The Interplay between Humans, Technology and User Authentication: A Cognitive Processing Perspective

Abstract: This paper investigates the interplay among human cognitive processing differences (field dependence *vs.* field independence), alternative interaction device types (desktop *vs.* touch) and user authentication schemes (textual *vs.* graphical) towards task completion efficiency and effectiveness. A four-month user study (N=164) was performed under the light of the field dependence-independence theory which underpins human cognitive differences in visual perceptiveness as well as differences in handling contextual information in a holistic or analytic manner. Quantitative and qualitative analysis of results revealed that field independent (FI) users outperformed field dependent users (FD) in graphical authentication, FIs authenticated similarly well on desktop computers as on touch devices, while touch devices negatively affected textual password entry performance of FDs. Users' feedback from a post-study survey further showed that FD users had memorability issues with graphical authentication and perceived the added difficulty when interacting with textual passwords on touch devices, in contrast to FI users that did not have significant usability and memorability issues on both authentication and interaction device types. Findings highlight the necessity to improve current approaches of knowledge-based user authentication research by incorporating human cognitive factors in both design and run-time. Such an approach is also proposed in this paper.

Keywords: Knowledge-based User Authentication; Human Cognitive Differences; Field Dependence-Independence; Task Performance; User Study.

1. INTRODUCTION

User authentication is a cornerstone of security in today's interactive systems [Koved and Stobert, 2016]. Derived from the Greek work *αὐθεντικός*; meaning real or genuine, user authentication is the act of confirming that a person interacting with a service is who he or she claims to be. Numerous user authentication schemes are currently deployed which can be classified into knowledge-based (*what the user knows*, *e.g.*, secret passwords, pictorial keys, sketches) [Biddle et al., 2012], token-based (*what the user has*, *e.g.*, credit cards) [Mare et al., 2016], and biometric-based (*what the user is*, *e.g.*, fingerprint, interaction behavior) [Renaud, 2005]. *Knowledge-based authentication schemes* are widely used today since: *a)* they are easy, fast and inexpensive to implement [Biddle et al., 2012]; and *b)* they don't entail the security and privacy flaws found in tokens (*e.g.*, loss or theft of credit card [Wang and Katabi, 2013]) and in biometrics (*e.g.*, users' fingerprints can be extracted from the objects they touch [Cao and Jain, 2016]). Research also indicates that knowledge-based approaches will continue to prevail in the next decays [Herley and van Oorschot, 2012], even in combination with other approaches (*e.g.*, token, biometric). Two important quality dimensions of an effective knowledge-based authentication scheme are related to its *security* and *usability* aspects. The security level determines its strength against adversary attacks, whereas usability levels are commonly determined by *memorability of selected secrets* and *task completion efficiency and effectiveness* [Biddle et al., 2012].

A plethora of knowledge-based authentication schemes have been proposed leveraging on different user experience and security factors, with the most prominent ones focusing on *text-based solutions* (passwords, PINs, etc.) [Herley and van Oorschot, 2012] and *graphical solutions* (pictorial, sketches, etc.) [Biddle et al., 2012; Nelson and Vu, 2010]. Text-based solutions require from users to memorize a secret that is represented by a sequence of textual characters, whereas graphical solutions are based on secret keys that typically consist of a sequence of images chosen by end-users from a pool of alternatives. Research on knowledge-based user authentication has become a complex endeavor since it embraces several parameters (human,

technology and design specific) that need to be taken into account. From the *human perspective*, recent research revealed that individual characteristics affect user authentication tasks, such as, users' age, gender and culture [Nicholson et al., 2013a; 2013b], cognitive disabilities [Ma et al., 2013] and cognitive processing abilities [Belk et al., 2015a; Belk et al., 2017a]. From the *technology perspective*, studies indicate that the device type, such as, desktop computers, tablets, smart phones, etc. have a main impact on users' performance and behavior in user authentication tasks [Melicher et al., 2016; von Zezschwitz et al., 2014; Schlöglhofer et al., 2012]. From the *design perspective*, research has shown that design characteristics, such as, authentication type (textual *vs.* graphical) [Brostoff and Sasse, 2000], pool of characters used in password policies [Komanduri et al., 2011], distribution of user-chosen images [Thorpe et al., 2014], image type (*e.g.*, faces *vs.* single-object images) in recognition-based graphical authentication [Mihajlov and Jerman-Blazic, 2011], image saliency [Alshehri & Crawford, 2016], image grid size during graphical key creation [Belk et al., 2017b], and image distortion [Hayashi et al., 2011] affect task completion performance and security.

Given that knowledge-based user authentication is primarily a cognitive task; users perceive, process, mentally represent and recall textual and graphical information, it is interesting to investigate whether and how state-of-the-art socio-cognitive theories can be adopted as an analysis framework aiming to assist the design of more user-centric and usable authentication schemes. In the aforementioned context, this paper contributes in further understanding the relationship between human cognitive processing differences (human factor), alternative interaction device types (technology factor) and alternative user authentication schemes (design factor).

Next, we present the background theory and motivation of this research endeavor. Thereafter, we describe the context and method of an empirical study, followed by an analysis of results and discussion of the main findings. Finally, we outline the implications and limitations of the reported research, and conclude the paper.

2. BACKGROUND THEORY AND MOTIVATION

Cognitive psychologists have studied the functioning of the human mind in terms of cognitive processes, providing evidence that variations in *cognitive processing styles and abilities* exist among individuals [Peterson et al., 2009]. Among a high number of theories on human cognitive differences [Kozhevnikov, 2007; Riding and Cheema, 1991], *field dependence-independence* is considered a highly researched and widely applied theory [Chen and Liu, 2008; Kozhevnikov, 2007; Rezaei and Katz, 2004; Riding and Cheema, 1991]. Field dependence-independence refers to the way individuals seek, retrieve, process, comprehend, organize and recall information. It highlights human cognitive differences along a continuum scale that distinguishes individuals into field dependent and field independent. The primary differentiation between field dependent and field independent individuals lies in visual perceptiveness [Hong et al., 2012] which is the ability to interpret the surrounding environment by processing information that is contained in visible light.

Field dependent (FD) individuals primarily view and organize the information and the structure presented by their visual field as a whole, proceed from the whole to the parts and organize information in loosely clustered wholes. Individuals termed as FD are not attentive to detail, cannot abstract an element from its context and tend to handle problems in a holistic way. Accordingly, given that FD individuals view the perceptual field as a whole, they are not efficient and effective in situations

where they are required to extract relevant information from a complex whole [Reardon and Moore, 1988; Witkin et al., 1977].

Field independent (FI) individuals view and organize the information and the structure presented by their visual field as a collection of parts, stress one or two aspects at a time, proceed from the parts to the whole and organize information in clear-cut groupings. When confronted with problems, FI individuals are good at extracting things from the context and prefer to handle them in a more analytical way. Given that FI individuals tend to experience items as discrete from their backgrounds and independent to the perceptual field, they are more successful in disembedding and isolating important information from a complex whole [Davis, 1991; Messick, 1993; Witkin et al., 1977; Goodenough and Karp, 1961].

Given the complex nature of human cognitive differences, various researchers have explained the aforementioned differences in a number of ways and therefore a standard and global definition has not been given [Peterson et al., 2009]. In particular, one group of researchers have attempted to explain and define field dependence-independence differences from a cognitive styles' perspective [Morgan, 1997], others from a cognitive abilities' perspective [Miyake et al., 2001; Rittschof, 2010], while others from a cognitive control's perspective [Jonassen and Grabowski, 1993]. Researchers have also used different labels for describing field dependence-independence [Riding and Cheema, 1991; Peterson and Deary, 2006], among others, Holists-Serialists [Pask and Scott, 1972], Wholists-Analysts [Riding and Cheema, 1991], Diverging-Converging [Hudson, 1966]. These dimensions have been related to the field dependence-independence theory (*e.g.*, Holists-Serialists in Clewley et al. [2010]) and used interchangeably in various user studies [Ling and Salvendy, 2009; Chen and Liu, 2008; Kinley et al., 2014; Clewley et al., 2011].

A number of researchers suggest that field dependence-independence is correlated with elementary cognitive processes of the human mind, such as working memory [Miyake et al., 2001; Rittschof, 2010]. Specifically, research findings have correlated field dependence-independence with Baddeley's visuospatial sketchpad and central executive system [Baddeley, 1992; 2012], since viewing shapes is primarily a visuospatial function, and the cognitive ability of extracting embedded shapes out of a complex whole involves the use of central executive functions, such as monitoring [Angeli et al., 2009]. Furthermore, a considerable number of studies confirmed that field dependence-independence affects search and browsing behavior of users in interactive systems [Chan et al., 2014; Kinley et al., 2014; Belk et al., 2013; Chen and Liu, 2008; Clewley et al., 2011; 2010], problem solving performance in educational and hypermedia environments [Angeli et al., 2004; 2009; 2013], gaming performance and behavior in cultural heritage games [Raptis et al., 2016], and digital-based puzzle problems [Hong et al., 2012].

