
- **Current Experiment Design**
- **Results**
- **Contributions**

Security Fatigue/ Depletion



A Barrier that interferes with productivity



- **Threshold of acceptance**
- **Burden**

- **Weariness**
- **Tiredness**
- **Turned-off, overwhelmed**

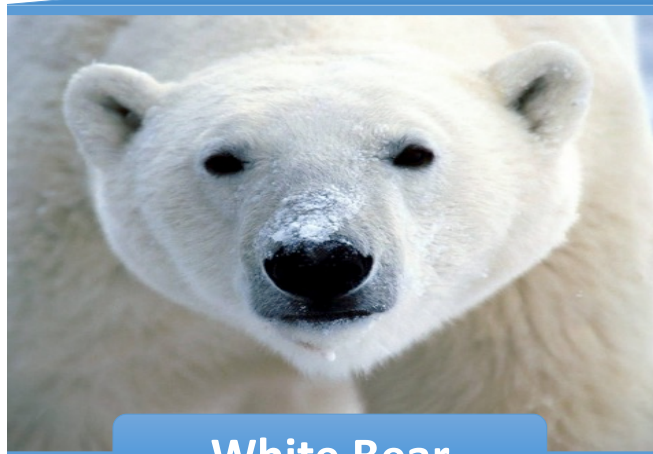
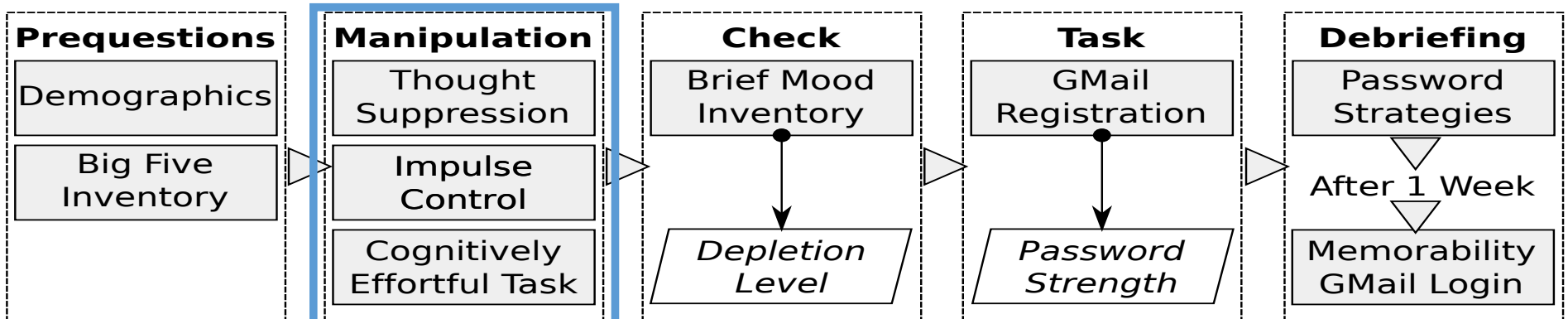
Cognitive Depletion [Baumeister]

- *Cognitive effort* needed to control impulses and to think about 'hard' problems.
- Limited energy to exert cognitive effort.
- Once cognitively depleted, the capacity to exert cognitive effort is incapacitated.

BAUMEISTER, R. et al. Ego depletion: is the active self a limited resource? *Personality and social psychology* 74 (1998), 1252–1265.

Previous Research

Past Empirical Research: Incidental Effort (LASER'16)



