Poster: Effect of Cognitive Depletion on Password Choice

Thomas Gross, Kovila P.L. Coopamootoo and Amina Al-Jabri
School of Computing Science
Newcastle University
Email: firstname.lastname@newcastle.ac.uk

***Abstract*—This lab experiment with 100 subjects is the first to investigate the impact of cognitive effort and depletion on the choice of user passwords. Two groups of 50 subjects each were asked to generate a password. One group was cognitively depleted, the other was not. Password strength was measured and compared across groups. We find that subjects who are cognitively depleted create worse passwords than undepleted subjects. Surprisingly, subjects who report mild cognitive exertion create better password than undepleted subjects.**

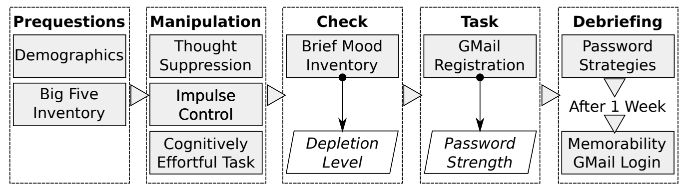

Figure 1. Experiment structure with cognitive depletion as manipulation.

I. Introduction

The Limited Strength model [1] of cognitive psychology predicts that cognitive effort is a limited resource. Consequently once depletion is reached, subsequent decisions are impeded. We investigate how password choice differs between depleted and non-depleted users. Two groups of 50 subjects each were asked to generate a password. Password strength was measured and compared across groups.

Using a stepwise linear regression we found that depletion level, personality traits and mood, predict password strength, with an overall adjusted $R^2 = .206$. The depletion level was the strongest predictor of password strength (predictor importance 0.371 and $p = .001$). Participants with slight effortful exertion created significantly better passwords than the undepleted control group.

The diminished capacity for strong password choices under depletion condition indicates that cognitive effort is necessary for the creation of strong passwords. It is however surprising that slight exertion of cognitive effort prior to the password creation leads to stronger passwords. This research is the first in-depth analysis of the impact of cognitive effort and depletion on password creation. Our findings open up new avenues for usable security research.

II. Method

The sample consisted of university students, $N = 100$, mostly non-computer science background, of which 50 were women. The mean age was 28.18 years ($SD = 5.241$) for the 83 participants who revealed their age.

The experiment group was artificially cognitively depleted while the control group was not depleted, completing non-depleting tasks with similar length and flavour. Cognitive depletion was induced via (a) a thought suppression task where participants were shown the photo of a white bear and asked not to think of the white bear [2], (b) an impulse control task where participants first learn a habit of cutting out all letter 'e' in a statistical text, followed by a new rule of crossing out all letters 'e' unless they are adjacent to a vowel, (c) a cognitively effortful task using Stroop [3] where the name of a color (e.g., 'red') is printed in a color not denoted by the name (incongruent color and name).

The procedure consisted of (a) pre-task questionnaires for demographics and personality traits (Big Five Inventory, BFI), (b) a manipulation to induce cognitive depletion, (c) a manipulation check on the level of depletion, (d) a password entry for a mock-up GMail registration, and (e) a debriefing and memorability check one week after the task with a GMail login mockup. We depict the procedure in Figure 1.

III. Results

We use two-tailed tests at an alpha level of .05.

A. Manipulation Check

Because induced cognitive depletion can differ across participants [4], we use the short form of a brief mood inventory (BMI) as manipulation check. We follow a method of Baumeister's research [5], including the dimensions (a) excited, (b) thoughtful, (c) tired, (d) happy, (e) worn out, (f) sad, (g) angry, (h) calm, rated on 5-point Likert scale from strongly disagree to strongly agree. A comparison across groups on tired and worn out suggested that the manipulation was successful (Mann-Whitney U, two-tailed, tired: $U = 368, Z = -6.299$, significance $p = .000 < .05$; wornout: $U = 669$, $Z = -4.145$, significance $p = .000 < .05$). The effect size of the manipulation for feeling tired is large ($r = 0.63$) and for feeling worn out is medium to large ($r = 0.42$).

B. Password Strength Score

The impact on the password strength score was analyzed with a multi-predictor forward stepwise linear regression.

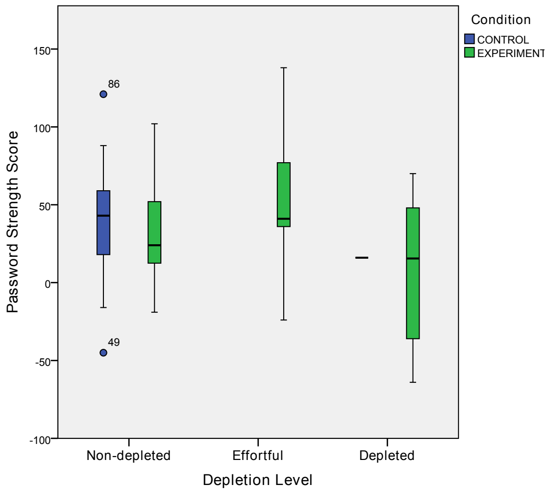

Figure 2. Password strengths by experiment condition and depletion level.