2.1 Research Motivation

This research endeavor is primarily motivated by the fact that FD and FI users have differences in visual perceptiveness [Hong et al., 2012], visual working memory [Miyake et al., 2001; Rittschof, 2010], visual search abilities [Angeli et al., 2009], in the way they are affected by contextual surroundings as well as in the way they organize and process information of their surrounding visual field [Davis, 1991; Messick, 1993; Goodenough and Karp, 1961; Reardon and Moore, 1988]. Accordingly, this work is based on the main hypothesis that FD and FI users might perform differently in various user authentication types (textual and graphical) and interaction device types (desktop computers and mobile touch), since the way the

task is performed in each case is different, both from a cognitive processing and task execution perspective as follows:

- textual information processing *vs.* graphical information processing;
- pure recall in textual authentication *vs.* recognition and recall in graphical authentication;
- typing text *vs.* selecting images;
- typing text on a standard keyboard *vs.* virtual keyboard, and;
- visual search task *vs.* non-visual search task.

Based on the aforementioned rational, this paper aims to investigate the interplay between *human cognitive factors (FD and FI)*, *technology factors (desktop computer and mobile touch-based device)*, and *user authentication design factors (textual and graphical)* towards task completion time (efficiency) and number of attempts (effectiveness) required for successful authentication (Figure 1).

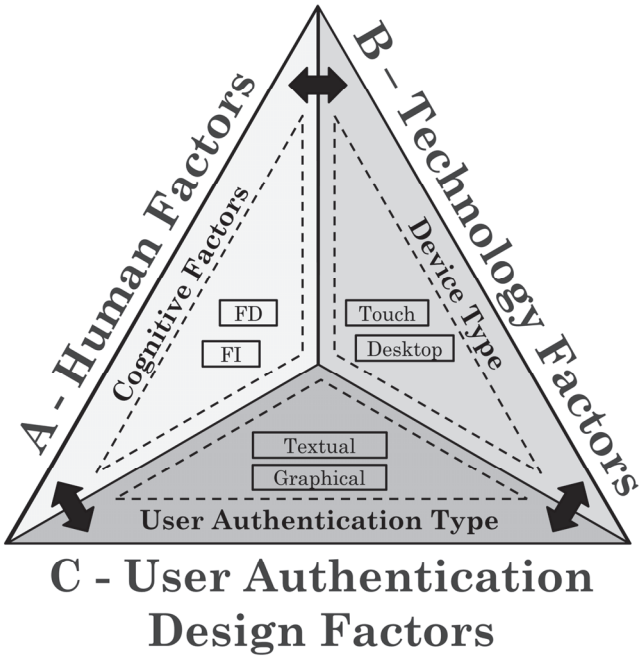


Fig. 1. Investigating the interplay between human cognitive factors (FD and FI), device type (desktop and touch), and user authentication type (textual and graphical).

3. METHOD OF STUDY

3.1 Research Instruments

The following research instruments were used: *i)* the GEFT paper-and-pencil test for classifying the participants into FD and FI groups; *ii)* two user authentication mechanisms (textual and graphical); and *iii)* two types of interaction devices (desktop computers and mobile touch-based devices). These are described next.

3.1.1. Cognitive Factor Elicitation. Users' field dependence-independence was measured through the Group Embedded Figures Test (GEFT) [Witkin et al., 1971] which is a widely used and accredited paper-and-pencil test [Kozhevnikov, 2007; Hong et al., 2012; Altun and Cakan, 2006]. The test measures the user's ability to find common geometric shapes in a larger design. The GEFT consists of a twenty-five item assessment, divided in 3 sections. The first section consists of 7 items which are only used for practice purposes and are not included in the assessment. Section 2 and

Section 3 consist of 9 items each. 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 the provided answers, the score of a participant might range between 0 and 18 in which the norm is 11.4 [Witkin et al., 1971]. For classifying participants into FD and FI groups we used a cut off score at 11 correct items which has been used in practice [Hong et al., 2012; Altun and Cakan, 2006]. Participants with a low score (below and equal to 11 correct items) were classified as FD, whereas participants with a high score (above 11 correct items) were classified as FI.

3.1.2. User Authentication Mechanisms and Policies. Two types of user authentication mechanisms were used: a *text-based password mechanism*, and a *recognition-based graphical authentication mechanism* (Figure 2). The text-based password mechanism required from users to recall and provide a fixed number of 8 characters which is a common length for personal computers [von Zezschwitz et al., 2013]. During user enrolment, the password should meet additionally the following requirements: include two numbers, one upper-case character and one special character. The policy did not allow the creation of a dictionary word by performing a dictionary check based on a method widely used in practice[1] [Komanduri et al., 2011]. The reasoning behind this policy choice was as follows: *a)* to generate a state-of-the-art real life security policy embracing lower- and upper-case letters, numbers, special characters, resulting in *"medium complexity internet passwords"* [von Zezschwitz et al., 2013; 2014]; *b)* prevent the participants to select an already known password from other contexts; and *c)* control the key creation length and type of characters used with the aim to isolate the possible influence that human cognitive characteristics have on user authentication task performance and prevent as much as possible the effect of other factors on the usability of the tasks.

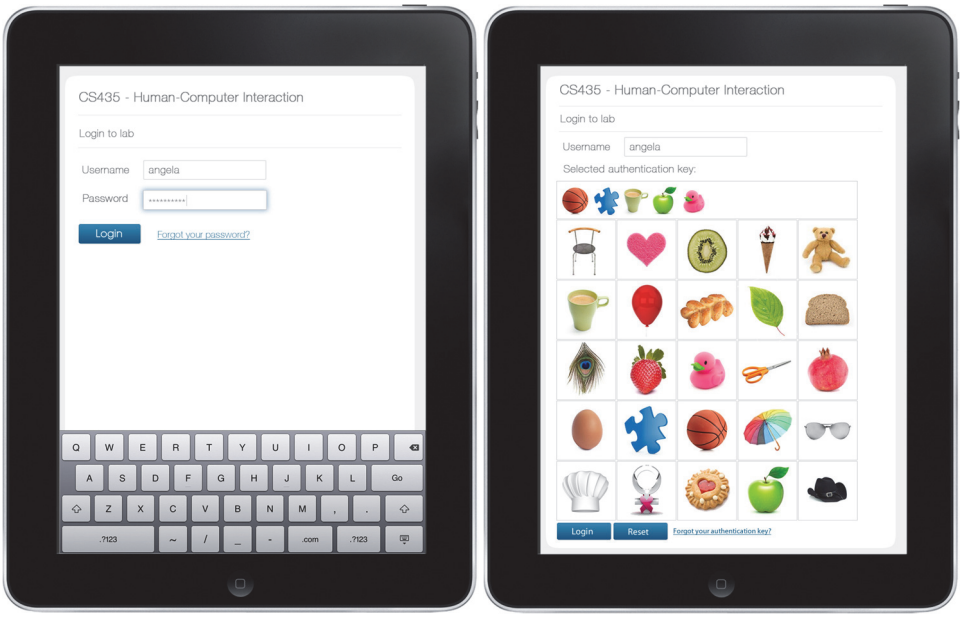


Fig. 2. Textual (left) and graphical (right) authentication mechanism deployed on a touch-based device.

[1] http://download.openwall.net/pub/wordlists

The graphical authentication mechanism was designed and developed following guidelines of accredited recognition-based graphical authentication mechanisms; such as DejaVu [Dhamija and Perrig, 2000], PassFaces [2009] and ImagePass [Mihajlov and Jerman-Blazic, 2011]. In particular, during user enrolment, users created their authentication key by selecting a fixed number of 5 images out of 120 images in a specific order. The same image could not be selected multiple times in a single authentication key. Single-object images were used since studies have shown that these are more memorable than faces and abstract images [Mihajlov and Jerman-Blazic, 2011; Chowdhury et al., 2013]. In addition, we intentionally did not use images that illustrate human faces since prior research has shown that user choice can be influenced by gender, race and attractiveness of the human face, and thus can be highly predictable and vulnerable to image-based dictionary attacks [Davis et al., 2004]. During user authentication, the 5 user-selected images were shuffled with 20 stable, system-assigned decoy images (Figure 2, right), and users were required to select these 5 images based on the predefined order. The provided image policy was based on existing approaches and is typical in recognition-based graphical authentication [Biddle et al., 2012; Renaud et al., 2013; Ma et al., 2013], in order to keep usability at reasonable levels.

3.1.3. Interaction Devices. Two types of interaction devices were used in the study: *standard desktop computers*, and *mobile touch-based devices*. The standard desktop computers had the following technical features: IBM Thinkcenter M73, 21-inch screen size monitor, Windows 7 operating system, IBM standard keyboard and mouse. For mobile touch-based devices we used Apple iPad 3 which has a 9.5-inch screen size. We intentionally used this particular touch-based device since prior research has shown that mobile screen sizes larger than 4.3 inches are more efficient during information seeking tasks [Raptis et al., 2013], as well as screen sizes between 5.7 and 9.7 inches have a main effect on enjoyment [Kim et al., 2011].

