University of Central Lancashire
UCLan

Central Lancashire Online Knowledge (CLoK)

Title	A Human-Cognitive Perspective of Users' Password Choices in Recognition-Based Graphical Authentication
Type	Article
URL	https://clok.uclan.ac.uk/id/eprint/26170/
DOI	https://doi.org/10.1080/10447318.2019.1574057
Date	2019
Citation	Katsini, Christina, Fidas, Christos, Belk, Marios, Samaras, George and Avouris, Nikolaos (2019) A Human-Cognitive Perspective of Users' Password Choices in Recognition-Based Graphical Authentication. International Journal of Human–Computer Interaction, 35 (19). pp. 1800-1812. ISSN 1044-7318
Creators	Katsini, Christina, Fidas, Christos, Belk, Marios, Samaras, George and Avouris, Nikolaos

It is advisable to refer to the publisher's version if you intend to cite from the work.
https://doi.org/10.1080/10447318.2019.1574057

For information about Research at UCLan please go to http://www.uclan.ac.uk/research/

All outputs in CLoK are protected by Intellectual Property Rights law, including Copyright law. Copyright, IPR and Moral Rights for the works on this site are retained by the individual authors and/or other copyright owners. Terms and conditions for use of this material are defined in the http://clok.uclan.ac.uk/policies/

RESEARCH PAPER

A Human Cognitive Perspective of Users' Password Choices in Recognition-based Graphical Authentication

Christina Katsini[1], Christos Fidas[2], Marios Belk[3], † George Samaras[4], Nikolaos Avouris[1]

[1] HCI Group, Interactive Technologies Lab, Department of Electrical and Computer Engineering, University of Patras, Greece
[2] Department of Cultural Heritage Management and New Technologies, University of Patras, Greece,
[3] School of Sciences, University of Central Lancashire, Cyprus,
[4] Department of Computer Science, University of Cyprus, Nicosia, Cyprus

ARTICLE HISTORY
Compiled December 4, 2018

ABSTRACT
Graphical password composition is an important part of graphical user authentication which affects the strength of the chosen password. Considering that graphical authentication is associated with visual search, perception, and information retrieval, in this paper, we report on an eye-tracking study (N=109) that aimed to investigate the effects of users' cognitive styles towards the strength of the created passwords and shed light into whether and how the visual strategy of the users during graphical password composition is associated with the passwords' strength. For doing so, we adopted Witkin's Field Dependence-Independence theory, which underpins individual differences in visual information and cognitive processing, as graphical password composition tasks are associated with visual search. The analysis revealed that users with different cognitive processing characteristics followed different patterns of visual behavior during password composition which affected the strength of the created passwords. The findings underpin the need of considering human cognitive characteristics as a design factor in graphical password schemes. The paper concludes by discussing implications for improving recognition-based graphical passwords through adaptation and personalization techniques based on individual cognitive characteristics.

KEYWORDS
Usable Security; Knowledge-based User Authentication; Recognition-based Graphical Passwords; Password Strength; Password Creation Strategy; Visual Behaviour; Eye-tracking; Human Factors; Cognitive Styles; Field Dependence-Independence

1. Introduction

Graphical user authentication (GUA) has been proposed as an alternative to alphanumeric passwords for gaining access to systems and services, with Android Pattern Unlock and Windows Picture Passwords being two commercial examples of such schemes. GUA entails users authenticating with the use of an image input and the aid of a mouse/stylus, a touch screen, or gestures. GUA schemes lie under two categories, depending on which memory function is triggered when they are being used; *recall* and *recognition*. *Recall-based* GUA schemes require users to remember and reproduce a drawing they have drawn before, often with the use of a background image as a cue. Examples of recall-based GUAs are DAS (Jermyn et al., 1999), Pass-Go (Weinshall et al., 2004), and PassPoints (Wiedenbeck et al., 2005). *Recognition-based* GUA schemes require users to identify and select target images from a challenge set. Examples of recognition-based GUAs are PassFaces (Brostoff and Sasse, 2000), DejaVu (Dhamija and Perrig, 2000), and ImagePass (Mihajlov and Jerman-Blažič, 2011).

GUA schemes have been proposed aiming to improve the usability and memorability issues associated with the alphanumeric schemes (Melicher et al., 2016), which have raised security issues since people have adopted workarounds for remembering their passwords such as reuse of passwords or parts of their passwords or use of very simple words (e.g., *password*, *123456*) which constitute them easily guessable. Despite the fact that graphical passwords are easier to remember, people still make predictable choices when using GUA schemes which affect the strength of the created graphical passwords (Salehi-Abari et al., 2008; Thorpe and Van Oorschot, 2007; Zhao et al., 2015) and introduce security vulnerabilities. User choices are affected by several factors related to the characteristics of the individuals and the characteristics of the GUA scheme.

Focusing on the characteristics of the individuals, researchers have investigated the effect of gender (Mihajlov et al., 2016) and age (Nicholson et al., 2013). However, GUA is an information processing task, and thus, individual cognitive characteristics are expected to have an impact on the way the users create their graphical passwords. Regarding human cognition, current research is limited to investigating memorability issues in the login phase (Belk et al., 2017a). However, considering that the strength of the chosen graphical password is determined during password composition, it is important to investigate the effect of individual cognitive differences in the registration phase. In this regard, Katsini et al. (2018b) revealed that human cognitive styles influence the graphical password choices when using a recall-based GUA scheme.

Considering the results of the study on the recall-based GUA scheme (Katsini et al., 2018b) and given that no study has been yet performed to investigate users' choices in the password composition phase of recognition-based GUA from a human cognitive perspective, in this paper we adopt such a perspective and investigate how users react to stimuli, through analysing their visual behaviour, aiming to understand how they decide on the graphical passwords they create. This will allow us to draw conclusions on the effects of cognitive styles on the strength of the created passwords, understand whether and how these are associated with the users' visual behaviour, and investigate whether this knowledge could be used to design adaptive mechanisms that would enable users to create stronger passwords.

The remainder of the paper is structured as follows: firstly, we discuss the related research. Then, we present a user-study we conducted to investigate the interplay effects between the users' cognitive style and the strength of the created passwords. Next, we present the main findings of the research endeavour, along with their implications. Finally, we discuss the study validity and conclude the paper.

2. Related Work

Considering that the password strength is associated with predictable user choices which are interrelated with the memorability of the created passwords, researchers have attempted to investigate which factors affect the user choices. They are affected, among others, by human factors, such as gender, age and cultural background (Mihajlov and Jerman-Blažič, 2011; Nicholson et al., 2013), and GUA scheme characteristics, such as grid size and image type (Belk et al., 2017b; De Luca et al., 2009), grid layout (Katsini et al., 2018c), and design space (Stobert and Biddle, 2013).

A major parameter that influences the password strength is the type of images used in the image grid. People make predictable choices when using images of people as they tend to select images of people with similar cultural background to theirs (Davis et al., 2004). When using images of single objects, people tend to select images of business, food and drinks, plants, and games (Mihajlov et al., 2016). Thorpe et al. (2014) revealed that password choices are affected by the use of colors, shapes, patterns, and/or letters for remembering the order of the selection. The results are supported by Mihajlov et al. (2016) who showed that people select images based on the shapes and the colors they contain. Thorpe et al. (2014) also showed that a simple effect such as gradually revealing the images of an image grid affects the users' password choice and Katsini et al. (2018c) showed that the image grid visualization (2D versus 3D layouts) affects the strength of the created passwords.