White Bear

1. Train impulse:
Cross-out letter 'e'
2. Change the game:
Cross-out 'e' unless adjacent to a vowel, such as in 'bear'.

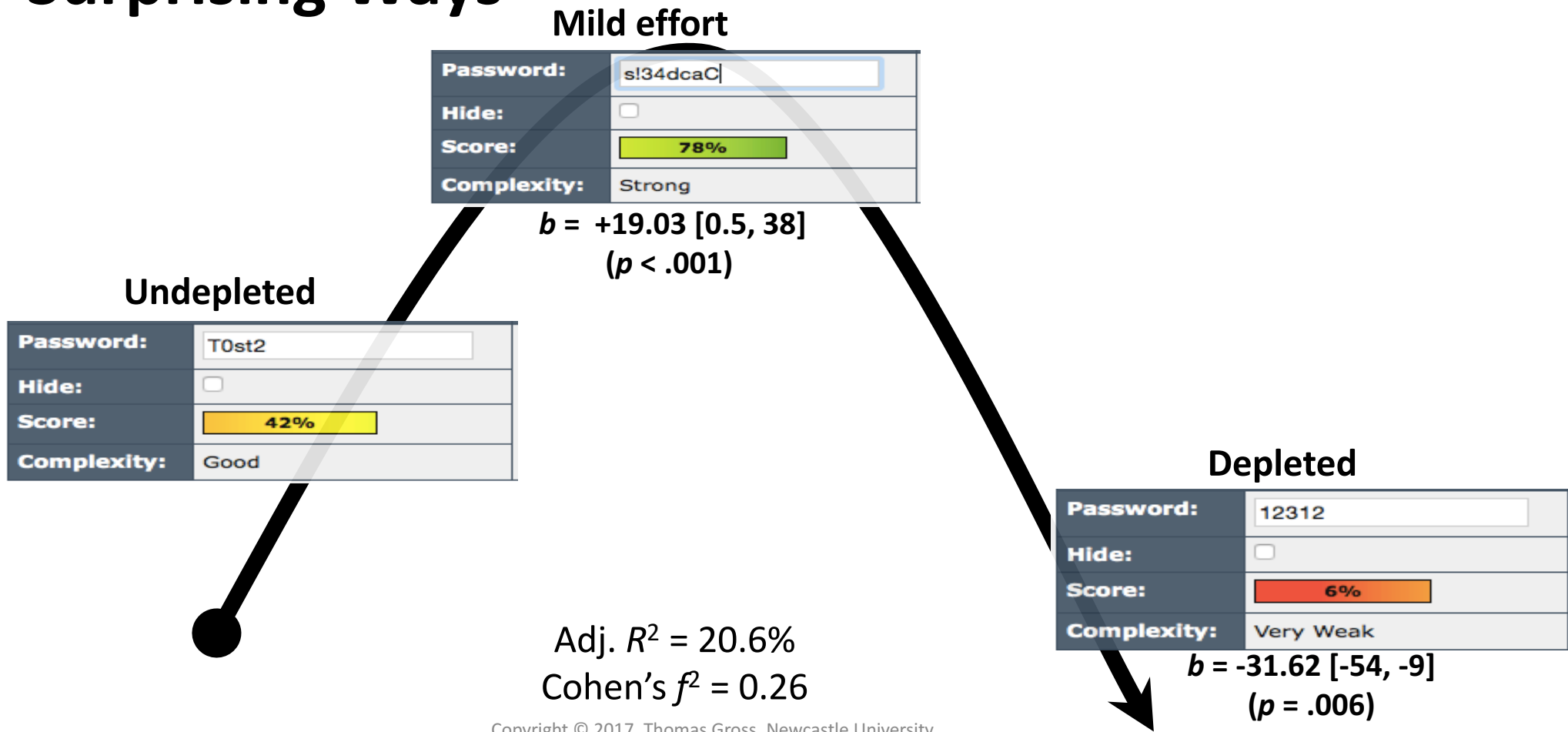
Impulse



Stroop

GROSS, T. et al. Effect of cognitive depletion on password. LASER 2016.

Cognitive Effort Impacts Password Strength in Surprising Ways



Current Scenario

CAPTCHA & PASSWORD

Why this scenario?

- CAPTCHA widely used across Web
- Balancing security & usability remains a challenge
- Incorrect CAPTCHA & form reset – result in solving the CAPTCHA before password
- CAPTCHA as a gateway
 - Frequent requests from the same IP
 - TOR Anonymizer users

Create account

Username [\(help me choose\)](#)

Password

Confirm password

Email address (optional)

To protect the wiki against automated account creation, we kindly ask you to enter the words that appear below in the box [\(more info\)](#):

CAPTCHA Security check



[Refresh](#)

Can't see the image? [Request an account](#)

Create your account

In a Nutshell

- **Background:** CAPTCHAs are known to be difficult for humans. Effort spent in one security task can impact subsequent tasks.
- **Aim:** Investigate how users' password choice differ after solving a CAPTCHA vs. not solving a CAPTCHA.
- **H₁:** *Solving a text/picture CAPTCHA causes weaker passwords than not solving a CAPTCHA.*
- **Method:** Three groups of 22 participants each create a password – one solves a text CAPTCHA, another a picture CAPTCHA, the other not.
- **Result:** Solving a CAPTCHA implies weaker passwords.