3.2 Procedure

The user study was applied in the frame of a university Computer Science laboratory course that was held two times per week during an academic semester, and students would authenticate through a login form for accessing their daily course's material (*i.e.*, daily lab exercise). All the students participated voluntarily and provided their consent that their interactions with the course's Web-site would be recorded anonymously in the context of an experimental user study of the researchers' group. In order to avoid bias, no details were provided about the aim of the research until the end of the study. Also, throughout the semester, users were able to opt out of the study any time they like. The user study lasted for four months and was split in three phases based on a mixed study design. The research design outline is depicted in Figure 3. Next, we describe in detail the procedure followed in each phase.

3.2.1. Phase A – User Classification. Seventeen controlled laboratory sessions with a maximum of ten participants per session were conducted during the first month of the study at times convenient to the participants. Each participant was first assigned a unique *userid* that was used by the system to track and store the data of each participant throughout the study. Then, participants were guided to an online consent form and each one read and agreed to participate. Thereafter, the GEFT paper-and-pencil test was administered aiming to highlight the participants' cognitive characteristics. Depending on their scores on the test, participants were classified as FD or FI. For the purpose of the study, before proceeding to Phase B, all

participants were classified into a specific group since this would be used to assign a balanced number of authentication types (textual or graphical) and device types (desktop or touch) to both user groups. The students' unique *userid* and GEFT classification was stored in the system's database that would be used in Phase B for assigning the particular user authentication and device type.

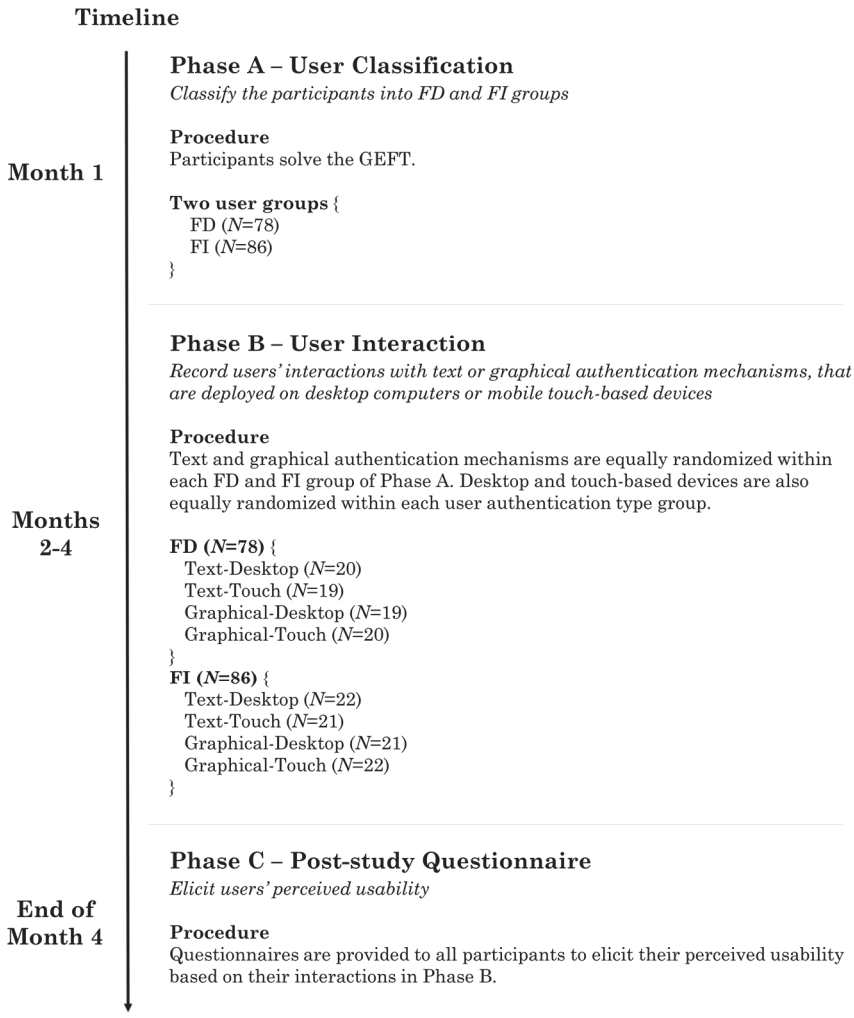


Fig. 3. Research design outline.

3.2.2. Phase B – User Interaction. Participants initially enrolled in the course's Web-site through a registration form. They were required to create a user profile by providing their unique *userid* assigned in Phase A, that was used to retrieve their field dependence-independence classification from the system's database. Participants then entered their age, gender and email (used as the unique username) as well as created a secret authentication key. Within each FD and FI group, half of the participants enrolled with a textual password mechanism and the other half with a graphical authentication mechanism. The user authentication type (text or graphical) was randomly assigned to the participants considering their field dependence-independence classification so that the assignment was balanced across user groups. Similarly, within each user authentication type group, the device type

(desktop or touch) was randomly assigned that was kept the same throughout the study. Then the participants interacted with the course's Web-site by utilizing the assigned device types (desktop computer or touch-based device) that were located at the course's laboratory room to access and read material of their weekly course. Aiming to control the frequency of access (twice per week) and prevent user interactions with other types of devices, the users' IP addresses were monitored so that they would access the authentication mechanisms only through the devices located at the laboratory room. The users' interactions with the authentication mechanisms were recorded for three months (two sessions per week; a maximum twenty-four sessions for each user).

3.2.3. Phase C – Post-study Questionnaire. Qualitative data were collected at the end of the study to elicit the users' perceptions, perceived usability and memorability based on their interactions with the user authentication mechanisms during Phase B.

3.3 Data Collection

The data captured during the user study were grouped in the following categories:

— *User Data* consists of data about the users' cognitive characteristics. Participants were classified as FD or FI based on their scores on the GEFT.

— *Task Data* consists of data about the task performed by the user. The total time in seconds required for successful authentication was recorded. Recording time started as soon the participants entered their username for user authentication type identification (textual or graphical) until they correctly provided the authentication key. Furthermore, the total number of attempts to successfully authenticate in each session was recorded. Based on the number of attempts, we classified a session as successful or failed, and accordingly, we measured the overall failure rate as follows: the total number of sessions that included a failed attempt, divided by the total number of all sessions [Brostoff and Sasse, 2000]. A session is considered as failed in case the participant needed more than one attempt to successfully authenticate.

— *Device Data* consists of data about the device type (standard desktop or mobile touch-based) in which user interaction takes place. A browser-based logging facility was utilized to detect the type of device used during interaction.

— *Authentication Key Data* represent the generated text-based and graphical authentication keys in plaintext which were gathered in order to perform a security analysis and accordingly contextualize the evaluation results. To avoid security and privacy issues related to the generated passwords of the users, the plaintext authentication keys were anonymously stored in a separate database (along with the assigned field dependence-independence classification and device type), than the actual system's database which encrypted the authentication keys accordingly. As such, the plaintext could not be correlated to any actual user.

3.4 Participants

A total of 164 individuals (72 Male, 92 Female) participated in the study. Participants varied from the age of 18 to the age of 25, with a mean age of 20.46 and were undergraduate students in Computer Science and Electrical Engineering. All participants had prior interaction experience with desktop-based and mobile touch-based devices and were familiar with text-based password mechanisms. No participant was familiar with the utilized recognition-based graphical authentication mechanism. This familiarity difference of our sample in using text-based passwords

in contrast to graphical authentication should be carefully considered when interpreting the results of the study.

Based on the users' scores on the GEFT, 78 participants (47.56%) were classified as FD and 86 participants (52.44%) were classified as FI. Figure 4 illustrates the frequency of users' scores on the GEFT. The mean score was 10.95 correct items (*sd*=3.91), which is comparable to the national mean of 11.4 in Witkin et al. [1971]. The lowest GEFT score was 3 correct items and the maximum 17 correct items.

These frequencies and distribution of scores are comparably similar to general public GEFT test scores as derived from the literature. For example, in Robinson et al. [2009], a total of 112 agricultural education graduates were tested utilizing the GEFT that generated a mean score of 12.88 (*sd*=3.89). The most frequent score was 15 (*n*=19), followed by scores of 14 and 18 (*n*=11). Overall, 32 individuals (31%) were classified as FD, and 71 individuals (69%) were classified as FI. In another study of Davis [2006], a total of 67 adults (mean age 45.1 years) were tested with the GEFT resulting in 38 FD users (56.7%), and 29 FI users (43.3%). The overall GEFT score mean of the population was 10.40 (*sd*=5.29). In Khatib and Hosseinpur [2011], the GEFT was administered to 60 adult professionals and students (age 20-35) which resulted in 23 FD users (38.3%), and 37 FI users (61.7%).