From a human cognitive factor perspective, Chiasson et al. (2009) investigated the effects of memory interference, an elementary cognitive process, on text-based and graphical passwords. They focused on password interference when using multiple passwords and their research revealed that people make significantly fewer errors when using graphical passwords than text-based passwords. Similar results were revealed by Meng et al. (2016), whose research revealed that individuals who used the pattern unlock mechanism outperformed those using the text-based mechanism, both in the short-term and in the long-term. In another research, Everitt et al. (2009) investigated the effects of password frequency, interference, and training. They found that participants who accessed four different infrequent graphical passwords each week had a failure rate more than ten times greater than participants accessing a single infrequent password. Stobert and Biddle (2013) explored the impact of different types of memory retrieval (free-recall, cued-recall, and recognition-based) on the usability and memorability of randomly assigned graphical passwords. Their study showed that cued-recall was better than free-recall and that recognition-based graphical passwords were more memorable than recall-based passwords. Huestegge and Pimenidis (2014) analysed the impact of memory load in terms of number of faces to be remembered and retention interval (hour, day, week, month) on performance in login tasks. They found no evidence for strong performance deterioration when memory load increased. Longer retention intervals yielded a slight increase in login failures, but post-hoc analyses indicated that this was mainly due to the worse performance of few individuals in the upper part of the age distribution of their sample. Search times increased with longer retention interval, probably reflecting decay or retrieval problems in long-term memory. Regarding elementary cognitive processing abilities, Belk et al. (2013b) reported that users with enhanced working memory and cognitive processing speed outperformed users with more limited processing abilities during textual and graphical login tasks.

Belk et al. (2017a) investigated the effects of cognitive processing styles towards users' login performance and memorability of alphanumeric and graphical passwords.

Their research revealed that users with individual differences in cognitive processing styles (i.e., verbals vs. imagers, field-dependent vs. field-independent) perform significantly different during login with alphanumeric and graphical passwords. Imagers outperformed Verbals (Belk et al., 2013a) and field-independent individuals outperformed field-dependent individuals (Belk et al., 2017a) in login tasks when using a recognition-based GUA scheme. Katsini et al. (2018b) revealed that field-independent individuals created stronger passwords when using a recall-based GUA scheme than field-dependent individuals.

The aforementioned research suggests that the cognitive styles affect the users during login tasks and there is evidence that the cognitive styles are reflected on the visual behaviour patterns of individuals (Raptis et al., 2016b). Recent research in the user authentication domain reveals that the visual behaviour during the password creation is correlated with the password strength (Katsini et al., 2018b), enables attackers to perform guessing attacks by identifying hot spots (Thorpe and Van Oorschot, 2007) and can be used as gaze-based strength quantification metric for user-created graphical passwords (Katsini et al., 2018d).

2.1. Motivation

From the discussion on the related work we can deduce that (a) existing research focuses rather on the login task, and not on password composition and (b) there is no research work which investigates whether users' individual cognitive factors affect the strength of recognition-based graphical passwords and whether they are reflected on their visual behaviour during graphical password composition. Given the importance of perception, visual search, and visual information comprehension in GUA tasks, we argue that investigating the effect of users' cognitive styles in graphical password composition tasks can provide important insights about the value of considering human cognitive differences in different technological contexts as a human design factor, in both design and run time, aiming to offer personalized features to end-users based on their individual cognitive styles.

Convinced that the eye plays an important role in understanding people's intentions and strategies while performing graphical password composition activities, in this paper we adopt a cognitive theory interrelated with the ability of an individual to extract visual information in graphical scenes to investigate how people choose graphical passwords. The *Field Dependence-Independence (FD-I)* theory suggests that individuals have different habitual approaches, according to contextual and environmental conditions, in processing graphical information, and accordingly characterizes individuals as being either field-dependent (FD) or field-independent (FI). FD individuals tend to follow a more holistic approach to process visual information and have difficulties in identifying details in complex visual scenes (Witkin et al., 1975). FI individuals tend to follow a more analytic approach to process visual information, pay attention to details, and easily separate simple structures from the surrounding visual context (Witkin et al., 1975). Our aim is to shed light into how people select their graphical passwords, whether (and how) their visual strategy is associated with the passwords they select, and investigate how this knowledge could be used to drive the design of adaptive mechanisms that could affect people's graphical password choices.

3. Method of Study

We designed a study in which the participants were asked to use a recognition-based GUA scheme to create a graphical password by selecting images from a large image set. Considering that users nowadays mainly use desktop and mobile devices for web browsing and logging in to services, we decided to use both types of interaction devices aiming to increase the external validity of our study. We adopted a mixed-methods research approach, based on the analysis of both qualitative and quantitative data, aiming to examine whether cognitive differences during password creation affect the strength of the created passwords, whether these are reflected in the users' visual behaviour and understand their password creation strategy. Based on the literature and our motivation, we formed the following hypotheses:

H_1. FI individuals create stronger passwords than FD individuals.
H_2. The visual behavior of users with different cognitive styles (FD vs FI) is correlated to the strength of the passwords they create.
H_3. FI individuals have different visual behavior than FD individuals when creating a password.

To investigate the hypotheses, we performed a statistical analysis, discussed in the next section. We also performed a qualitative analysis to shed light into the users' behavior and thoughts when creating a graphical password. For the qualitative analysis, we used the eye-tracking recordings and an informal discussion with the users at the end of the session.

3.1. Study Instruments

In our study, we used the following study instruments: i) a recognition-based GUA scheme; ii) the GEFT test for classifying the participants into FD and FI groups; iii) the interaction devices, and iv) the eye-tracking apparatus.

3.1.1. Graphical User Authentication (GUA) scheme

We used the recognition-based GUA scheme designed by Belk et al. (2017a), which is depicted in Figure 1. This mechanism was designed and developed based on the guidelines of well-cited recognition-based GUA mechanisms, such as DejaVu (Dhamija and Perrig, 2000), PassFaces (Brostoff and Sasse, 2000), and ImagePass (Mihajlov and Jerman-Blažič, 2011). During user enrollment, the user is presented with a total of 120 images of objects in a single screen in a 9x14 image grid. The images cannot be sorted or replaced and no adding or removing images is permitted. Each image can only be used once in a graphical password. The use of single-object images was based on recent research revealing that these are more memorable than faces and abstract images (Chowdhury et al., 2013; Mihajlov and Jerman-Blažič, 2011). To log in, the user is provided with a set of 25 images which includes the 5 images selected by the user and a set of 20 decoy images. The decoy images are randomly selected after the password creation from the remaining images, they are shuffled with the 5 images of the password and they are randomly positioned on the login grid, counterbalancing the security aspect of the login stage (Belk et al., 2017a).

The selected GUA scheme policy requires a password of 5 ordered images and a large pool of selection images to ensure a comparable theoretical entropy to that of the commercially used 6-characters long password-based authentication mechanisms

(Katsini et al., 2016). The entropy is measured in bits and is calculated through the following equation (O'Gorman, 2003):

$$H_{max} = log_2 K_p$$

The password key space K_p refers to the range of all possible values of password combinations and is calculated by multiplying the character pool available for each of the password characters. The maximum theoretical entropy of the GUA used in our study is 34.41 bits.

Figure 1. The GUA mechanism used in our study. Users selected five unique images to form their graphical password.

3.1.2. GEFT

To classify the participants as either FD or FI we used the Group Embedded Figures Test (GEFT) by Oltman et al. (1971), which is the original classification FD-I tool. GEFT is a credible and validated time-administered instrument (Khatib and Hosseinpur, 2011), which measures the ability of an individual to identify a simple figure within a complex background. The test consists of three sections. In each section, the

individual is asked to identify and outline a given simple pattern in a visually complex context within a given amount of time. The first section consists of seven pattern-recognition problems and the individual has two minutes to solve them. The second and third section consist of nine pattern-recognition problems each, and the individual has five minutes to complete each section. A raw score is calculated by summing the correct answers in the last two sections, as the first section is used only for practice purposes. Thus, the score ranges between 0 and 18 and the individuals are classified as either FD or FI with the use of a cut-off score. In our study, the cut-off score was determined to be 11, as it has been widely used in the literature (Angeli et al., 2009; Hong et al., 2012; Raptis et al., 2016a). The cut-off score means that the participants who score 11 or lower would be classified as FD, and those who score between 12 and 18 as FI.

3.1.3. Interaction devices

The study was conducted using two interaction devices: a standard desktop computer and a mobile touch-screen device. The standard desktop computer had the following technical specifications: Intel core i7 with 8GB RAM, Windows 10 operating system, and an LG Monitor 22" 22MP48D at a screen resolution of 1920x1080 pixels (i.e., 16:9) with a frequency of 50Hz. The mobile touch-based device was a Samsung P1000 Galaxy Tab tablet computer with a 7.0" monitor at a screen resolution of 600x1024 pixels (i.e., 3:4).

