Influences of Mixed Reality and Human Cognition on Picture Passwords: An Eye Tracking Study

Christos Fidas[1], Marios Belk[2,3], George Hadjidemetriou[3], Andreas Pitsillides[3]

[1]Department of Cultural Heritage Management and New Technologies
University of Patras, Patras, Greece
`fidas@upatras.gr`
[2]School of Sciences, University of Central Lancashire, Cyprus Campus, Larnaca, Cyprus
`mbelk1@uclan.ac.uk`
[3]Department of Computer Science, University of Cyprus, Nicosia, Cyprus
`ghadji12@cs.ucy.ac.cy, andreas.pitsillides@cs.ucy.ac.cy`

Abstract. Recent research revealed that individual cognitive differences affect visual behavior and task performance of picture passwords within conventional interaction realms such as desktops and tablets. Bearing in mind that mixed reality environments necessitate from end-users to perceive, process and comprehend visually-enriched content, this paper further investigates whether this new interaction realm amplifies existing observed effects of individual cognitive differences towards user interactions in picture passwords. For this purpose, we conducted a comparative eye tracking study (*N*=50) in which users performed a picture password composition task within a conventional interaction context *vs.* a mixed reality context. For interpreting the derived results, we adopted an accredited human cognition theory that highlights cognitive differences in visual perception and search. Analysis of results revealed that new technology realms like mixed reality extend and, in some cases, amplify the effect of human cognitive differences towards users' visual and interaction behavior in picture passwords. Findings can be of value for improving future implementations of picture passwords by considering human cognitive differences as a personalization factor for the design of user-adaptive graphical passwords in mixed reality.

Keywords: Picture Passwords; Human Cognition; Mixed Reality; Eye Tracking; Visual Behavior; Usability; Security.

1 Introduction

Mixed reality technologies are being continuously embraced by researchers and practitioners for developing immersive applications and services which favor multi-modal human computer interaction techniques like touch-, hand gesture- or gaze-based. Such technology advancements open unprecedented opportunities for designing new visually enriched interaction experiences for end-users in a variety of application domains that include healthcare, military training, aviation, interactive product management, remote working, games, etc. [29, 30, 31].

A cornerstone user activity in mixed reality environments is related to *user authentication*. User authentication is an act which aims at verifying that a user is who she claims to be and therefore has eligible rights to access sensitive information and services. Since mixed reality contexts introduce new challenges and opportunities for designing visually enriched user experiences, researchers have explored existing and alternative user authentication schemes (*e.g.*, pin, passwords, patterns, graphical) in mixed and virtual reality contexts, aiming to gain new knowledge on the interplay between human behavior, usability, and security in such schemes [1-5].

In this context, picture passwords, which require users to draw secret gestures on a background image to unlock a device or application, have been introduced as viable mixed reality user authentication schemes since they leverage on hand gesture interaction modalities. Figure 1 depicts Microsoft's Picture Gesture Authentication (PGA), a widely deployed picture password scheme that has been introduced in Windows 8^{TM} (and further deployed in Windows 10^{TM}) as a promising alternative login experience to text-based passwords. Picture passwords necessitate from humans to perform visual search and visual memory processing tasks, aiming to view, recognize and recall graphical information. Given that individuals differ in the way they perceive and process visual information [6-8], researchers have investigated the effects of human cognitive differences towards human behavior, experience and security of graphical passwords within conventional environments, such as desktop and mobile [9-12].

Fig. 1. Example of Microsoft Windows 10 PGA on a traditional desktop computer [18]. Users are required to draw three gestures on a background image to create their graphical password.

Research Motivation. Given the increased adoption of mixed reality technologies in a variety of application domains [13], we are motivated in investigating effects of human cognitive differences and mixed reality technology towards user's interaction and visual behavior within graphical password composition tasks. Such new knowledge would allow application designers to draw conclusions on the interplay between human cognitive and design factors of graphical passwords within mixed reality, and apply this knowledge for the provision of human cognitive-centered password experiences that are best-fit to each user's cognitive characteristics, and consequently assist visual information search and processing.

For doing so, we adopted an accredited human cognition theory and conducted a between-subjects eye tracking study (N=50) in which users performed a picture password composition task that was seamlessly deployed in mixed reality and traditional desktop contexts. To the best of our knowledge, this is amongst the first works which investigate the effects of human cognition and mixed reality picture password composition towards users' interaction and visual behavior.