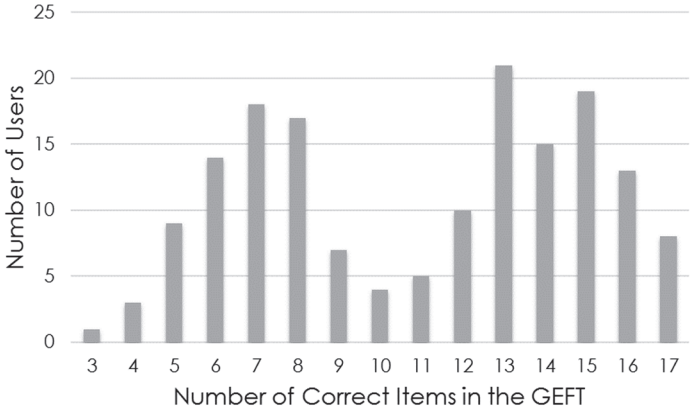


Fig. 4. Frequency of users' scores in the GEFT.

Throughout the three months, in Phase B of the study, 3674 authentication sessions have been recorded. A mean of 22.4 sessions were recorded per user (*sd*=1.65; *min*: 15; *max*: 24) along the three months (two sessions per week). Table 1 summarizes the number of authentication sessions per group for all combinations of field dependence-independence groups, user authentication type and device type.

Table 1. Number of authentication sessions per group combination.

	Textual		**Graphical**	
	Desktop	**Touch**	**Desktop**	**Touch**
FD	426 (11.59%)	441 (12%)	428 (11.65%)	460 (12.52%)
FI	478 (13.01%)	475 (12.93%)	473 (12.87%)	493 (13.42%)

4. ANALYSIS OF USER INTERACTIONS

4.1 Assumptions for Statistical Analysis and Details

The quantitative data were analyzed using R [R Core Team, 2015] with the *lme4* package [Bates et al., 2015], using a mixed effects analysis for task efficiency (time to login) and mixed logistic regression for task effectiveness (successful *vs.* unsuccessful attempts). These models were chosen since they enabled us to handle all the variables of the study while accounting for repeated-measures of individuals (24 sessions across a period of three months). Another advantage of such statistical models is that they can handle missing data of users, *e.g.*, a user that has not participated in some sessions across the three months of the study can be used in the analysis without requiring removing the user from the sample, as opposed to an analysis of a repeated-measures ANOVA [Pinheiro and Bates, 2000]. Furthermore, in order to account for multiple testing, we adjusted the alpha level with the Dunn-Sidak correction and accordingly report the corrected p-values in the analysis. Using the *dunn.test* package in R [Dinno, 2015], we defined the following function and accordingly corrected the original p-values:

function(P,N){1 - ((1 - P) ^ N)} where *P* the given p-value, and *N*=8 comparisons

4.1.1. Task Efficiency. For task efficiency (time to login), we performed a mixed effects analysis of the relationship between the time to successfully authenticate, and users' field dependence-independence, user authentication type and interaction device type. As fixed effects, we entered field dependence-independence (FD and FI), user authentication type (textual and graphical) and interaction device type (desktop and touch) into the model. As random effects, we used subjects in order to account for non-independence of measures. Visual inspection of residual plots revealed that linearity and homoscedasticity were violated. In this respect, we performed a log transformation on the dependent variable (time to login), and further inspected residual plots, histograms and Q-Q plots of the residuals, indicating that there were no obvious deviations from linearity, homoscedasticity and normality. P-values were obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question [Winter and Grawunder, 2012]. In the analyses that follow, we report descriptive statistics and comparisons between factors based on the non-transformed data, whereas significance testing is performed on the transformed data.

4.1.2. Task Effectiveness. A mixed logistic regression was conducted, with task effectiveness (successful *vs.* unsuccessful attempts) being the dependent variable. In all models, field dependence-independence (FD and FI), user authentication type (textual and graphical) and interaction device type (desktop and touch) where entered as fixed effects, and the subjects as random effects. We first constructed models that included the effects of field dependence-independence, user authentication type and interaction device type and further constructed a model that included interactions between field dependence-independence and user authentication type. For significance testing we tested the full model against a model without the effects in question by obtaining their likelihood ratio tests.

4.2 Task Completion Time Comparisons

Figure 5 illustrates the means of task completion time (in seconds) for all combinations of field dependence-independence (FDI), user authentication type (UA) and device type (Device).

4.2.1. Main Effects of Single Factors on Task Completion Time. The analysis revealed that users' field dependence-independence affected the time needed to authenticate ($\chi^2(1)=20.599$, $p<0.001$), with FI users being faster by 2.6 seconds ± 0.5 (standard errors (SE)). Furthermore, the authentication type (textual *vs.* graphical) has a main effect on authentication task completion time ($\chi^2(1)=21.79$, $p<0.001$). Interactions with textual password mechanisms were more efficient than graphical authentication mechanisms by 2.45 seconds ± 0.5 SE. It is important to mention that text-based password interactions on touch-based devices was rather long, compared to recent typical estimates [von Zezschwitz et al., 2014]. A possible explanation might be based on the fact that an unusual text-based policy was provided to the users (requiring fixed character length and fixed number of each character set), preventing them to freely create text-based passwords from other existing accounts. Finally, there was an effect of the device type towards time needed to authenticate ($\chi^2(1)=8.15$, $p=0.03$). Desktop-based user interactions were significantly faster than touch-based interactions by 1.12 seconds ± 0.5 SE.

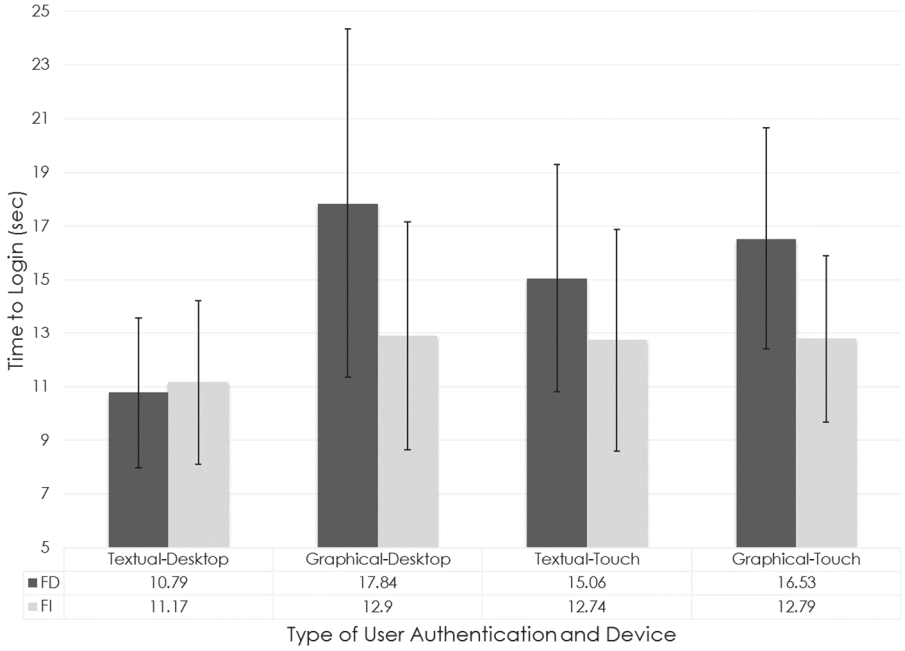


Fig. 5. Means of login time (sec) per field dependence-independence, user authentication and device group.

4.2.2. Main Interaction Effects of Field Dependence-Independence and User Authentication Type on Task Completion Time. There was a statistically significant interaction effect between users' field dependence-independence and user authentication type ($\chi^2(1)=8.46$, $p=0.02$). Pairwise comparisons between FD and FI users revealed that in the case of graphical authentication mechanisms, FI users significantly outperformed FD users in task completion time (FI-FD: $MD=-4.184$, $SE=0.68$; $\chi^2(1)=29.88$, $p<0.001$). In the case of textual password mechanisms, no significant differences in task completion time were recorded (FI-FD: $MD=-0.987$, $SE=0.642$; $\chi^2(1)=2.29$, $p=0.13$). A possible interpretation of the latter result might be based on the fact that all users were familiar with textual password mechanisms since the sample included experienced users and all of them had already registered in at least one online password protected account in the past. Furthermore, pairwise comparisons between user authentication types for each user group revealed the following: for FD users,

significant differences in task completion time were observed between textual and graphical authentication mechanisms (FD: Textual-Graphical: MD=-4.17, SE=0.80; $\chi^2(1)=22.279$, $p<0.001$). FD users were significantly faster in authenticating on textual passwords compared to graphical authentication mechanisms. In the case of FI users, significant differences were observed between the two authentication types (FI: Textual-Graphical: MD=-0.97, SE=0.512; $\chi^2(1)=4.47$, $p=0.03$). Nonetheless, descriptive statistics show that the mean difference is smaller than that of FD users (4.17 sec *vs.* 0.97 sec) since FI users were more efficient in authenticating through graphical authentication mechanisms compared to FD users.