3.1.4. Eye-tracking apparatus

The eye-tracking equipment used in the study was the Tobii Pro Glasses 2. The system captures data at 50Hz (i.e., the point of gaze was measured every 20 milliseconds). Fixations were extracted using a customized velocity threshold identification (I-VT) algorithm (Komogortsev et al., 2010), based on the I-VT algorithm provided by Tobii.

3.2. Metrics

3.2.1. Password strength metrics

To measure the generated graphical passwords' strength, we used two password metrics: password *guessability* and password *complexity*. To measure the password guessability, we used a brute-force approach which checked all possible combinations of graphical passwords comprising of five unique images starting from the upper left of the grid and traversing it row by row. We selected this for implementing the exhaustive password search based on research findings which revealed that when users are browsing a page of images they tend to scan the image grid in a horizontal pattern line by line (Djamasbi et al., 2011; Shrestha and Lenz, 2007). Our aim for selecting this approach was to reveal whether the relative position of the images on the image grid has an impact on the images that people with different cognitive styles use to create their password and whether this introduces security vulnerability to the GUA scheme. The password guessability was measured in terms of number of guesses needed to crack each password. The higher the number of guesses, the more difficult the password to crack. To measure the password complexity, we used the equation developed by Sun et al. (2014):

$$PS_P = S_p \times log_2 (L_p + I_p + O_p)$$

In the above equation, S_p is the size of the password (i.e., total number of images); L_p is the physical length of the password (i.e., the sum of the Euclidean distances between the selected images of the password); I_p is the total number of intersections (i.e., when two non-consecutive line segments have a common point); and O_p is the number of overlaps of the password pattern (i.e., when a line segment of the password pattern is covered by another segment). The higher the score, the more complex the password is.

3.2.2. Visual behavior metrics

In the present work we focused on visual behavior metrics which are related to visual decision-making tasks (Raptis et al., 2016b) and can contribute towards drawing conclusions for the strength of the created graphical passwords. Our visual behavior metrics are: the number of fixated images, the total fixation duration, and the Euclidean scanpath distance. The number of fixated images metric is used to measure how many unique images the user has fixated on during the password creation task. This metric is associated with the actual choice pool of the users (in parallel to the key-pool for the text-based authentication schemes), meaning that an image can only be chosen by the users, as part of their password, only if they have fixated on it. The total fixation duration metric is the total duration of fixations of an individual within the image grid, considering visits and revisits to the images, during the graphical password composition task. It is associated with the visual attention, as large duration reflects on high level of attention (Rayner, 1977). The Euclidean scanpath distance metric is the aggregation of the straight-line distance between two sequential fixated points of the image grid. It reflects on the difference between fixation sequence alignment distances and reveals the way a user searches for information, such as analytic style (Feusner and Lukoff, 2008). Typically, the visual behaviour metrics are performed on areas-of-interest (AOIs), which are clustered sub-regions of the displayed stimuli in which the eye-tracking metrics are applied (Steichen et al., 2014). In our study, each image of the GUA mechanism is an AOI.

3.3. Participants

A total of 109 individuals (50 females and 59 males) participated in our study. Their age ranged between 18 and 47 years ($M = 30.5, SD = 7.3$). Twenty-eight participants were undergraduate university students; thirty-six participants were postgraduate students; forty-five participants were professionals. Participants did not have any vision problems or had corrected to normal vision (i.e., wore glasses). The recruitment took place by communicating the research via social media, mailing lists, posting flyers on bulletin boards at various places on campus, and directly contacting acquaintances of the research team. To increase study validity, we aimed to recruit participants that had no experience with a recognition-based GUA mechanism to avoid any familiarity effects. Based on the users' GEFT scores, 53 participants were classified as FD and 56 participants as FI (GEFT score: $M = 11.27$, $SD = 3.51$, $min = 3$, $max = 18$). The scores of the participants were normally distributed according to Shapiro-Wilk's test ($p = .085$). Regarding the interaction device, the first half of the participants were

assigned to the desktop device and the rest half to the tablet. 54 participants used the desktop computer (26 FDs, and 28 FIs), and 55 participants used the tablet computer (27 FDs, 28 FIs).

3.4. Procedure

To increase the ecological validity of the users' interactions with the GUA mechanism, participants were asked to create a password that they would use for their email account since people tend to create more secure passwords when they are asked to secure their email account (Florencio and Herley, 2007). They were informed about the data that would be collected during the study session and that it would be stored anonymously and used only for research purposes by the research group, and they all provided their consent. No details in regards to the research objective were provided prior to the study, to avoid any bias effects. Participants were informed that they could opt out at any point of the study session.

Each participant visited our lab at a previously agreed date and time. The study was conducted in a quiet room in the lab. The procedure involved the following steps: first, the participants were introduced to the task and familiarized with the eye-tracking equipment. Then, the participants wore the eye-tracking equipment; participants wearing glasses were allowed to wear the eye-tracking equipment on top of their glasses. The eye-tracking equipment was calibrated based on the procedure described in the user-manual provided by Tobii AB (2017). Then, the participants were asked to look at the screen and create a graphical password, by creating a username in the first step and then selecting five images in a specific order from a set of 120 images using either a desktop or a tablet computer. The position of each image on the image grid remained the same for all participants during the registration phase. Given that participants had no prior experience with GUA schemes, instructions on the authentication policy were provided and participants were free to ask any questions before proceeding to the password creation. Next, the participants completed the GEFT test; the researchers calculated the GEFT score of the participants and classified them as either FD or FI based on the cut-off score, as discussed in GEFT section. To ensure participants did not create their passwords at random and increase the external validity of the study, they were asked to use the login details they have created to answer a short questionnaire on demographics. Finally, an informal discussion on the strategy of the users took place, after answering the questionnaire.

4. Results

4.1. The effect of FD-I on password strength

To investigate H_1, we performed a between-subjects two-way MANOVA test, with cognitive style (FD or FI) and interaction device (desktop or tablet computer) as the independent variables, and the combined security metric (i.e., guessability and complexity) as the combined dependent variable. The residual analysis for the guessability test revealed that the data was not normally distributed across all independent variables, and thus we transformed them using the two-steps approach proposed by Templeton (2011). After this transformation, the test met the required assumptions (i.e., linear relationship between the dependent variables; no evidence of multicollinearity; no univariate or multivariate outliers in the data; dependent variables data were

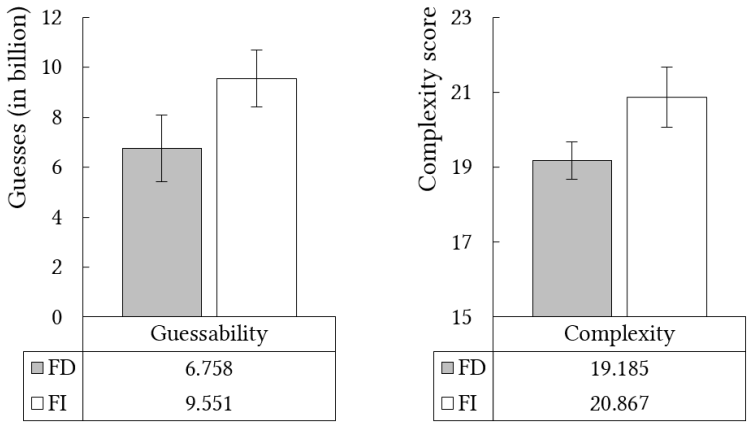

Figure 2. FIs created stronger (left) and more complex (right) passwords than FDs.

normally distributed; homogeneity of covariance matrices and variances).