2 Human Cognition Theory

We adopted Witkin's field dependence-independence theory (FD-I) [9, 14, 15] which suggests that humans have different habitual approaches, according to contextual and environmental conditions, in retrieving, recalling, processing and storing graphical information [8]. Accordingly, the theory distinguishes individuals as being field dependent and field independent. *Field dependent (FD)* individuals view the perceptual field as a whole, they are not attentive to detail, and not efficient and effective in situations where they are required to extract relevant information from a complex whole. *Field independent (FI)* individuals view the information presented by their visual field as a collection of parts and tend to experience items as discrete from their backgrounds. With regards to visual search abilities, studies have shown that FIs are more efficient in visual search tasks than FDs since they are more successful in dis-embedding and isolating important information from a complex whole [14, 15].

3 Method of Study

3.1 Null Hypotheses

***H_{01}*.** There is no interaction effect between FD-I differences and the technological context (desktop *vs.* mixed reality) towards time needed to create a picture password; *by investigating this research question we examine the effects of mixed reality's multimodal interaction capabilities towards task efficiency of FD-I users.*

***H_{02}*.** There is no interaction effect between FD-I differences and the technological context (desktop *vs.* mixed reality) towards users' visual behavior; *by investigating this research question we examine the effects of mixed reality's enriched visual content presentation capabilities towards gaze behavior of FD-I users.*

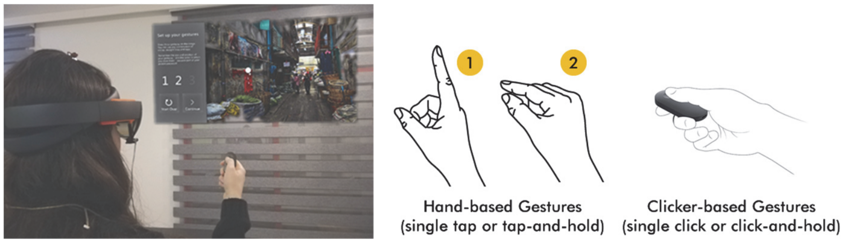
***H_{03}*.** There is no correlation between the time to create a picture password and visual behavior; *by investigating this research question we examine the interdependencies between FD-I users' interaction and visual behavior in mixed reality's environments.*

3.2 Research Instruments

Cognitive Factor Elicitation. Users' FD-I was measured through the Group Embedded Figures Test (GEFT) [16] which is a widely accredited and validated paper-and-pencil test [14, 15]. The test measures the user's ability to find common geometric shapes in a larger design. The GEFT consists of 25 items; 7 are used for practice, 18 are used for assessment. In each item, a simple geometric figure is hidden within a complex pattern, and participants are required to identify the simple figure by drawing it with a pencil over the complex figure. Based on a widely-applied cut-off score [14, 15], participants that solve 11 items and less are FD, while 12 items and above are FI.

Graphical Password Scheme. We developed a picture password mechanism, coined HoloPass, following guidelines of Microsoft Windows 10^{TM} Picture Gesture Authentication (PGA) [17] in which users draw passwords on a background image that acts as a cue (Figure 2-left). Implementation details and suitability of HoloPass is reported in [18]. Three gestures were implemented, *i.e.*, *dot*, *line*, *circle* which can be achieved through *hand-based gestures* or *clicker-based gestures* (Figure 2-right). For each gesture, the following data are stored: for dots, the coordinates of the point, for lines the coordinates of the starting and ending point, and for circles the coordinates of the point's center, radius and direction.

Interaction Devices. The picture password scheme was deployed on a conventional desktop computer and a mixed reality device. The desktop computer was a typical PC, with Intel core i7, 8GB RAM, 21-inch monitor, standard keyboard/mouse. For mixed reality we used Microsoft HoloLens which is a popular and widely adopted head mounted display for mixed reality, and features see-through holographic lenses. To measure the users' visual behavior and fixations, we have used and integrated Pupil Labs' eye tracker [19] in HoloLens using Pupil Labs' Binocular Add-on.



Hand-based Gestures
(single tap or tap-and-hold)
Clicker-based Gestures
(single click or click-and-hold)

Fig. 2. A user interacting with HoloPass that resembles PGA in mixed reality (left); and types of user input through hand gestures or using the HoloLens clicker (right) [18].