4.2.3. Main Interaction Effects of Device Type and User Authentication Type on Task Completion Time. The analysis revealed a statistically significant interaction between device and authentication type on the time to authenticate ($\chi^2(1)=9.526$, $p=0.01$). Pairwise comparisons of user authentication types revealed that text-based and graphical authentication interactions on standard desktop computers had significant differences in task completion time (Desktop: Textual-Graphical: MD=-4.18, SE=0.67; $\chi^2(1)=31.404$, $p<0.001$), whereas in the case of touch-based interactions, no significant differences where observed between the two user authentication types (Touch-based: Textual-Graphical: MD=-0.76, SE=0.73; $\chi^2(1)=1.58$, $p=0.20$). The latter result has been caused by the significant increase of time to complete the text-based authentication task on touch-based devices as revealed by a pairwise comparison of device type which showed that text-based passwords were completed significantly faster on standard desktop computers compared to touch-based devices (Text-based: Touch-Desktop: MD=2.83, SE=0.57; $\chi^2(1)=21.385$, $p<0.001$). In contrast, in graphical authentication no significant differences were observed between desktop and touch-based interactions (Graphical: Touch-Desktop: MD=-0.58, SE=0.82; $\chi^2(1)=0.0018$, $p>0.05$).

4.2.4. Main Interaction Effects of Field Dependence-Independence, User Authentication Type and Device Type on Task Completion Time. The analysis revealed a statistically significant three-way interaction between users' field dependence-independence, user authentication type and device type ($\chi^2(4)=27.65$, $p<0.001$). Next, a simple two-way interaction was carried out to determine the statistical significance of the simple two-way interaction (authentication type*device type) for FD users and for FI users. There was a statistically significant simple two-way interaction between user authentication type and device type for FD users ($\chi^2(3)=43.178$, $p<0.001$), as well as for FI users ($\chi^2(3)=8.44$, $p=0.037$). Given the statistically significant simple two-way interaction, we followed these with simple main effects. We investigated the effect of device type at every level of user authentication type. There was a statistically significant simple main effect of device type for FD users in text-based password authentication ($\chi^2(1)=25.59$, $p<0.001$), but not for FD users in graphical authentication ($\chi^2(1)=0.03$, $p>0.05$). Furthermore, we investigated the effect of device type at every level of user authentication type for FI users. There was no statistically significant simple main effect of device type for FI users in text-based password authentication ($\chi^2(1)=3.1642$, $p=0.07$), as well as in the case of graphical authentication ($\chi^2(1)=0.0326$, $p>0.05$). For FD users, time to authenticate in desktop-based interactions was 10.80±2.81 seconds and 15.06±4.24 seconds in touch-based interactions, a difference of 4.26 seconds. Such a result provides initial evidence that the device type affects FD users when interacting with text-based passwords since the shift from a standard keyboard to a virtual keyboard significantly affects the users' visual field and context of use.

4.2.5. Over Time Effects on Task Completion Time. The task completion time was analyzed throughout the 24 sessions aiming to investigate the impact of experience on time to complete the authentication task. Figure 6 illustrates the task completion time comparison between FD and FI users for all combinations across the 24 sessions of the study.

The analysis revealed a main effect of session trials on the time to successfully authenticate indicating that task completion efficiency improves as users gain more experience with the user authentication mechanisms and the device ($\chi^2(1)=205.36$, $p<0.001$). Descriptive statistics reveal that in text-based password interactions (Figure 6, left), FD users logged constantly the highest times to authenticate on touch-based devices across the 24 sessions (FD-Text-Touch), compared to all other combinations (FD-Text-Desktop, FI-Text-Desktop, FI-Text-Touch). Such a result further indicates the increased difficulty of FD users in interacting with text-based passwords on touch-based devices due to the interaction shift from standard keyboard input to touch-based virtual keyboard input. Furthermore, graphical-based interactions (Figure 6, right) revealed that FD users logged the highest times to authenticate on both desktop-based and touch-based devices throughout the 24 sessions compared to FI users, indicating their increased difficulty, compared to FI users, in completing the graphical authentication task. Nevertheless, in the case of FD users interacting with graphical authentication on desktop computers, a considerable decrease of time is observed throughout the 24 sessions. In particular, results revealed that while FD users needed significant more time to login on graphical authentication than FI users in the initial sessions, as they gained experience, time difference between FD-FI users was considerably decreased.

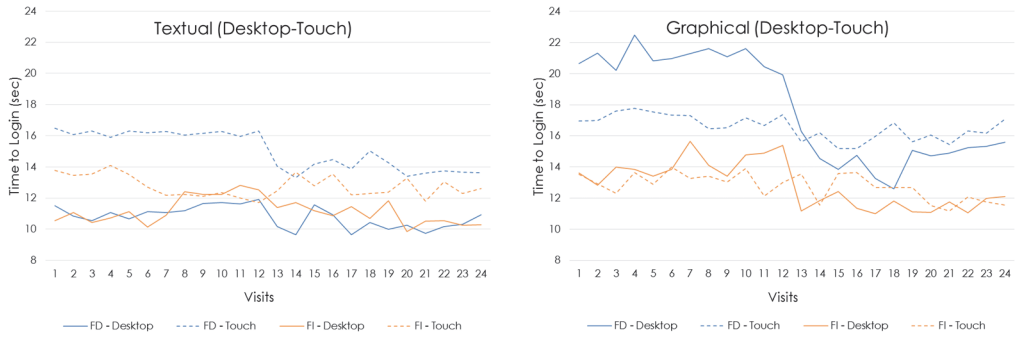


Fig. 6. Task completion time comparison between FD and FI users for all combinations of user authentication and device types throughout the 24 sessions.

4.3 Failure Rate Comparisons

Failure rate was calculated based on the number of sessions that included a failed attempt. A session is considered as failed in case the participant needed more than one attempt to successfully authenticate. For example, in case the participant needed three attempts to authenticate (*i.e.*, the first and second attempt failed and the third succeeded), this session is considered as failed. In contrast, a session is considered as successful when the participant authenticated successfully at first attempt. Among 3674 user authentication sessions, 503 attempts failed (13.69% overall failure rate). Table 2 summarizes the total number of sessions with failed attempts, categorized by field dependence-independence group, user authentication type and device type.

Table 2. Sessions with failed attempts per field dependence-independence, user authentication and device.

	Textual		Graphical	
	Desktop	**Touch**	**Desktop**	**Touch**
FD	48 (9.54%)	59 (11.73%)	80 (15.9%)	90 (17.89%)
FI	53 (10.54%)	55 (10.93%)	71 (14.12%)	47 (9.34%)
Total	101 (20.08%)	114 (22.66%)	151 (30.02%)	137 (27.24%)

Over the span of 24 sessions for each user, we entered a flag indicating whether the particular session was a successful or a failed attempt. Accordingly, we performed a mixed logistic regression with the task effectiveness (successful *vs.* failed) as the dependent variable. The independent variables were used as fixed effects (field dependence-independence, user authentication type, interaction device type), and the subjects as random effects. The analysis revealed that the effect of field dependence-independence on failure rate was not significant ($\chi^2(1)=4.33$, $p=0.21$). Descriptive statistics show that FI users needed less attempts to authenticate than FD users (FD-FI: $MD=-0.35$, $SE=0.16$). Furthermore, the effect of user authentication type on failure was also not significant ($\chi^2(1)=5.06$, $p=0.17$). Descriptive statistics show that users scored higher failure rates on graphical authentication than text-based authentication (Graphical-Textual: $MD=-0.37$, $SE=0.16$). Furthermore, the analysis revealed that FD users had high failures rates in graphical authentication on both desktop and touch-based devices. In particular, 33.8% of all failed attempts (15.9% on desktop computers, 17.89% on touch-based devices) were caused by FD users interacting with graphical authentication mechanisms. Figure 7 illustrates the failure rate per FD user, grouped by user authentication type and interaction device.