Difference from previous study

- Sequence of Security tasks
- Effort Induced via solving CAPTCHA is *Incidental* to Password Choice
 - CAPTCHA is known to be difficult for users
- Yet *Integral* to Security context in general

Similarities with previous study

- Experimental/Scientific Methodology (guidelines cited below)
- Same measurements (manipulation check, dependent variable)
- Same password task
- Lab experiment
- Student sample

Study Design

Conditions: text-CAPTCHA, 2-digit multiplication (48 X 97), CONT

DV: password strength

N = 40, Newcastle University

Pre-Study

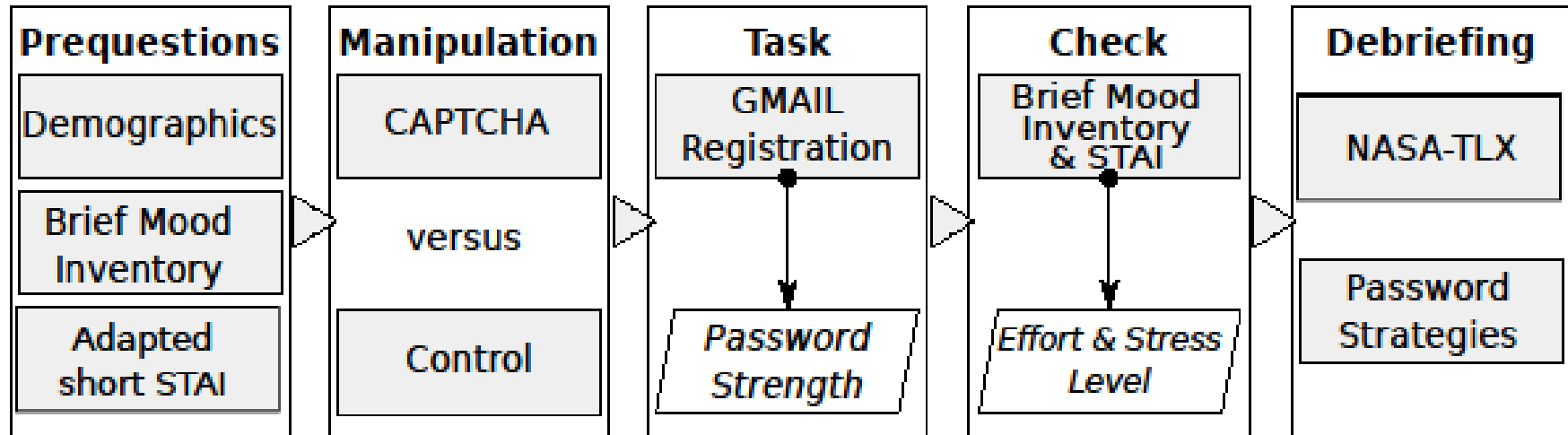
Conditions: text-CAPTCHA, picture-CAPTCHA, CONT

DV: password strength

N = 66, University of Derby

Main Study

Main Experiment Structure



Sampling (N=66)

Students: 38 local nationals;
non-computer science.
Gender: 30 female
Age M=21.79 (SD=3.223).

Random Block Assignment.