3.3 Sampling and Procedure

A total of 50 individuals (10 females) participated in the study, ranging in age from 18 to 40 (*m*=24.46; *sd*=3.58). Based on their scores on the GEFT; 24 participants (48%) were FD; 26 participants (52%) were FI. No participant was familiar with picture passwords and all had no or limited prior experience with mixed reality devices. The study involved the following steps: *i)* participants were informed that the collected data would be stored anonymously for research purposes, and they signed a consent form; *ii)* they were familiarized with the picture password and equipment, following an eye-calibration process; *iii)* participants then created a picture password to unlock a real service in order to increase ecological validity; and finally *iv)* they were asked to log in to ensure that the passwords were not created at random.

3.4 Data Metrics

For interaction behavior we measured time required to create the picture password which started as soon the user was shown with the task until the user successfully completed the password creation. For visual behavior we used the following measures: *i)* fixation count and duration; and *ii)* transition entropy [25] between Areas of Interests (AOIs) which measures the lack of order aiming to capture eye movement variability.

4 Analysis of Results

In the analysis that follows, data are mean ± standard error. Residual analysis was performed, outliers were assessed by inspection of a boxplot, normality was assessed using Shapiro-Wilk's normality test for each cell of the design and homogeneity of variances was assessed by Levene's test. There were no outliers, residuals were normally distributed and there was homogeneity of variances.

4.1 Password Creation Time Differences

To investigate H_{01}, we ran a two-way ANOVA to examine the effects of FD-I and interaction context on graphical password creation time (Figure 3-left). There was a significant effect of FD-I on the time to create the picture password in both interaction context, $F(1, 50)=4.846$, $p=.033$, $partial\ \eta^2=.095$. FD users spent significantly more time to create a picture password than FI users, in both interaction contexts (FD-Desktop: 37.25±19.34; FD-HoloLens: 29.16±14.29; FI-Desktop: 26.28±13.78; FI-HoloLens: 17.87±12.22). An analysis across groups (FD and FI) revealed that mixed reality interactions were completed faster in both groups compared to desktop contexts.

4.2 Visual Behavior Differences

To investigate H_{02}, a two-way MANOVA was run with two independent variables (FD-I and interaction context) and two dependent variables (fixation count and mean fixation duration). The combined fixation metrics were used to measure visual behavior. The interaction effect between FD-I and interaction context on the combined dependent variables was not statistically significant, $F(2, 45)=.745$, $p=.48$, $Wilks'\ \Lambda=.968$, $partial\ \eta^2=.032$. There was a statistically significant main effect of interaction context on the combined dependent variables, $F(2, 45)=13.302$, $p<.001$, $Wilks'\ \Lambda=.628$, $partial\ \eta^2=.372$. Follow up univariate two-way ANOVAs were run, and the main effect of intervention considered. There was a statistically significant main effect of interaction context for fixation duration, $F(1, 50)=24.640$, $p<.001$, $partial\ \eta^2=.349$, but not for fixation count, $F(1, 50)=.722$, $p=.4$, $partial\ \eta^2=.015$. As such, Tukey pairwise comparisons were run for the differences in mean fixation duration between interaction contexts. The marginal means for fixation duration were 981.38 ± 35.42 for desktop interactions, and 732.7 ± 35.42 for mixed reality interactions. For FD users, there was a statistically significant mean difference between the desktop-based fixation duration and the mixed reality fixation duration of -230.73 (95% CI, -376.16 to 85.3), *p*=.003, while for FI users the difference was -266.61 (95% CI, 406.34 to 126.89), *p*<.001.

We further ran a two-way ANOVA to examine the effects of FD-I and interaction context on transition entropy (Figure 3-right). There was a significant effect of FD-I on transition entropy, *F*(1, 50)=27.089, *p*<.001, *partial* η^2=.371. FD users had significantly higher transition entropy than FI users since they had higher randomness and variability in their visual behavior. There was also a significant effect of interaction context on transition entropy, *F*(1, 50)=5.259, *p*=.027, *partial* η^2=.102 with mixed reality interaction triggering higher transition entropies than conventional interaction contexts.