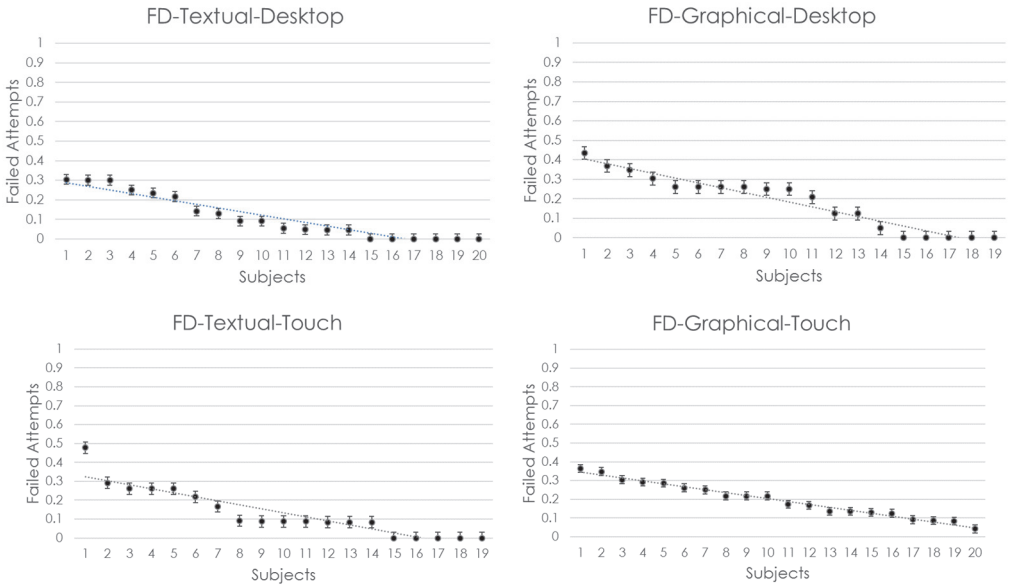


Fig. 7. Failure rate per FD user, grouped by user authentication type and interaction device type.

4.4 Authentication Key Resets

The number of user authentication key resets was calculated per user field dependence-independence groups and user authentication type (Table 3).

Table 3. Number of authentication key resets.

	Textual	Graphical
FD	5	10
FI	3	7
Total	8	17

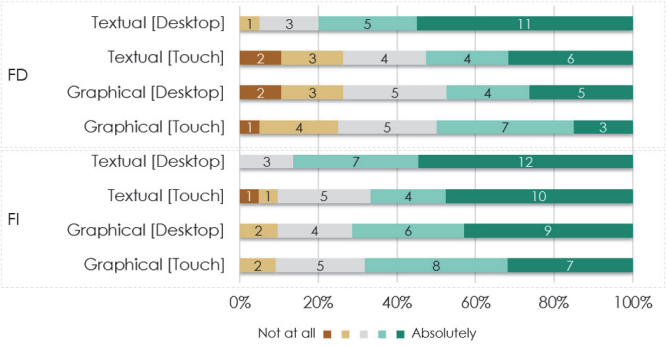
In total, 25 requests were logged for authentication key reset. Among those, 17 users reset their key 1 time, while 4 users reset their key 2 times. The non-parametric Mann-Whitney U test did not reveal significant differences in authentication key requests between all combinations of field dependence-independence and user authentication type. Descriptive statistics show that graphical authentication triggered more key resets than the textual throughout.

4.5 Users' Perceived Usability

A post-study survey was run in order to validate findings of the quantitative analysis as well as to elicit the users' perceptions, perceived usability and memorability based on their interactions with the authentication mechanisms during Phase B. The survey investigated the following factors: *i)* perceived speed; *ii)* perceived ease-of-use; and *iii)* perceived memorability. Example statements of the survey were: "*The [device x authentication] is easy to use*" and "*The [device x authentication] is fast to use*", where *device* and *authentication* was respectively the device and authentication type used by each particular user. Users rated the statements through a 5-point Likert scale (1: Not at all – 5: Absolutely).

In the analysis that follows, the non-parametric Mann-Whitney U test was run on specific group combinations (*field dependence-independence* X *user authentication type* X *device type*) in order to correlate these with the main findings of the quantitative analysis, aiming to increase internal validity and triangulate the results.

4.5.1. Effects of User Authentication and Device Type on Perceived Speed. A high number of participants rated the authentication mechanisms as fast to use (65.85%), whereas 20.73% rated these as neutral (Figure 8).



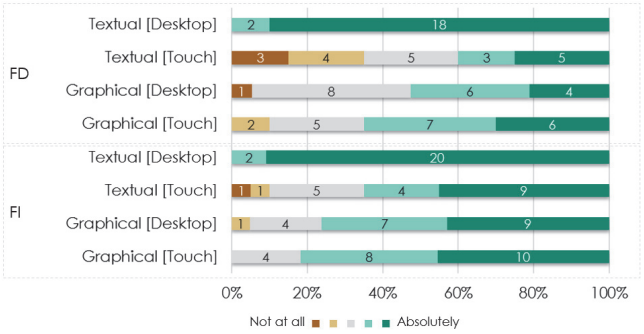
Mann-Whitney U test statistics

Group 1	Group 2	U	z	p
FD-Textual-Desktop	FD-Textual-Touch	124.5	-1.942	0.065
FD-Textual-Desktop	***FD-Graphical-Desktop***	***112.5***	***-2.283***	***0.028***
FI-Textual-Desktop	FI-Graphical-Desktop	189	-1.106	0.269
FI-Textual-Desktop	FI-Textual-Touch	193.5	-0.992	0.321

Fig. 8. Users' responses on the statement: "*The [device x authentication] is fast to use*".

Among the FD user group, a high number of participants rated the graphical authentication mechanism as inefficient to use (12.82%). In addition, a considerable high number of FD users rated touch-based interactions (both text-based and graphical) as not fast to use (12.82%). The Mann-Whitney U test was run in order to determine if the users' responses significantly differ between specific groups. The results revealed that a high number of FD users rated text-based passwords as faster to use on desktop computers than touch-based devices, however, these differences were not significant (U=124.5, z=-1.942, p=0.065). Furthermore, comparisons between text-based passwords and graphical authentication on desktop computers revealed that FD users rated text-based passwords as faster to use (U=112.5, z=-2.283, p<0.05). In contrast, in the case of FI users no significant differences in ratings were observed between all combinations. The results further support the quantitative measures since FD users recorded high times in completing the graphical authentication task across device, and text-based passwords that were deployed on touch-based devices. In contrast, as revealed through the quantitative analysis, FI users did not score significant differences in ratings between user authentication types and device types.

4.5.2. Effects of User Authentication and Device Type on Perceived Ease-of-use. Analogous to perceived speed, the majority of participants (73.17%) agreed that the user authentication mechanisms they interacted with were easy to use. 18.9% provided a neutral rate whereas 7.92% stated that the user authentication mechanism was not easy to use. Among those negative statements, 5.48% (FD users: 7; and FI users: 2) were rated by participants that interacted with textual authentication mechanisms deployed on touch-based devices (Figure 9).



Mann-Whitney U test statistics

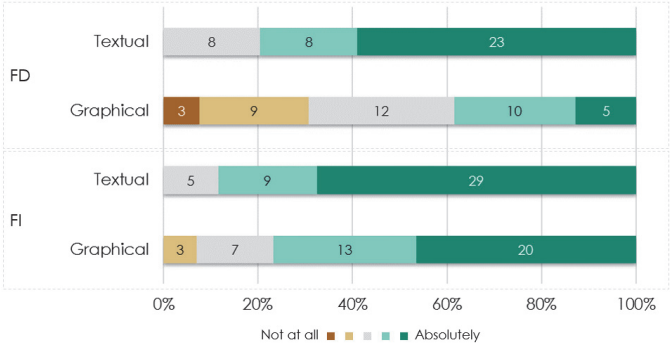
Group 1	Group 2	U	z	p
FD-Textual-Desktop	***FD-Textual-Touch***	***58***	***-4.281***	***<0.001***
FI-Textual-Desktop	***FI-Textual-Touch***	***112***	***-3.331***	***0.001***
FD-Graphical-Desktop	FI-Graphical-Desktop	141	-1.666	0.117
FD-Graphical-Touch	FI-Graphical-Touch	168	-1.388	0.165
FD-Textual-Desktop	***FD-Graphical-Desktop***	***50***	***-4.388***	***<0.001***
FI-Textual-Desktop	***FI-Graphical-Desktop***	***115***	***-3.409***	***0.001***

Fig. 9. Users' responses on the statement: "*The [device x authentication] is easy to use*".

The Mann-Whitney U test was run to determine whether differences exist between several groups on the users' responses regarding ease of use. The analysis further confirms that users (both FD and FI) had increased difficulties in text-based authentication on touch-based devices, compared to desktop computers, as significant differences were observed between desktop-based and touch-based interaction for both user groups. In particular, FD users had significant more positive rates on desktop-based interactions than touch-based interactions (U=58, z=-4.281, p<0.001), so did FI users (U=112, z=-3.331, p<0.02). Furthermore, comparisons of graphical

authentication between FD and FI users revealed no significant differences in user responses for both desktop-based and touch-based interactions. Comparisons between text-based and graphical authentication mechanisms that were deployed on desktop computers revealed that in both FD and FI user groups, text-based passwords were rated as more easy-to-use (FD: U=50, z=-4.388, p<0.001; FI: (U=115, z=-3.409, p<0.02). Such results indicate that although FI users were significantly more efficient and effective in graphical authentication than FD users, their scores in perceived ease of use did not significantly differ. Furthermore, FI users perceived text-based passwords as more easy-to-use than graphical authentication in desktop-based interactions, although the quantitative analysis revealed that FI users interacted similarly well on both mechanisms. In the same line, for text-based passwords deployed on touch-based devices, FI users had higher negative rates than in desktop-based interactions.