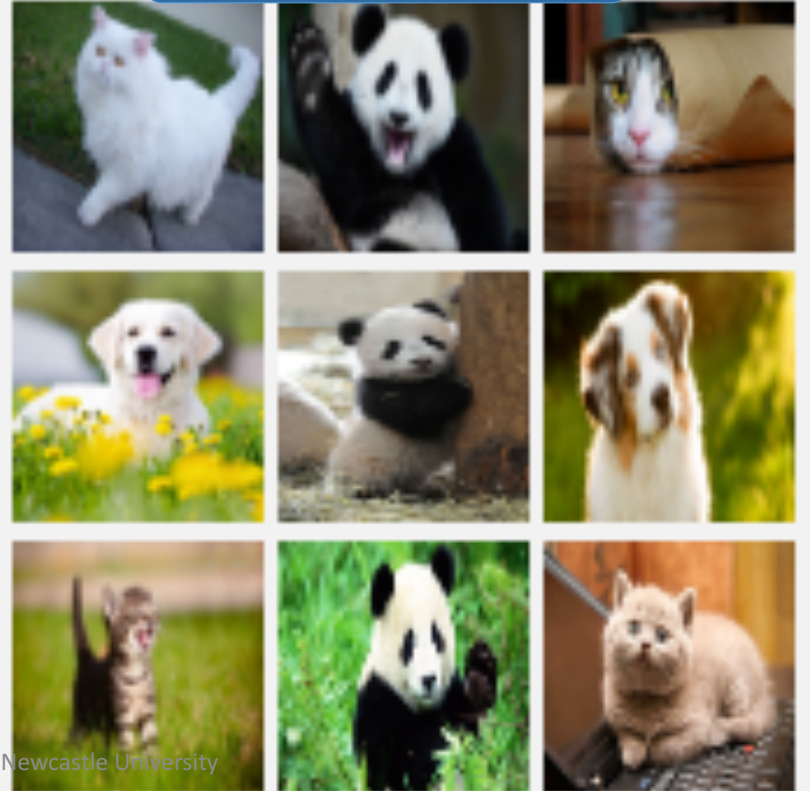
N = 22 in each group (text CAPTCHA, picture CAPTCHA, control)

Effort Manipulation

Text CAPTCHA



Picture CAPTCHA



Measurement: Password Strength

Passwordmeter.net

Test Your Password		Minimum Requirements	
Password:	<ul style="list-style-type: none"> Minimum 8 characters in length Contains 3/4 of the following items: <ul style="list-style-type: none"> Uppercase Letters Lowercase Letters Numbers Symbols 	
Hide:	<input checked="" type="checkbox"/>		
Score:	69%		
Complexity:	Strong		

Additions	Type	Rate	Count	Bonus
Number of Characters	Flat	$+(n*4)$	11	+ 44
Uppercase Letters	Cond/Incr	$+\left((len-n)*2\right)$	2	+ 18
Lowercase Letters	Cond/Incr	$+\left((len-n)*2\right)$	8	+ 6
Numbers	Cond	$+(n*4)$	1	+ 4
Symbols	Flat	$+(n*6)$	0	0
Middle Numbers or Symbols	Flat	$+(n*2)$	1	+ 2
Requirements	Flat	$+(n*2)$	4	+ 8

Deductions				
Letters Only	Flat	$-n$	0	0
Numbers Only	Flat	$-n$	0	0
Repeat Characters (Case Insensitive)	Comp	-	4	- 1
Consecutive Uppercase Letters	Flat	$-(n*2)$	0	0
Consecutive Lowercase Letters	Flat	$-(n*2)$	6	- 12
Consecutive Numbers	Flat	$-(n*2)$	0	0
Sequential Letters (3+)	Flat	$-(n*3)$	0	0
Sequential Numbers (3+)	Flat	$-(n*3)$	0	0
Sequential Symbols (3+)	Flat	$-(n*3)$	0	0

Password strength score additive or deductive features (e.g., numbers/symbols)

Additional Penalties

- dictionary word (-25),
- user's real name (-50),
- username (-50),
- user's student id (-50)

Other: zxcvbn

Measurement: Effort & Stress

- Brief Mood Inventory (BMI) [Tice et al., 2007]
 - Short form measuring excited, thoughtful, tired, happy, worn out, sad, (g) angry, calm
- Cognitive Workload (NASA-TLX) [Hart & Staveland, 1988]
 - Mental workload via dimensions of mental demand, physical demand, temporal demand, performance, effort and frustration
- State Trait Anxiety Inventory [Spielberger, 1970]
 - Y-1 questionnaire measuring stress (Y-1 for state anxiety as opposed to trait anxiety in Y-2)

Tice et al. Restoring the self: Positive affect helps improve self-regulation following ego depletion. *Journal of Experimental Social Psychology* 43, 3 (2007), 379–384.
Hart G., & Staveland, L. E. Development of NASA-TLX (task load index): Results of empirical and theoretical research. *Advances in psychology* 52 (1988), 139–183.
Spielberger, C. D., et al. STAI manual for the State-Trait Anxiety Inventory. Palo Alto, 1970.