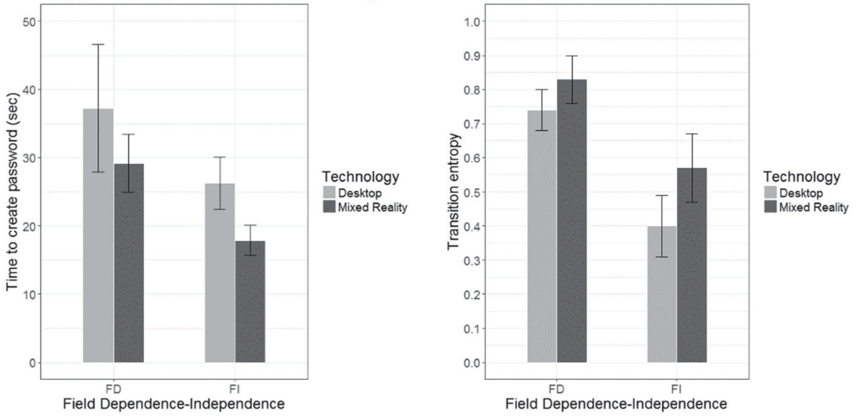


Fig. 3. Time to create (left) and transition entropy (right) per user group.

4.3 Correlation between Time to Create a Picture Password and Visual Behavior

To investigate H_{03}, we performed a Pearson's Product Moment correlation test, between time to create the password and transition entropy (Figure 4). The analysis revealed a strong positive correlation between creation time and transition entropy for desktop interactions (*r*=.505, *p*=.01) as well as for mixed reality interactions (*r*=.438, *p*=.028). The higher the transition entropy, the more disordered the visual behavior is. These results explain the previous analyses, since FD users spent significantly more time and had higher transition entropies than FI users.

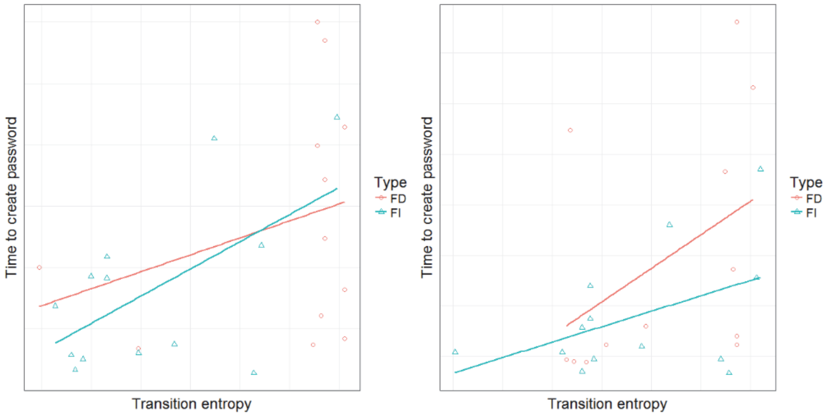


Fig. 4. Scatter-plots depicting creation time of passwords and transition entropy for desktop interactions (left) and mixed reality interactions (right).

5 Interpretation of Results

Interpretation with regards to H_{01}. Mixed reality scaffolded more efficient graphical password task execution for both user groups (FD and FI) compared with the desktop context. A between cognitive factor analysis revealed that within mixed reality, FI users were significantly faster than FD users. This can be explained due to FI users' positive adaptation and independence in regards with contextual and field changes (desktop *vs.* mixed reality). This finding suggests that the device, and eventually the field change, towards mixed reality interactions (context-wise and interaction-wise) was adopted more efficiently and effectively by FI users compared to FD users. This further supports previous findings which state that FD users depend on their surrounding field whereas FI users are not significantly influenced by their surrounding field and context of use [24, 26, 27]. Furthermore, this finding can also be explained by the fact that FD users follow a more holistic and exploratory approach during visual search compared to FI users that primarily focus on specific focal points of an image during interaction. Based on qualitative feedback, the increased amount of time for FDs did not negatively affect their interaction experience.

"I was excited to draw a password on an image. At first, I spent some time to view the whole content and then I made my selections" ~ P20 - FD individual

"It is much easier to draw my password than using the virtual keyboard. I created my password in no time by selecting the people in the image" ~ P24 - FI individual

Interpretation with regards to H_{02}. The interaction context has a main effect on the fixation duration during picture password composition. Users in the mixed reality interaction context fixated longer on areas of the image than users during desktop-based interactions. With regards to transition entropy, results revealed significant differences among FD and FI users. Specifically, FD users had significantly higher transition entropies (higher randomness in eye movements) compared to FI users. Hence, these observable differences in eye gaze behavior among FD and FI users allows to better explain the previous finding related to task completion efficiency.

"The most difficult part was finding where to draw the gestures, but I believe that adds up to the security of the password" ~ P15 - FD individual

"It is a more creative way to create a password and escapes the dullness of the keyboard" ~ P30 - FI individual

Interpretation with regards to H_{03}. A strong positive correlation between password creation time and transition entropy was revealed which further supports Finding 2 and Finding 3. The higher the transition entropy, the more disordered the visual behavior is. These results explain the previous analyses, since FD users spent significantly more time and hence triggered higher transition entropies compared to FI users.