4.5.3. Effects of User Authentication Type on Perceived Memorability. Comparison of ratings were conducted across device types since the recall and memorability of authentication key is primarily affected by human cognitive functions and not task execution [Baddeley, 1992; 2012]. Therefore, we grouped the users' responses based on field dependence-independence and user authentication type across device types. The results can be interpreted based on the quantitative analysis (task completion time and failure rate) given that a considerable number of FD users rated graphical authentication keys as not memorable, in contrast to textual password mechanisms, where no user rated low memorability. Figure 10 depicts the users' ratings regarding the memorability of their user authentication key.



Mann-Whitney U test statistics

Group 1	Group 2	U	z	p
FD-Textual	***FD-Graphical***	***305.5***	***-4.732***	***<0.001***
FI-Textual	***FI-Graphical***	***711***	***-2.067***	***0.039***
FD-Textual	FI-Textual	751.5	-0.944	0.354
FD-Graphical	***FI-Graphical***	***421.5***	***-4.006***	***<0.001***

Fig. 10. Users' responses on the statement: *"The user authentication key is memorable".*

A total of 9.14% rated the graphical authentication as not memorable, among them 7.31% were rated by FD users (12) and 1.82% by FI users (3). The Mann-Whitney U test revealed significant differences in memorability ratings between text-based passwords and graphical authentication across FD and FI users. In particular, a significant higher number of responses rated the text-based password as memorable in contrast to graphical authentication (FD: U=305.5, z=-4.732, p<0.001; FI: (U=711, z=-2.067, p<0.04). Furthermore, no significant differences were observed between FD and FI users in regards with memorability ratings in text-based passwords. On the other hand, a significant higher number of FI users rated the

graphical authentication as more memorable than FD users (U=421.5, z=-4.006, p<0.001).

5. SECURITY ANALYSIS

We present a security analysis based on the generated authentication keys with the aim to contextualize the evaluation results and get empirical insights about the participants' intentions in using secure authentication keys during the study.

The analysis focused on resistance to offline brute-force attacks since this is an important and highly researched metric for evaluating the security of authentication mechanisms [Komanduri et al., 2011; Biddle et al., 2012]. An evaluation of other types of attacks (*e.g.*, online guessing attacks, shoulder surfing, phishing, etc.) is out of the scope of this analysis since these types of attacks are already bound to the core design of each mechanism and are not highly affected by the user study's dynamics. For example, online guessing attacks can be prevented in both authentication mechanisms through Human Interaction Proof mechanisms (*e.g.*, CAPTCHA) that can be enabled after multiple unsuccessful user logins. The interested reader may find an analysis and theoretical comparison of such attacks between text-based passwords and graphical authentication mechanisms in Biddle et al. [2012].

5.1 Ethical and Privacy Considerations

Aiming to effectively quantify the generated authentication keys, we required access to the plaintext of the generated keys which raises several security, privacy and ethical issues. From a security perspective, in such a scenario, a participant could reuse a password from an existing account since research has shown that people reuse passwords across multiple accounts [Florencio and Herley, 2007]. Access to the plaintext password by a malicious person raises severe security issues for the participants. In order to avoid this issue, and at the same time quantify the security strength of the keys for the purpose of this research, we hashed and stored the authentication key in the database of the actual user authentication scheme, along with the rest data of the user (*e.g.*, username, demographics, etc.). Instead, the plaintext authentication key (along with the plain FD/FI classification and assigned interaction device type) was stored in a separate location during user enrolment (or during key reset), without any binding information that could relate the authentication key to any particular user.

5.2 Strength of Textual Keys

We assessed the strength of the textual password keys using Carnegie Mellon University's Password Guessability Service (PGS) [Ur et al., 2015]. PGS estimates plaintext passwords' "guessability"; how many guesses a particular password-cracking algorithm with particular training data would take to guess a password. For running the password guessability calculations, PGS uses four high-level approaches to password cracking: *i)* using the software tool oclHashcat; *ii)* using the software tool John the Ripper; *iii)* using probabilistic Markov models; and *iv)* using a probabilistic context-free grammar implementation (PCFG).

The password cracking approach based on Markov models could not crack any of the passwords, while the Hashcat approach cracked 4 out of 82 passwords (*min number of guesses:* 10^{10}; *max number of guesses:* 10^{12}). PCFG was more effective than the other approaches, by cracking 37 out of 82 passwords. Figure 11 illustrates the number of guesses the approach required to crack each of the 37 passwords, and in which group the password was used (field dependence-independence and device type).

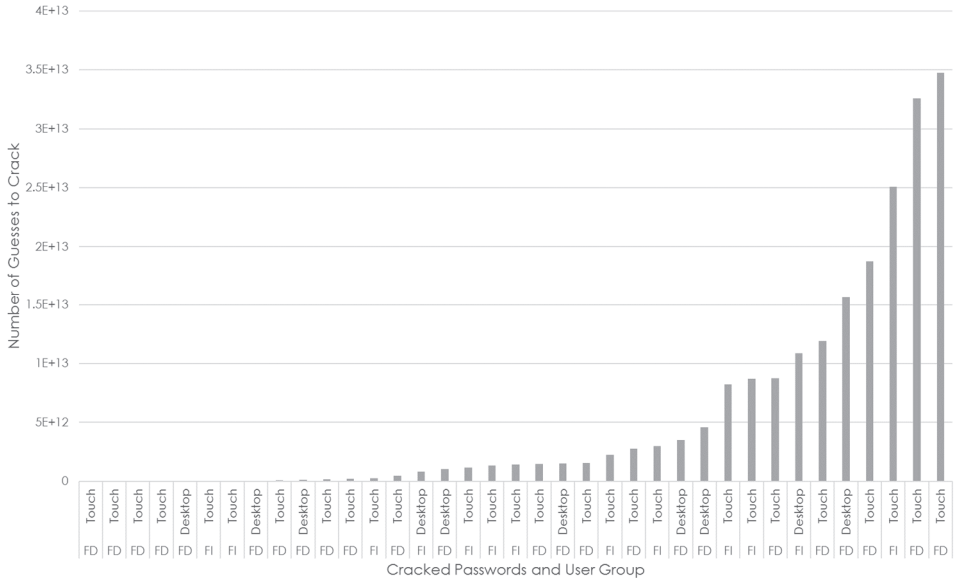


Fig. 11. Number of guesses required for cracking the 37 passwords using PCFG.

Noteworthy is the fact that text-based passwords used on touch-based devices were successfully cracked by PCFG (27 out of 37). This denotes that the interaction device type might have an influence on password selection as users created easier to crack text-based passwords on touch-based devices than on desktop computers. Such a result is in line with prior research that has shown that password-entry on mobile devices can increase insecure behavior [von Zezschwitz et al., 2014].

5.3 Strength of Graphical Keys

The security analysis for graphical keys focuses on: *i)* the theoretical security strength of graphical keys, which is measured by the average combinations required to crack a password selected at random; and *ii)* the practical security strength of graphical keys, which is calculated by measuring the resistance of user-selected passwords to an offline brute-force attack. For calculating practical strength, we implemented a brute-force attack that checks all possible permutations of graphical keys comprising five images, starting from the upper left corner of the image grid and traversing it row-by-row. We measure practical strength by calculating the average "guesses" performed per user until each corresponding password is guessed correctly. Table 4 summarizes the theoretical and practical security strength. The results suggest that the practical security strength across user groups closely matches the corresponding theoretical strength which indicates that, in practice, the graphical approach is effective in terms of security, with respect to the theoretical limits of an attacking algorithm.

Table 4. Theoretical vs. practical security strength per user group.

Theoretical Strength	Practical Strength (Variation)			
	FD-Desktop	FD-Touch	FI-Desktop	FI-Touch
3,187,800	3,014,422 (-5.43%)	3,100,783 (-2.72%)	3,178,991 (-0.27%)	3,108,143 (-2.49%)

6. DISCUSSION OF MAIN FINDINGS

Results have shown several main effects between users' field dependence-independence, user authentication type and device type. Next, we report the main findings and interpretations based on the analysis of results.