The analysis revealed that FI individuals created stronger passwords, in terms of guessability and complexity, than FD individuals, $F_{2,104} = 3.582$, $p = .031$, $Wilks'$ $\Lambda = .936$, *partial* $\eta^2 = .064$. Follow-up univariate two-way ANOVAs were performed. There were statistically significant differences in the adjusted mean for both password guessability ($F_{1,105} = 4.144$, $p = .022$, *partial* $\eta^2 = .038$) and complexity ($F_{1,105} = 4.852$, $p = .015$, *partial* $\eta^2 = .044$). FIs' passwords required 2.768 billion more guesses (95% CI, 1.201 billion to 5.628 billion) to be cracked than FDs'. Moreover, FIs created 1.702 (95% CI, .170 to 3.228) more complex passwords than FDs. No effects of the interaction device were found. The results are depicted in Figure 2.

4.2. Association between password strength and visual behavior

To assess the relationship between the strength of the created passwords and the visual behavior (H_2) taking into consideration the influence of the FD-I cognitive style, we ran a series of partial Pearson correlations. All assumptions were met (i.e., linear relationship between the variables; no significant outliers; dependent variables data were normally distributed). Regarding the number of the fixated images, the analysis revealed a strong correlation with both password guessability ($r = .646$, $p = .012$) and complexity ($r = .685$, $p = .008$). Regarding the fixation duration, the analysis revealed a moderate correlation with password guessability ($r = .385$, $p = .031$) and a strong correlation with password complexity ($r = .545$, $p = .025$). Regarding the Euclidean scanpath distance, the analysis revealed a moderate correlation with both password guessability ($r = .301$, $p = .037$) and complexity ($r = .305$, $p = .033$). The correlation analysis results are summarized in Table 1. Therefore, the users' visual behavior influenced the strength of the created graphical passwords.

4.3. The effect of FD-I on visual behavior

To investigate H_3, we performed a between-subjects two-way MANOVA test, with cognitive style (FD or FI) and interaction device (desktop or tablet computer) as the independent variables and the combined eye-tracking metrics (i.e., number of fixated

Table 1. Pearson correlation analysis between the strength of the graphical passwords and the visual behavior metrics.

	Number of fixated images	Mean fixation duration per image	Euclidean scanpath distance
Password guessability	r = .646 p = .012	r = .385 p = .031	r = .301 p = .037
Password complexity	r = .685 p = .008	r = .545 p = .025	r = .305 p = .033

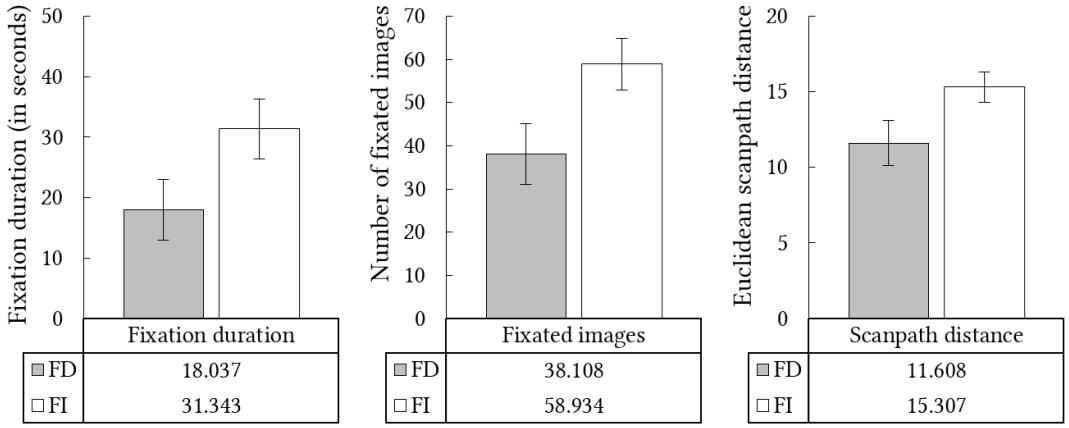

Figure 3. FIs fixated GUA for longer time (left), fixated more GUA images (middle), and produced longer scanpaths (right) than FDs.

images, total fixation duration, and Euclidean scanpath distance) as the dependent variables. To normalize the data across all independent variables, we followed the two-steps approach proposed by Templeton (2011). After the transformation, the test met the required assumptions.

The analysis of the results revealed that that FI individuals have different visual behaviour than FD individuals when creating a password, $F_{3,103} = 14.309$, $p < .001$, $Wilks'\ \Lambda = .706$, *partial* $\eta^2 = .294$. Follow-up univariate two-way ANOVAs were performed. There were statistically significant differences in the adjusted mean for fixation duration ($F_{(1,105)} = 27.898$, $p < .001$, *partial* $\eta^2 = .210$), number of fixated images ($F_{1,105} = 36.640$, $p < .001$, *partial* $\eta^2 = .259$), and Euclidean scanpath distance ($F_{1,105} = 23.785$, $p < .001$, *partial* $\eta^2 = .185$).

Regarding fixation duration, FIs fixated on the images for 16.112 seconds (95% CI, 10.603 to 22.163) more than FDs. Regarding the number of fixated images, FIs fixated on 20.810 more images (95% CI, 13.991 to 27.622) than FDs. Regarding the Euclidean distance of the scanpath, FIs produced 3.692 longer scanpaths (95% CI, 2.296 to 5.088) than FDs. No effects of the interaction device were found. The results are depicted in Figure 3.

4.4. Analysis of password creation strategies

To provide supportive insights for the conclusions drawn from the quantitative analysis and shed light in the users' password creation strategies, we performed a qualitative analysis on the informal discussions' data and supported the results with the qualitative data of their visual behavior during the task. After completing the task, participants were asked about the strategy they followed to create their graphical passwords. FDs reported creating their authentication keys in a random way based on images that caught their attention or images related to things they like, mainly located at the top of the image grid. Some reported that they memorized the name of the objects in the images they selected and others that they created a pattern (e.g., Z-shape). They considered they had too many images on the grid and they got anxious with the number of the images. They focused on the goal, which was to select five images, rather than think of how they would create a strong and easy-to-remember password. When they were asked about the categories of the images they have seen, they reported of having seen shoes, electronics, food, fruits, and flowers. Very few of them reported sports, vegetables, and other categories located at the bottom part of the image grid. When asked to log in after completing the GEFT test, they had difficulties in remembering the selected password. Most reported that they were expecting to view the same image grid during login and they could remember where the images were located on the registration image grid and had difficulties with identifying them on the login image grid. In addition, FD users answered negatively in the question of whether they would remember their graphical key after a month. Some interesting comments received by FD users were the following:

> *"I thought the shoes would be easy to remember. But, there were only four shoes provided so I selected the egg only because I thought I would remember it. When logging in, I could not remember the order of the images because the shoes were not provided in the same order."* $\sim P_{02}$
>
> *"I selected the first five images because I thought the images would be provided in the same order during login so it would be easy to remember the password."* $\sim P_{32}$
>
> *"The [selected] images relate to things currently related to my life; for example I'm planning to buy a camera, that's why I selected it. I'm an engineer that's why I added the calculator. I play football and basketball and I ate a banana in the morning."* $\sim P_{48}$
>
> *"I selected the flower images because they appealed to me, I didn't think of a specific order, I just started adding them from left to right until I got five of them. If the order of the images changed, I would not remember the correct order."* $\sim P_{93}$

Observing the visual behavior of the FD individuals allowed us to draw conclusions about their strategy during password composition. Their attention was drawn by images located at the top of the image grid. They selected the first image after scanning a very small part of the image grid. They mainly focused on the first two rows of the image grid, as there was a lot of visual information they had to explore and none of the images stood out. In their eyes, the whole grid was seeking their attention and given their inherent difficulties in exploring visually complex scenes they limited their search activity at the top few rows of the grid. In Figure 4:left a typical heatmap of an FD individual is depicted.