"I checked out the whole image to see all the items. I tried to avoid objects that were obvious for someone to guess my password so I tried to find less obvious objects to select" ~ P33 - FD individual

"I focused on specific objects and made my selections" ~ P42 - FI individual

6 Conclusions

This paper revealed underlying effects between individual cognitive differences and mixed reality interaction realms towards users' eye gaze behavior and task execution during picture password composition tasks. Analysis of eye-tracking data further validated that user's individual differences of visual information perception and processing are reflected by their eye gaze behavior in both conventional and mixed reality interaction realms, but with a stronger effect within mixed reality interaction contexts. As such, the enriched visual content presentation of mixed reality environments has a rather catalyst effect, in terms of visual content exploration and task execution, for FD users than FI users. A comparative analysis between the conventional and mixed reality interaction contexts revealed that the technology shift towards a visually enriched content presentation triggered FD users to explore longer and comprehensively the content. Hence, FD users spent more time and produced longer fixation durations and transition entropies within mixed reality environments when compared to FI users.

Bearing in mind that transition entropies of users have been correlated with security strength of graphical passwords [9, 28] such findings can be of value for mixed reality researchers and experience designers for considering: *a)* users eye gaze patterns as early predictors of password security strength [28]; and *b)* human cognitive characteristics as important design factors in picture password schemes [9, 24, 34]. We anticipate that this work will inspire similar research endeavors (*e.g.*, see the approaches discussed in [9, 10, 23, 24, 32, 33] on how human factors can be incorporated in personalized user authentication schemes) aiming to incorporate novel authentication schemes based on eye tracking methods and users' eye gaze patterns.

Acknowledgements

This research has been partially supported by EU Horizon 2020 Grant 826278 “Securing Medical Data in Smart Patient-Centric Healthcare Systems” (Serums). We thank all participants for their time and valuable comments provided during the studies.

References

1. Yu, Z., Liang, H.N., Fleming, C., & Man, K.L. (2016). An exploration of usable authentication mechanisms for virtual reality systems. In IEEE APCCAS 2016, IEEE, 458-460.
2. Roesner, F., Kohno, T., & Molnar, D. (2014). Security and privacy for augmented reality systems. Commun. ACM 57, 4, 88-96.
3. George, C., Khamis, M., Zezschwitz, E.V., Burger, M., Schmidt, H., Alt, F., & Hußmann, H. (2017). Seamless and secure vr: Adapting and evaluating established authentication systems for virtual reality. In USEC 2017.
4. Yadav, D. K., Ionascu, B., Ongole, S. V. K., Roy, A., & Memon, N. (2015). Design and analysis of shoulder surfing resistant PIN based authentication mechanisms on google glass. In Financial Cryptography and Data Security 2015, Springer Verlag.
5. Schneegaß, S., Oualil, Y., & Bulling, A. (2016). SkullConduct: Biometric user identification on eyewear computers using bone conduction through the skull CHI ‘16, ACM, 1379-1384.
6. Mayer, R.E., Massa, L.J., 2003. Three facets of visual and verbal learners: cognitive ability, cognitive style, and learning preference. J. Edu. Psychol. 95 (4), 833–846
7. Riding, R., Cheema, I., 1991. Cognitive styles –an overview and integration. Edu. Psychol. 11 (3–4), 193–215
8. Witkin, H.A., Moore, C.A., Goodenough, D.R., Cox, P.W., 1975. Field-dependent and field-independent cognitive styles and their educational implications. Res. Bull. Ser. (2), 1–64
9. Katsini, C., Fidas, C., Raptis, G., Belk, M., Samaras, G., & Avouris, N. (2018). Influences of human cognition and visual behavior on password security during picture password composition. ACM SIGCHI Human Factors in Computing Systems (CHI 2018), ACM, paper 87
10. Belk, M., Fidas, C., Katsini, C., Avouris, N., Samaras, G. (2017). Effects of Human Cognitive Differences on Interaction and Visual Behavior in Graphical User Authentication. INTERACT (3) 2017: 287-296
11. Belk, M., Germanakos, P., Fidas, C., Samaras, G. (2014). A personalisation method based on human factors for improving usability of user authentication tasks. User Modeling, Adaptation, and Personalization (UMAP 2014), 13-24
12. Raptis, G., Fidas, C.A., Avouris, N. (2018). Effects of mixed-reality on players' behaviour and immersion in a cultural tourism game: A cognitive processing perspective, International Journal of Human-Computer Studies, Elsevier, 69-79
13. Bellini, H., Chen, W., Sugiyama, M., Shin, M., Alam, S., Takayama, D., 2016. Virtual & Augmented Reality: Understanding the Race for the Next Computing Platform. Technical Report. The Goldman Sachs Group
14. Angeli, C., Valanides, N., & Kirschner, P. (2009). Field dependence-independence and instructional-design effects on learners' performance with a computer-modeling tool. Computers in Human Behavior, 25(6), 1355-1366.
15. Hong, J., Hwang, M., Tam, K., Lai, Y., & Liu, L. (2012). Effects of cognitive style on digital jigsaw puzzle performance: A GridWare analysis. Comp. in Hum. Beh., 28(3), 920-928.
16. Oltman, P.K., Raskin, E., Witkin, H.A., 1971. GEFT. Consulting Psychologists Press, USA.