Manipulation Check

Effort Exerted

Diff_Tiredness

$X^2 (2) = 12.736, p = .002 < .05$

text-CAPTCHA resulting in
more effort spent

Performance

Time to solve CAPTCHA

- text-based CAPTCHA condition took longer to solve than the picture-CAPTCHA condition, $t(42)=15.271, p < .001$.
- [Note for overall time in study - CONT: M = 551, text-CAPTCHA: M = 776, picture-CAPTCHA: M = 579]

Success at completing CAPTCHA

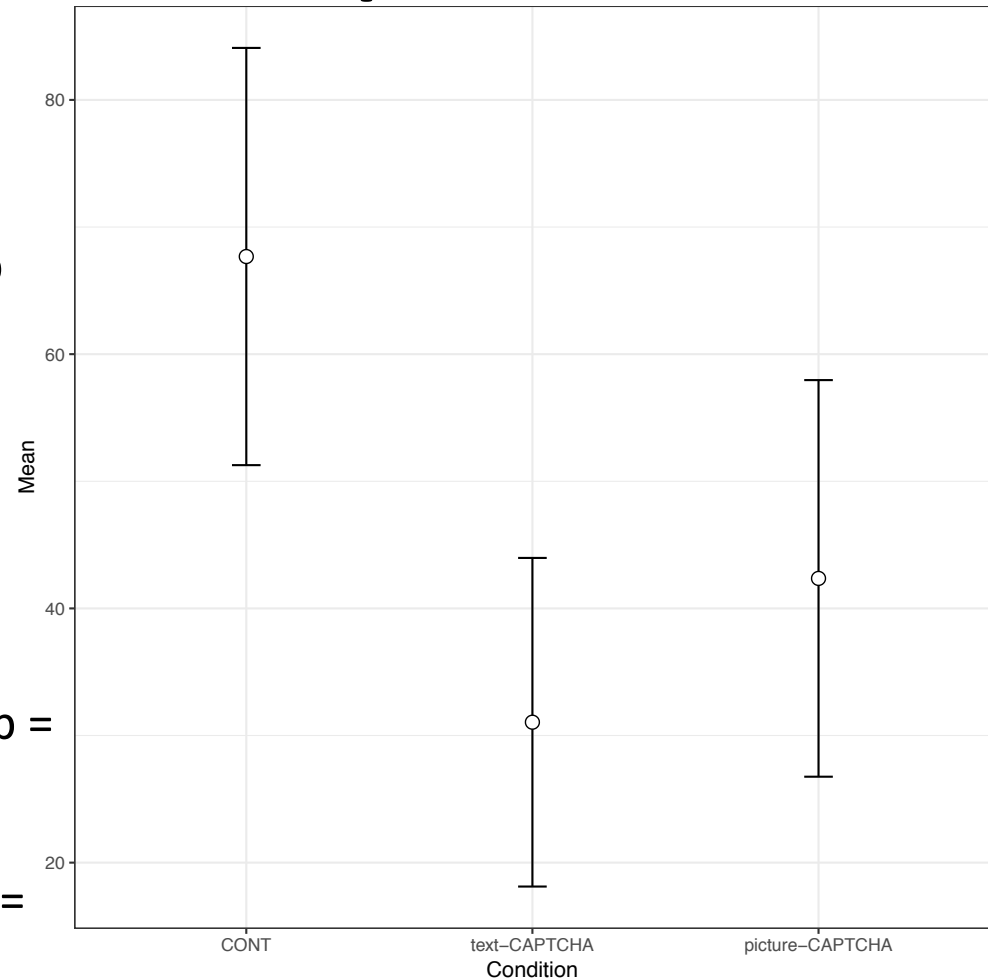
- text-based CAPTCHA: 5
- picture-CAPTCHA: 20