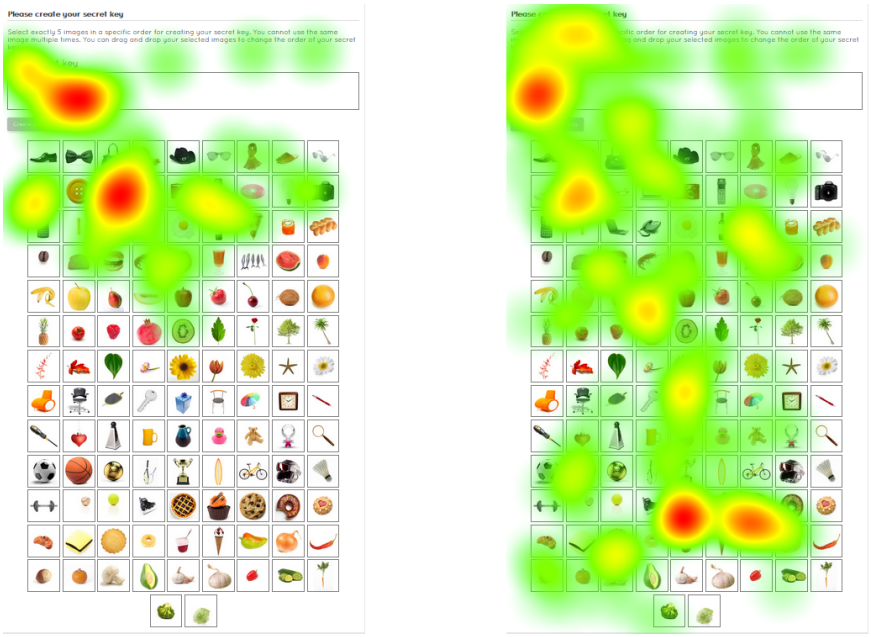

Figure 4. Heatmaps for FDs (left) and FIs (right)

On the other hand, FI users took more time to read the instructions carefully, browse the images and select their authentication keys. Their selection strategy was mainly based on their daily routine and included images related to their hobbies. Others created stories starting with an image that caught their attention and associated the order of the selected images with the stories they made up. Some reported of selecting objects with common visual features (e.g., color) or functional features (e.g., shoes) and ordering them alphabetically or visually (e.g., biggest object first). At first, they scanned the image grid carefully to see the available options and then they moved to selecting the images of their graphical password. They carefully thought of the order that they would select them, and they often used the drag-and-drop option to change the order of the images on the password before submitting it. When they were asked about the categories of the images they have seen, they reported almost all the categories of the images, and they could even remember their location on the image grid (i.e. top, middle, bottom). They were all positive that they would remember the created keys after a month, and they did not face any difficulty with remembering their key when they were asked to log in immediately after completing the GEFT test. Some interesting comments received by FI users were the following:

> *"I selected and ordered the images based on my everyday gym routine."* $\sim P_{29}$

> *"I selected the first image because I thought it was appealing. Then I found all images of the same color and ordered the objects alphabetically to create a password that I would remember later on."* $\sim P_{52}$

> *"I liked the cookie image, so I decided to use it. Then, I tried to find images with 'missing' parts of the cookie and created a story of adding one element each time to create the whole cookie."* $\sim P_{67}$

"I selected all the images that create an outfit. I would have started with clothes, but no such image was provided so I used the high heels, the purse, the scarf, the sunglasses and the button." $\sim P_{88}$

Observing the visual behavior of the FI individuals reveals their inherent ability to explore visually complex scenes (Aydın et al., 2013). They started with scanning the image grid almost sequentially row by row and spent half the password creation time in investigating the image grid. When they had an idea of the categories of the images, they decided on the approach they would use to create the password. Based on this approach, they searched for suitable images. In this second search round, they did not follow a row-by-row approach, but, because they already had an idea of the location of the images on the grid, they instead jumped from one point to another quickly. In Figure 4:right a typical heatmap of an FI individual is depicted.

To sum up, the password creation strategies of FD individuals were less organized and were based mainly on images that caught their attention and they often used the location of the images on the image grid to memorize their passwords. On the other hand, FI individuals were self-motivated when creating their passwords and either used objects with specific visual or functional features or created a story and used the images to fill their story.

5. Findings and Implications

5.1. Findings

The results of the study suggest that the human cognitive factors affect the graphical passwords' strength when using a recognition-based GUA scheme. The unique characteristics of the FD and FI individuals influenced the strategy they followed to create a graphical password, which is reflected in their visual behavior. The analysis of the strength of the graphical passwords of the FDs and the FIs revealed that FIs created more difficult to guess and more complex passwords than FDs (Figure 2). Observing the users' visual behavior allowed us to draw conclusions in regards to their graphical password creation strategy and associate this with the quantitative results of the study. A summary of the findings of the study is provided in Table 2.

Overall, FIs followed a more analytic image exploration approach, as they fixated on more images, produced more fixations and for longer time and had bigger Euclidean scanpath distances. Rather than selecting images that immediately caught their attention, they inherently created a goal of exploring their options before deciding on which images they would use to create their password (Figure 4: right). Then, their self-motivational characteristic led them to create the goal of making a memorable password, after having decided which approach would be a better fit based on the available images. Observing their visual behavior confirmed this strategy since they explored the image grid at first and then they decided on how they would create their password and searched for those images that would better meet their goal. Their analytic nature guided them to order their images based on a pattern (e.g., everyday habits) and, unlike the FDs, they consciously picked among similar images the one that best matched their goal. For example, they were aware that there are two ice cream images on the image grid and they selected the cone ice cream instead of the ice cream cup.

On the other hand, FDs were overwhelmed with the number of images available and the holistic approach they followed did not help them create strong passwords.

Table 2. Effects of FD-I cognitive style on password strength and on individuals' visual behavior, while creating their graphical password using the recognition-based GUA scheme.

Metric		Finding
Password strength	Guessability	• More guesses were required to crack FIs' passwords than FDs' passwords: $F_{1,105} = 4.144$, $p = .022$, *partial* $\eta^2 = .038$ (Figure 2) • Password guessability was strongly correlated with visual behavior (Number of fixated images: $r = .646$, $.012$; Mean fixation duration per image: $r = .385$, $p = .031$; Euclidean scanpath distance: $r = .301$, $p = .037$; Table 1)
	Complexity	• FIs created more complex passwords than FDs: $F_{1,105} = 4.852$, $p = .015$, *partial* $\eta^2 = .044$ (Figure 2). • Password complexity was strongly correlated with visual behavior (Number of fixated images: $r = .685$, $.008$; Mean fixation duration per image: $r = .545$, $p = .025$; Euclidean scanpath distance: $r = .305$, $p = .033$; Table 1)
Visual behavior	Fixation duration on images	• FIs fixated on the GUA images for longer periods than FDs, $F_{(1,105)} = 27.898$, $p < .001$, *partial* $\eta^2 = .210$ (Figure 3)
	Number of fixated images	• FIs fixated on more GUA images than FDs, $F_{1,105} = 36.640$, $p < .001$, *partial* $\eta^2 = .259$ (Figure 3)
	Scanpath length	• FIs produced longer fixations than FDs, $F_{1,105} = 23.785$, $p < .001$, *partial* $\eta^2 = .185$ (Figure 3)

They explored a very small number of images located mainly at the top half of the image grid (Figure 4: left) and thus they decreased the probability of selecting stronger passwords as they unwittingly made their selection from a smaller image pool. They selected images based on the position of the images on the image grid that related to their interests, but did not pay attention to their order despite that they were aware of the password policy of the GUA scheme. Observing their visual behavior during the graphical creation task revealed that they spent very few time to explore the image grid and rather immediately started selecting their images. They had not planned on their password creation strategy, but were rather guided by the images. They selected images that were associated with social interactions (e.g., images that represent team sports), which reflects their inherent tendency for communication. Given that they remember faces easier than objects (Pithers, 2002), the use of images of single-objects may have hindered the FDs from visually exploring the image grid.

The reported study suggests that the current *one-size-fits-all* approach is not the best solution for GUA as the GUA characteristics and the design space may be unintentionally either favoring or hindering people with specific cognitive characteristics. For example, the large number of images frustrated the FDs and thus they did not create strong passwords. Considering that cognitive styles rarely change through lifes-

pan (Wapner and Demick, 1992), studies like the reported one could drive the design of personalized GUAs that adapt to individual cognitive styles. Simple mechanisms, such as use of different types of images or use of different image grid layouts, could be elaborated to help users of different cognitive styles improve the strength of the passwords they create.