The linear regression accounted for 20.6% of the variability (adjusted $R^2 = .206$) corresponding to an effect size measured in Cohen's f^2 of 0.26. This is a medium to large effect.

1) Effect of Depletion Level: The automated data preparation of the regression grouped tiredness into three classes: (a) disagree and neither agree nor disagree, (b) slightly agree, (c) strongly disagree. We call this classification *depletion level* and label the classes as (a) non-depleted, (b) effortful, and (c) depleted.

The depletion level was the most important predictor in the regression (significance $p = .001 < .05$, importance$= 0.371$). The effortful level, that is only slightly depleted, had a coefficient of 50.65 (significance $p = .000 < .05$). The non-depleted level had a coefficient of 31.62 (significance $p = .006 < .05$). We summarize the descriptive statistics of password strength score by depletion level in Table I and depict them in Figure 2.

A one-way ANOVA gave a significant effect of the depletion level on the password strength, $F(2, 97) = 5.449$, $p = .006 < .05$. Because of the unequal sample sizes, we use Scheffé and Games-Howell as robust post-hoc tests. Both Scheffé and Games-Howell reported the depleted case significantly different from the effortful case, but neither of them significantly different from the non-depleted case.

2) Effects of Mood.: BMI Thoughtfulness and Calmness had significant effects. Strong disagreement to thoughtfulness implied stronger passwords (significance $p = .018 < .05$, importance $= 0.251$, coefficient 40.072). Strong disagreement to calmness implied stronger passwords (significance $p = .012 < .05$, importance $= 0.172$, coefficient 38.799).

3) Effects of Personality Traits.: Of the Big Five personality traits, the BFI Agreeableness score was the most important predictor on the password strength (significance $p = .025 < .05$, importance $= 0.137$, oefficient 14.649), where higher agreeableness significantly implied stronger passwords. The BFI Extraversion was a notable yet non-significant negative predictor on password strength (significance $p = .108 > .05$, importance $= 0.069$, coefficient -11.538).

Table I
PASSWORD STRENGTH BY CONDITION AND DEPLETION LEVEL.

Condition	Depletion Level	N	Mean	Std. Dev.
Control	non-depleted	49	40.65	30.97
	depleted	1	16.00	-
	Total	50	40.16	30.87
Experiment	non-depleted	23	33.3	33.81
	effortful	17	57.12	45.07
	depleted	10	11.10	45.97
	Total	50	36.96	42.99
Total		100	38.56	37.27

IV. DISCUSSION

We applied the methodology of previous depletion studies [1], [6], [5] to a ubiquitous security context. We observe that strong depletion leads to weaker passwords while mild cognitive exertion leads to stronger passwords than undepleted states. This finding corresponds with Kahneman's observation that initial effort during an activity introduces a bias towards exerting further cognitive effort [7] and also with Selye's arousal curve, an inverse U-shaped relation showing optimum human performance under moderate stress.

V. CONCLUSION

This study is the first to show that cognitive effort is necessary for the creation of strong passwords, that cognitive depletion implies weaker passwords than in an undepleted state and that mild cognitive effort implies better passwords than in an undepleted state. Our results ask for a deeper investigation of human dimensions of usable security, especially user traits and current cognitive and affective state. They suggest new design paradigms for password policies and HCI interventions to support the user's password creation.

ACKNOWLEDGMENT

This work was supported by the EU FP7 Project FutureID (GA nº318424) and the EPSRC Research Institute in Science of Cyber Security (RISCS; EP/K006568/1 ChAISe).

REFERENCES

[1] R. F. Baumeister, K. D. Vohs, and D. M. Tice, "The strength model of self-control," *Current directions in psychological science*, vol. 16, no. 6, pp. 351–355, 2007.

[2] D. M. Wegner, D. J. Schneider, S. R. Carter, and T. L. White, "Paradoxical effects of thought suppression." *Journal of personality and social psychology*, vol. 53, no. 1, p. 5, 1987.

[3] J. R. Stroop, "Studies of interference in serial verbal reactions." *Journal of experimental psychology*, vol. 18, no. 6, p. 643, 1935.

[4] V. Job, C. S. Dweck, and G. M. Walton, "Ego depletion is it all in your head? implicit theories about willpower affect self-regulation," *Psychological science*, 2010.

[5] D. M. Tice, R. F. Baumeister, D. Shmueli, and M. Muraven, "Restoring the self: Positive affect helps improve self-regulation following ego depletion," *Journal of Experimental Social Psychology*, vol. 43, no. 3, pp. 379–384, 2007.

[6] R. Baumeister, E. Bratslavsky, E. Muraven, and D. Tice, "Ego depletion: is the active self a limited resource?" *Personality and social psychology*, vol. 74, pp. 1252–1265, 1998.

[7] D. Kahneman, *Thinking fast and slow*. Farrar, Strauss, 2011.