Checking Results

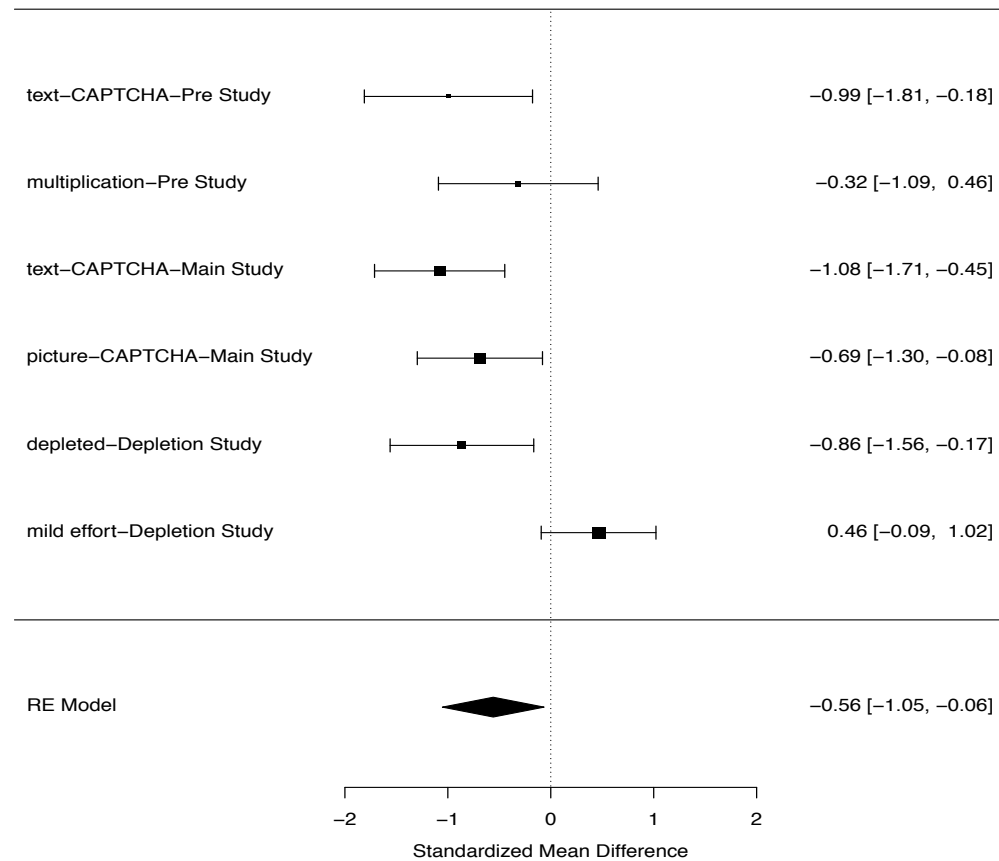
- text-CAPTCHA took longer to solve

Results: One-way ANOVA with the password strength score as DV

- $F(2,63) = 6.716$, $p = .002 < .05$
- Effect size in Cohen's $f = .42$ (medium to large effect)
- Eta-squared = $.176$ [0.043, 0.296]
- Tukey HSD PostHoc test, password strength
 - lower in text-CAPTCHA than in CONT with $p = .002 < .05$, effect size in Hedges' $g = 1.08$ [0.44, 1.71]
 - lower in picture-CAPTCHA than in CONT, $p = .042 < .05$, effect size in Hedges' $g = 0.69$ [0.08, 1.29]



Meta-Analysis: Password Strength as DV



Forest Plot showing difference between mean password strength across treatments vs. control, with overall effect size of 0.56

Contributions

- First study investigating fatigue in a sequence of security tasks
- Findings
 - Weaker password with CAPTCHA (supported in meta-analysis)
- Methodological
 - An example for methods replication
 - Documentation for forward reproducibility
 - Manipulation extended from psychological tasks to common CAPTCHA

Limitations

- Sampling bias
- Sample size and power statistics
 - At least 80% power, yet CIs not especially tight
- Post questionnaires
 - Subjective evaluation of CAPTCHA
- Stress
 - Participants seemed to have evaluated the GMail task

Things to take away

- Password choice is poorer after solving a CAPTCHA
- Positioning of CAPTCHA wrt password selection is important
- Overall security is impacted when effortful security tasks are designed in sequence
 - Would this cause rejection of security overall?
 - What combined effort can the user bear?

Questions/Discussion

kovila.coopamootoo@newcastle.ac.uk