5.2. Implications

The main findings of the study underpin the value of considering cognitive styles (e.g., FD-I) as a human design factor, in both design and run-time and transforming the inter-dependencies between human and GUA design factors into formal representations modeling users' individual characteristics, and accordingly provide adaptive and personalized GUA schemes. In this context, there are two major challenges for applying recognition-based GUA schemes that cater for the cognitive styles of the users: the adaptation and personalization of the image grid based on the users' cognitive style and the run-time elicitation of the cognitive style.

Adaptation and personalization of image-grid layout in recognition-based graphical authentication based on FD-I differences: The exploration of a rather small part of the image grid by the FDs indicates the importance of displaying images in ways that are easily perceived by users and ensure equal distribution of attention to all images regardless of where they are located on the screen. Considering that in this study we used a single 9x14 image grid, and that visual behavior revealed that both FD and FI users fixated only on part of the images provided, alternative approaches of laying out the images should be considered, especially when providing a large number of images in order to achieve comparable theoretical entropy to that of the text-based mechanisms. The analysis of the results revealed that FD users selected images randomly, but the strength of the created passwords could be increased if the images were presented in a way that would guide the users to pay attention to all the provided images, for example by illustrating multiple, smaller grids of images and preventing password selection until the user has scanned a certain percentage of the images (Katsini et al., 2018c). This could be used in combination with password strength metrics that derive from the users' visual behavior during the password creation task, such as gaze-based entropy introduced in Katsini et al. (2018d). The observed differences on the visual behavior of the FDs and the FIs necessitate the use of adaptation and personalization techniques for guiding users to scan the entire image grid through adjusting the GUA interface, based on the individual cognitive strategies. Given that FI individuals scanned more images than FD individuals, alternative layouts could improve the visibility of the images located at the bottom part of the image grid. Given that FIs scanned around sixty images on average, a suitable approach could be providing two image grids, each consisting of sixty images. For the FD individuals, given that they fixated on around forty images on average they could be provided with three image grids, consisting of forty images each, to ensure better visibility of the total choice pool. In both cases, it should be investigated whether the users should be guided to view all image grids or if it should be left to them to decide on their approach. FI individuals, due to their analytic nature may be more likely to explore the whole image grid, while FD individuals may require more help from the mechanism for viewing all images of the grid. For example, they could be hindered from selecting a password unless they have visited at least two of the three image grids. To improve users' behavior by guiding them to scan more images and assist them to

create stronger and more complex passwords, several mechanisms can be used, such as gradually revealing the image grid (Thorpe et al., 2014), use of 3D presentation techniques (Katsini et al., 2018c), and use of saliency masks for those images that are most likely to attract visual attention to spots on the image grid that were given much attention based on eye-gaze data (Katsini et al., 2018b).

Run-time elicitation of FD-I characteristics based on visual behavior differences of users: Aiming to apply adaptation and personalization techniques in graphical password schemes based on human cognitive styles, as discussed before, the elicitation of users' FD-I characteristics is required. Nowadays, the FD-I style is identified through in-lab time-consuming techniques which require human intervention (e.g., facilitator, test reviewer). These are credible and validated explicit elicitation techniques, but they make real time integration of such human cognitive factors impractical in interactive system design, because the process cannot be conducted solely by electronic equipment and human intervention is necessary. To overcome this issue, an eye-tracking-based implicit method elicitation could be used, considering that there is a strong correlation between the cognitive style and the visual behavior of the users (Raptis et al., 2016b). Raptis et al. (2017) proposed a multi-factorial framework, based on eye-tracking, to implicitly elicit the users' FD-I style in real time. According to this framework, a set of areas of interest are defined within a visual activity, based on the activity characteristics (e.g., type, difficulty, sequence). Based on the areas of interest, suitable low-level (e.g., number of fixations, fixation duration) and/or high-level (e.g., scan-paths) eye-tracking metrics are used to elicit implicitly and in real time the cognitive style, through intelligent and automatic modelling processes. Accordingly, the framework proposed by Raptis et al. (2017) could be bootstrapped based on the eye-tracking differences of FD-I users as described above. Each of the images of the GUA mechanism used in our study is defined as an area of interest, which enables the framework to measure the differences between the FD and the FI individuals based on the metrics discussed earlier (i.e., fixation duration and fixation count) and elicit the cognitive style of the users in real time before selecting the first image during graphical password creation. Recent studies in implicit cognition-based user-modeling when using graphical user authentication schemes (Katsini et al., 2018a; Raptis et al., 2017) are moving towards this direction.

6. Study validity

This research work entails several limitations inherent to the multidimensional character and complexity of the factors investigated. Regarding internal validity, the study environment and the study procedure remained the same for all participants. The methodology and statistical tests used to answer the research objectives met all the required assumptions, despite the rather limited size of the sample. Moreover, the confounding of factors other than FD-I cognitive style (e.g., gender, age, device) was considered, and no effects were revealed. Focusing on the study instruments, we used GEFT test to classify an individual as either FD or FI based on a cut-off score. Considering that the GEFT test highlights cognitive differences along a continuum scale, the use of a cut-off score might not classify correctly individuals that fall in between the two end points (e.g., Field-mixed (Angeli et al., 2009)). However, it is important to stress that the distribution of users' scores on the GEFT test in our sample is comparably similar to general public GEFT test scores as shown in several studies

with individuals with different demographics (Khatib and Hosseinpur, 2011; Rittschof, 2010). Finally, the guessability of the passwords was measured using an exhaustive key search. Despite that there are other more advanced algorithms for guessing the graphical passwords based on user choices (Davis et al., 2004), the scope of this paper is to use the password guessability as a measure for comparing passwords created by FDs and FIs rather than compare cracking algorithms for graphical passwords.

With regards to the ecological validity of the study, the sessions were performed in times and days convenient for each participant. We used only one recognition-based GUA mechanism, although numerous others are available and worth studying too. However, the selected recognition-based GUA mechanism has been used in other studies (Belk et al., 2017a; Katsini et al., 2017) and is based on popular recognition-based GUA schemes, such as DejaVu (Dhamija and Perrig, 2000), PassFaces (Brostoff and Sasse, 2000), and ImagePass (Mihajlov and Jerman-Blažič, 2011). To increase the ecological validity, participants were asked to create a password that they would use for an important account, such as email account, since people tend to create more secure passwords when they are asked to secure such types of accounts (Florencio and Herley, 2007). We also note that the use of an eye-tracking technology was a limitation, as the individuals rarely use such equipment when interacting with authentication schemes. However, the fact that the eye-tracking technology used in the study were wearable glasses, made the participants feel more comfortable after a while, as they could interact with the system as they would normally do.

Regarding the study external validity (i.e., to what extent an effect can be generalized), our participants' age-span was limited, ranging from 18 to 47. However, the fact that the ages were normally distributed and that the GEFT scores are comparably similar to general public across populations with varying demographics (Angeli et al., 2009; Hong et al., 2012; Khatib and Hosseinpur, 2011) contribute to the study generalizability. The participation of older subjects is required and will be considered in our future research attempts, to increase the external validity of the study. It is also important to stress out that participants used popular devices (desktop and tablet computer) to create their graphical passwords. We expect that our results will be replicated for: a) other types of recognition-based GUA schemes, b) other types of interaction device (e.g., virtual reality head-mount displays), and c) applications that embrace visual decision-making tasks which can be found in varying domains, besides security, such as e-shopping, e-learning, gaming, and engineering. Finally, given the self-reporting and complex nature of the presented experimental design, similar research attempts are required in order to acquire a deeper understanding about the interplay between human cognitive differences, visual behavior, and graphical authentication, and thus, increase the validity of this research.

7. Ethical Considerations

Our research incorporated appropriate consideration of ethical issues into the design, conduction, and analysis of the research study. Our research involved interaction with human subjects, and it was performed under context-specific ethical guidelines. Researchers actively respected the human rights and dignities of all those involved in the user study and appropriately addressed questions of consent, capacity, power relations, deception, confidentiality, and privacy. All individuals who participated voluntarily in our research study agreed to and signed a consent form that their interactions with the graphical authentication scheme would be recorded and analyzed anonymously as

part of an experimental user study of the researchers' group. Participants could bring the experiment to an end at any time and for any reason. During the experiment, the researchers had the authority to bring it to an end, if there was probable cause to believe that it could harm the participants in any way. Moreover, all participants were informed about the experiment procedure and the rights they had as volunteers. However, no further details about the aim of the study were provided to them in order to avoid bias effects. Finally, the research team conducted the study and analyzed the data with rigor, respect, and responsibility.