17. Johnson, J.J., Seixeiro, S., Pace, Z., van der Bogert, G., Gilmour, S., Siebens, L., & Tubbs, K. (2014). Picture gesture authentication. https://www.google.com/patents/US8910253
18. Hadjidemetriou, G., Belk, M., Fidas, C., & Pitsillides, A. (2019). Picture passwords in mixed reality: Implementation and evaluation (2019). In ACM CHI '19 Extended Abstracts, ACM Press.
19. Kassner, M., Patera, W., & Bulling, A. (2014). Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. In ACM UbiComp 2014, 1151-1160.
20. Zhao, Z., Ahn, G., Seo, J., Hu, H. (2013). On the security of picture gesture authentication. In USENIX Security 2013. USENIX Association, 383-398.
21. Zhao, Z., Ahn, G., & Hu, H. (2015). Picture gesture authentication: Empirical analysis, automated attacks, and scheme evaluation. ACM Trans. Inf. Syst. Secur. 17, 4, 37 pages.
22. Biddle, R., Chiasson, S., & van Oorschot, P. (2012). Graphical passwords: Learning from the first twelve years. ACM Computing Surveys, 44(4), 41.
23. Liu, D., Dong, B., Gao, X., & Wang, H. (2016). Exploiting eye tracking for smartphone authentication. Applied Cryptography and Network Security, 457-477.
24. Belk M., Fidas C., Germanakos P., & Samaras G. (2017) The interplay between humans, technology and user authentication: a cognitive processing perspective, Computers in Human Behavior, 76, 184-200.
25. Krejtz, K. et al. 2015. Gaze Transition Entropy. ACM Trans. on Appl. Perc. 13, 1, 1-20.
26. Davis, J. K. (1991). Educational implications of field dependence–independence, 149-175.
27. Messick, S. (1993). The matter of style: Manifestations of personality in cognition, learning, and teaching. Princeton, NJ: Educational Testing Service.
28. Katsini, C., Raptis, G. E., Fidas, C., & Avouris, N. (2018). Towards gaze-based quantification of the security of graphical authentication schemes. In ACM ETRA '18, ACM press
29. Bogle, A. (2016). ebay launches a world-first virtual reality department store. https://mashable.com/2016/05/18/ebay-virtual-reality-shopping.
30. Medenica, Z, Kun, L., Paek, T., Palinko, O. (2011). Augmented reality vs. street views: a driving simulator study comparing two emerging navigation aids. MobileHCI '11, 265-274.
31. Kim, S. and Dey, A.K. (2009). Simulated augmented reality windshield display as a cognitive mapping aid for elder driver navigation. In ACM CHI '09, 133-142.
32. Constantinides, A., Belk, M., Fidas, C., & Samaras, G. (2018). On cultural-centered graphical passwords: Leveraging on users' cultural experiences for improving password memorability. ACM UMAP '18, ACM Press, 245-249.
33. Belk, M., Pamboris, A., Fidas, C., Katsini, C., Avouris, N., & Samaras, G. (2017). Sweet-spotting security and usability for intelligent graphical authentication mechanisms. ACM WI '17, ACM Press, 252-259.
34. Katsini, C., Fidas, C., Raptis, G., Belk, M., Samaras, G., & Avouris, N. (2018). Eye gaze-driven prediction of cognitive differences during graphical password composition. ACM IUI '17, ACM Press, 147-152.