8. Conclusion

This paper revealed that FD-I cognitive style of users influenced the strength of the created passwords, in terms of generated password guessability and complexity, when using a recognition-based GUA scheme. It also revealed that individual FD-I cognitive style was correlated with the users' visual behavior (i.e., the total fixation duration, the number of unique fixated images, and the Euclidian scanpath length) during the password creation task.

We interpreted the differences on the visual behavior between the FIs and the FDs from a cognitive science theory perspective, which was backed up also by a qualitative analysis of the strategies of the users when creating their graphical passwords. Hence, we argue that recognition-based graphical passwords can be improved through adaptation and personalization techniques to assist users with different individual cognitive styles during graphical password composition tasks. Such implications were also discussed in the paper.

Bearing in mind that recent research attempts (Bulling et al., 2012; Cantoni et al., 2015; Galdi et al., 2016; Khamis et al., 2017; Song et al., 2016) try to incorporate novel authentication schemes based on eye-tracking methods and users' gaze patterns, studies like the reported one provide valuable insights for further understanding the complexity and interplay among users' high-level cognitive factors and eye gaze behavior within graphical authentication tasks.

Acknowledgments

We would like to thank all the participants who voluntarily took part in our study. This research was supported by the General Secretariat for Research and Technology (GSRT) and the Hellenic Foundation for Research and Innovation (HFRI) - 1st Proclamation of Scholarships for PhD Candidates / Code: 617

References

Angeli, C., Valanides, N., and 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.

Aydın, Ü. A., Acartürk, C., and Çağıltay, K. (2013). The role of visual coherence in graphical passwords. *Np, nd Web*, 10.

Belk, M., Fidas, C., Germanakos, P., and Samaras, G. (2013a). Security for diversity: Studying the effects of verbal and imagery processes on user authentication mechanisms. In *Lecture*

Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), volume 8119 LNCS, pages 442–459. Springer.

Belk, M., Fidas, C., Germanakos, P., and Samaras, G. (2017a). The interplay between humans, technology and user authentication. *Comput. Hum. Behav.*, 76(C):184–200.

Belk, M., Fidas, C., Katsini, C., Avouris, Nikolaos Samaras, G., Avouris, N., and Samaras, G. (2017b). Effects of Human Cognitive Differences on Interaction and Visual Behavior in Graphical User Authentication. In *IFIP TC13 Human-Computer Interaction (INTERACT 2017)*, volume 10515 LNCS.

Belk, M., Germanakos, P., Fidas, C., and Samaras, G. (2013b). Studying the effect of human cognition on user authentication tasks. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 7899 LNCS, pages 102–113. Springer.

Brostoff, S. and Sasse, M. A. (2000). Are passfaces more usable than passwords? a field trial investigation. In McDonald, S., Waern, Y., and Cockton, G., editors, *People and Computers XIV — Usability or Else!*, pages 405–424, London. Springer London.

Bulling, A., Alt, F., and Schmidt, A. (2012). Increasing the security of gaze-based cued-recall graphical passwords using saliency masks. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*, page 3011, New York, New York, USA. ACM Press.

Cantoni, V., Galdi, C., Nappi, M., Porta, M., and Riccio, D. (2015). GANT: Gaze analysis technique for human identification. *Pattern Recognition*, 48(4):1027–1038.

Chiasson, S., Forget, A., Stobert, E., van Oorschot, P. C., and Biddle, R. (2009). Multiple password interference in text passwords and click-based graphical passwords. In *Proceedings of the 16th ACM conference on Computer and communications security - CCS '09*, page 500, New York, New York, USA. ACM Press.

Chowdhury, S., Poet, R., and Mackenzie, L. (2013). A comprehensive study of the usability of multiple graphical passwords. In Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., and Winckler, M., editors, *Human-Computer Interaction – INTERACT 2013*, pages 424–441, Berlin, Heidelberg. Springer Berlin Heidelberg.

Davis, D., Monrose, F., and Reiter, M. K. (2004). On User Choice in Graphical Password Schemes. In *In 13th USENIX Security Symposium*.

De Luca, A., Denzel, M., and Hussmann, H. (2009). Look into my eyes! In *Proceedings of the 5th Symposium on Usable Privacy and Security - SOUPS '09*, page 1, New York, New York, USA. ACM Press.

Dhamija, R. and Perrig, A. (2000). Deja vu-a user study: Using images for authentication. In *USENIX Security Symposium*, volume 9, page 4.

Djamasbi, S., Siegel, M., and Tullis, T. (2011). Visual hierarchy and viewing behavior: An eye tracking study. In *International Conference on Human-Computer Interaction*, pages 331–340. Springer.

Everitt, K. M., Bragin, T., Fogarty, J., and Kohno, T. (2009). A comprehensive study of frequency, interference, and training of multiple graphical passwords. In *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*, page 889, New York, New York, USA. ACM Press.

Feusner, M. and Lukoff, B. (2008). Testing for statistically significant differences between groups of scan patterns. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*, ETRA '08, pages 43–46, New York, NY, USA. ACM.

Florencio, D. and Herley, C. (2007). A large-scale study of web password habits. In *Proceedings of the 16th international conference on World Wide Web*, pages 657–666. ACM.

Galdi, C., Nappi, M., Riccio, D., and Wechsler, H. (2016). Eye movement analysis for human authentication: a critical survey. *Pattern Recognition Letters*, 84:272–283.

Hong, J.-C., Hwang, M.-Y., Tam, K.-P., Lai, Y.-H., and Liu, L.-C. (2012). Effects of cognitive style on digital jigsaw puzzle performance: A gridware analysis. *Computers in Human Behavior*, 28(3):920–928.

Huestegge, L. and Pimenidis, L. (2014). Visual Search in Authentication Systems Based on

Memorized Faces: Effects of Memory Load and Retention Interval. *International Journal of Human-Computer Interaction*, 30(7):604–611.

Jermyn, I., Mayer, A., Monrose, F., Reiter, M. K., and Rubin, A. D. (1999). The Design and Analysis of Graphical Passwords.

Katsini, C., Belk, M., Fidas, C., Avouris, N., and Samaras, G. (2016). Security and usability in knowledge-based user authentication: A review. In *Proceedings of the 20th Pan-Hellenic Conference on Informatics*, PCI '16, pages 63:1–63:6, New York, NY, USA. ACM.

Katsini, C., Fidas, C., Belk, M., Avouris, N., and Samaras, G. (2017). Influences of users' cognitive strategies on graphical password composition. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '17, pages 2698–2705, New York, NY, USA. ACM.

Katsini, C., Fidas, C., Raptis, G. E., Belk, M., Samaras, G., and Avouris, N. (2018a). Eye gaze-driven prediction of cognitive differences during graphical password composition. In *23rd International Conference on Intelligent User Interfaces*, IUI '18, pages 147–152, New York, NY, USA. ACM.

Katsini, C., Fidas, C., Raptis, G. E., Belk, M., Samaras, G., and Avouris, N. (2018b). Influences of human cognition and visual behavior on password strength during picture password composition. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 87:1–87:14, New York, NY, USA. ACM.

Katsini, C., Raptis, G. E., Fidas, C., and Avouris, N. (2018c). Does image grid visualization affect password strength and creation time in graphical authentication? In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*, AVI '18, pages 33:1–33:5, New York, NY, USA. ACM.

Katsini, C., Raptis, G. E., Fidas, C., and Avouris, N. (2018d). Towards gaze-based quantification of the security of graphical authentication schemes. In *Proceedings of the Tenth Biennial ACM Symposium on Eye Tracking Research & Applications*, ETRA '18, New York, NY, USA. ACM.

Khamis, M., Hasholzner, R., Bulling, A., and Alt, F. (2017). GTmoPass: two-factor authentication on public displays using gaze-touch passwords and personal mobile devices. In *Proceedings of the 6th ACM International Symposium on Pervasive Displays - PerDis '17*, pages 1–9, New York, New York, USA. ACM Press.

Khatib, M. and Hosseinpur, R. M. (2011). On the Validity of the Group Embedded Figure Test (GEFT). *Journal of Language Teaching and Research*, 2(3).

Komogortsev, O. V., Gobert, D. V., Jayarathna, S., Do Hyong Koh, and Gowda, S. M. (2010). Standardization of Automated Analyses of Oculomotor Fixation and Saccadic Behaviors. *IEEE Transactions on Biomedical Engineering*, 57(11):2635–2645.

Melicher, W., Mazurek, M. L., Kurilova, D., Segreti, S. M., Kalvani, P., Shay, R., Ur, B., Bauer, L., Christin, N., and Cranor, L. F. (2016). Usability and Security of Text Passwords on Mobile Devices. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, pages 527–539, New York. ACM Press.

Meng, W., Li, W., Jiang, L., and Meng, L. (2016). On Multiple Password Interference of Touch Screen Patterns and Text Passwords. In *Chi '16*, pages 4818–4822.

Mihajlov, M. and Jerman-Blažič, B. (2011). On designing usable and secure recognition-based graphical authentication mechanisms. *Interacting with Computers*, 23(6):582–593.

Mihajlov, M., Jerman-Blažič, B., and Ciunova Shuleska, A. (2016). Why That Picture? Discovering Password Properties in Recognition-Based Graphical Authentication. *International Journal of Human–Computer Interaction*, 32(12):975–988.

Nicholson, J., Coventry, L., and Briggs, P. (2013). Age-related performance issues for pin and face-based authentication systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 323–332. ACM.

O'Gorman, L. (2003). Comparing passwords, tokens, and biometrics for user authentication. *Proceedings of the IEEE*, 91(12):2021–2040.

Oltman, P. K., Raskin, E., and Witkin, H. A. (1971). *Group Embedded Figures Test*. Consulting Psychologists Press, Palo Alto CA, USA.

Pithers, R. (2002). Cognitive learning style: a review of the field dependent-field independent approach. *Journal of Vocational Education & Training*, 54(1):117–32.

Raptis, G. E., Fidas, C. A., and Avouris, N. M. (2016a). Do Field Dependence-Independence Differences of Game Players Affect Performance and Behaviour in Cultural Heritage Games? *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play - CHI PLAY '16*, pages 38–43.

Raptis, G. E., Fidas, C. A., and Avouris, N. M. (2016b). Using eye tracking to identify cognitive differences: A brief literature review. In *Proceedings of the 20th Pan-Hellenic Conference on Informatics*, PCI '16, pages 21:1–21:6, New York, NY, USA. ACM.

Raptis, G. E., Katsini, C., Belk, M., Fidas, C., Samaras, G., and Avouris, N. (2017). Using Eye Gaze Data and Visual Activities to Infer Human Cognitive Styles. In *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization - UMAP '17*, pages 164–173.

Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. *Memory & Cognition*, 5(4):443–448.

Rittschof, K. A. (2010). Field dependence–independence as visuospatial and executive functioning in working memory: implications for instructional systems design and research. *Educational Technology Research and Development*, 58(1):99–114.

Salehi-Abari, A., Thorpe, J., and Van Oorschot, P. C. (2008). On purely automated attacks and click-based graphical passwords. *Proceedings - Annual Computer Security Applications Conference, ACSAC*, pages 111–120.

Shrestha, S. and Lenz, K. (2007). Eye gaze patterns while searching vs. browsing a website. *Usability News*, 9(1):1–9.

Song, C., Wang, A., Ren, K., and Xu, W. (2016). EyeVeri: A secure and usable approach for smartphone user authentication. In *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*, pages 1–9. IEEE.

Steichen, B., Wu, M. M. a., Toker, D., Conati, C., and Carenini, G. (2014). Te,Te,Hi,Hi: Eye gaze sequence analysis for informing user-adaptive information visualizations. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8538:183–194.

Stobert, E. and Biddle, R. (2013). Memory retrieval and graphical passwords. In *Proceedings of the Ninth Symposium on Usable Privacy and Security - SOUPS '13*, page 1, New York, New York, USA. ACM Press.

Sun, C., Wang, Y., and Zheng, J. (2014). Dissecting pattern unlock: The effect of pattern strength meter on pattern selection. *Journal of Information Security and Applications*, 19(4-5):308–320.

Templeton, G. F. (2011). A two-step approach for transforming continuous variables to normal: Implications and recommendations for is research. *CAIS*, 28(1):41–58.

Thorpe, J., Al-Badawi, M., MacRae, B., and Salehi-Abari, A. (2014). The presentation effect on graphical passwords. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*, pages 2947–2950, New York, New York, USA. ACM Press.

Thorpe, J. and Van Oorschot, P. C. (2007). Human-Seeded Attacks and Exploiting Hot-Spots in Graphical Passwords * 1. *16th USENIX Security Symposium*, page 8.

Tobii AB (2017). Tobii Pro Glasses 2: User Manual.

Wapner, S. and Demick, J. (1992). Field Dependence-Independence in Adult Development and Aging. *Field Dependence-independence Bio–psycho-social Factors Across the Life Span*, pages 245–268.

Weinshall, D., Kirkpatrick, S., Tao, H., Adams, C., Sampangi, R. V., Hawkey, K., Dhamija, R., Perrig, A., De Angeli, A., Coutts, M., Coventry, L., Johnson, G. I., Cameron, D., Fischer, M. H., Davis, D., Monrose, F., and Reiter, M. K. (2004). Pass-Go: A Proposal to Improve the Usability of Graphical Passwords. *IJ Network Security*, 7(2):273–292.

Wiedenbeck, S., Waters, J., and Birget, J. (2005). Authentication using graphical passwords: effects of tolerance and image choice. *Proceedings of the ...*, pages 1–12.

Witkin, H. A., Moore, C. A., Goodenough, D. R., and Cox, P. W. (1975). Field-Dependent and Field-Independent Cognitive Styles and Their Educational Implications. *ETS Research Bulletin Series*, 1975(2):1–64.

Zhao, Z., Ahn, G.-J., and Hu, H. (2015). Picture Gesture Authentication: Empirical Analysis, Automated Attacks, and Scheme Evaluation. *ACM Transactions on Information and System Security*, 17(4):1–37.

About the Authors

A short biography for each author:

Christina Katsini is a Ph.D. Candidate at the University of Patras, Greece. Her interests lie in understanding how people interact with systems and services, and in designing for the people with the people. In her Ph.D. research, she is investigating user choices in graphical user authentication from a human cognitive perspective.

Christos Fidas (Ph.D., http://cfidas.info) is an electrical and computer engineer, and senior researcher with an interest in cultural heritage informatics, usable and secure information systems, and human socio-cultural and cognitive factors. He is an Assistant Professor at the Department of Cultural Heritage Management and New Technologies, University of Patras, Greece.

Marios Belk (Ph.D., http://mbelk.info) is a computer scientist and usable security researcher with an interest in user authentication, human interaction proofs, user modeling, personalization, and cognitive sciences; he is a Lecturer in Cybersecurity and Data Analytics at the University of Central Lancashire, Cyprus Campus.

† **George Samaras** (Ph.D., http://www.cs.ucy.ac.cy/ cssamara) was a computer scientist and senior researcher with an interest in human-centered computing, cognitive and visual computing, adaptive and personalization systems, and data management; he was a full Professor at the Department of Computer Science, University of Cyprus, Cyprus. Prof. Samaras deceased in July, 2018.

Nikolaos Avouris (MSc, Ph.D., https://sites.google.com/view/avouris) is an electrical and computer engineer with a research interest in human-computer interaction. He is a Professor of Software Technology and Human-Computer Interaction in Electrical and Computer Engineering Department of University of Patras, Greece. He is Head of Interactive Technologies Lab and HCI Group.