Finding A - Task completion time in graphical authentication significantly differs between FD and FI users across interaction device types. Results have shown a main effect of field dependence-independence on graphical authentication (*Figure 1, Axis A-Human (FD, FI), and Axis C-User Authentication (graphical)*). In particular, FI users needed significantly less time to complete the graphical authentication task compared to FD users across device type. An interpretation of the usability results could be based on the particular stimuli and interaction design of the graphical authentication mechanism, *i.e.*, in the case of graphical authentication, homogeneous objects and structure/organization are illustrated to the users, in which the surrounding framework might dominate the perception of the aiming items within. Accordingly, when FD users interact with these types of stimuli, they might find it difficult to locate the information they are seeking because other information might mask what they are looking for. On the other hand, FI users find it easier to recognize and select the important information from its surrounding field due to their improved dis-embedding skills and visual search task abilities [Angeli et al., 2009]. Furthermore, when information is presented in an ambiguous, unstructured format, FI users impose their own structure on the information, while FD users attempt to first understand and learn that information as it is presented and without restructuring it. This might explain the added difficulty and time in completing the graphical authentication task compared to FI users. This finding is also in accordance with previous research, which has shown that FI users have a tendency to use their own internal references whereas FD users rely more on external frames of reference [Chen and Liu, 2008]. A recent eye tracking study [Katsini et al., 2017; Belk et al., 2017a] that was based on a similar experimental setup, revealed that FI users spent significantly more time and fixated cumulatively on more images during graphical key creation compared to FD users. Such a behavior could explain Finding A, since FI users might have better memorized their graphical key by paying attention to detail and through deeper information processing, thus positively affecting the task login efficiency. Finally, given that FI users have an enhanced visual working memory in contrast to FD users [Miyake et al., 2001; Rittschof, 2010], FI users might have been positively affected compared to FD users in graphical tasks since the images are primarily processed through the visual working memory sub-system [Baddeley, 1992; 2012].

Finding B - In desktop-based environments, graphical authentication is less usable than text-based authentication for FD users, but not for FI users. In line with Finding A, results revealed that in desktop-based interactions, FD users needed significantly more time to complete the graphical authentication task compared to the text-based task. A possible interpretation could be based on the familiarity users had in text-based authentication over graphical authentication. On the other hand, FI users did not have significant differences in task completion time and failure rate between the two user authentication mechanisms. In this respect, since the familiarity factor did not affect the graphical authentication mechanism, FI users had better performance than FD users in graphical authentication, and recorded similar performance with text-based password authentication which might

be accredited to their dis-embedding skills, improved visual working memory and visual search task abilities [Angeli et al., 2009].

Finding C - Time to complete text-based authentication significantly differs between desktop computers and touch-based devices, but not for graphical authentication. Results revealed that the device affects text-based password interactions (*Figure 1, Axis B-Device (desktop, touch), and Axis C-User Authentication (textual)*). In particular, analyses of touch-based user interactions revealed that in general, the time to complete text-based password tasks was significantly longer than desktop-based user interactions. Such a result increases the external validity of prior research results which revealed that entering text-based passwords on mobile touch-based devices is more time consuming and considered a more demanding task compared to standard desktop computers [von Zezschwitz et al., 2014; Findlater et al., 2011]. In regards with graphical authentication mechanisms, no significant differences were observed between desktop-based and touch-based interactions. Such a result might be explained by the fact that selecting images through the computer mouse or through touch on the screen is fairly the same task execution process (*i.e.*, selecting the images through computer mouse clicks *vs.* finger touch on the screen).

Finding D - The interaction device in text-based authentication significantly affects FD users' task completion time, but not FI users. Analysis of results has shown that the task performance of FD and FI users has been affected by specific combinations of user authentication and device types, given the users' different cognitive processing abilities, and adaptation skills within contextual shifts (desktop *vs.* touch-based) (*Figure 1, Axis A-Human (FD, FI), Axis B-Device (desktop, touch), and Axis C-User Authentication (textual, graphical)*). Results revealed that FD users had a significant increase of time to complete the text-based password task on touch-based devices compared to desktop computers, whereas FI users did not have significant differences in task completion time between the two interaction device types. Furthermore, the analysis of desktop-based interactions revealed no significant differences between FD and FI users in text-based passwords, which might be explained by their familiarity with text-based passwords. However, in touch-based interactions, FI users were significantly faster in completing the textual password task compared to FD users due to their positive adaptation and independence in regards with contextual and field changes (desktop *vs.* touch-based). These results suggest that the device, and eventually the field change, towards touch-based interactions (context-wise and interaction-wise) was adopted more efficiently and effectively by FI users compared to FD users. These findings strengthen the claim of previous research, 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 [Davis, 1991; Messick, 1993; Goodenough and Karp, 1961; Reardon and Moore, 1988].

Finding E - Security strength in graphical authentication is similar across field dependence-independence groups and interaction device type, but not in text-based authentication. The security strength of graphical authentication keys is similar among FD and FI users across interaction device types. In addition, the theoretical and practical strength of graphical keys is fairly similar indicating that users across different groups chose random authentication keys with respect to the supported security policy. Security strength of text-based passwords for both FD and FI groups, that were deployed on desktop computers was higher than touch-based devices. This indicates insecure user behavior given the difficulty of password

entry on visual keyboards and increases external validity of prior research [von Zezschwitz et al., 2014] that observed similar differences in user behavior. It is worthwhile mentioning that in both authentication types, text-based and graphical, the security analysis indicates acceptable security levels and randomness in the selection of authentication keys which suggests that participants' selections were not influenced by any particular biased factors which could decrease the ecological validity of the study. This is also supported by the analysis of CMU PGS which suggests that users generated strong and hard to guess text-based passwords since the majority of passwords were not easily cracked.

7. IMPLICATIONS

The discussion of results suggests that human cognitive differences in field dependence-independence, the interaction device and the authentication type affect task completion performance of user authentication. FD and FI individuals have particular characteristics that influence their performance when interacting with different user authentication types and different interaction devices. This denotes that knowledge-based user authentication mechanisms could eventually embrace both text-based password and graphical authentication mechanisms, bootstrapped on the unique characteristics of field dependence-independence and the interaction device, with the aim to improve the task usability and eventually provide a positive user experience.

Accordingly, the main results of this study could be transformed into specific context-based recommendation rules, and further applied in a method for suggesting the "best-fit" user authentication design factors by considering human and technology factors. Algorithm #1 presents the method in pseudo-code for recommending a user authentication type given a user's individual context model ($UCM_j(u_i)$) that consists of a set of human and technology factors and their corresponding values.

ALGORITHM 1. USER AUTHENTICATION RECOMMENDATION (*UAR*)

Input: A set of individual context models *UCM* and a user u_i.
Output: Assign user authentication type *ua_type* of user u_i with a recommendation *r*.

```
1:  method: SecurityPolicyA.UAR(UCM, u_i)
2:      r = null;
3:      // create UF by extracting a set of tuples (fc_i, val) from UCM for user u_i
4:      UF = { (fc_i, val) : (u_k, fc_z, val) ∈ UCM and u_i = u_k }
5:      if ( (cognitive, FI) ∩ UF ≠ ∅ and
6:                              ( (device, desktop) ∩ UF ≠ ∅ or (device, touch) ∩ UF ≠ ∅ ) )
7:          r = any; // give the option to choose between textual or graphical authentication
8:      else if ( (cognitive, FD) ∩ UF ≠ ∅ and (device, desktop) ∩ UF ≠ ∅ )
9:          r = textual;
10:     else if ( (cognitive, FD) ∩ UF ≠ ∅ and (device, touch) ∩ UF ≠ ∅)
11:         r = any;
12:     else if ( (device, touch) ∩ UF ≠ ∅ )
13:         r = graphical;
14:     else if ( (device, desktop) ∩ UF ≠ ∅ )
15:         r = any;
16:     else
17:         r = any;
18:     UCM = UCM ∪ (u_i, ua_type, r);
19: end method
```

The method would run during user registration or during user login in the system. Given a specific user u_i, the method extracts the values of all the factors (*UF*) from the user's individual context model and accordingly recommends a specific user authentication type based on several context-based recommendation rules which reflect the observed main effects of this user study. In case a user has already enrolled with a particular user authentication type, and over time, the user's history log of interactions suggests a different authentication type (*e.g.*, the user frequently uses a different device type as the previous one; shift from desktop computer to mobile touch-based device), the system would provide the new recommendation in the form of a notification alert, suggesting the user to change the authentication type.

7.1 Recommendation Rules

RECOMMENDATION RULE #1. Results have shown that FI users perform equally well on both text-based password mechanisms and graphical authentication mechanisms, since results did not reveal significant differences between the two user authentication types in terms of time and failure rate (Finding A and B). Accordingly, following the approach in Forget et al. [2014] that proposed a multi-type authentication framework in which users select the preferred user authentication type, Recommendation Rule #1 gives the option to FI users to choose the preferred user authentication type (*Algorithm #1, line 5-7*).

RECOMMENDATION RULE #2. Results have shown that FD users needed significantly more time to complete the graphical authentication task than the text-based task on desktop computers (Finding B). Accordingly, Recommendation Rule #2 recommends a text-based password mechanism to FD users when interacting on desktop computers, since a graphical authentication mechanism would significantly decrease the task usability for those users (*Algorithm #1, line 8-9*).

RECOMMENDATION RULE #3. Results have shown that the interaction device negatively affects FD users in text-based password authentication (Finding D). In addition, given that inferential statistics did not reveal significant differences between text-based passwords and graphical authentication on touch-based devices for FD users, Recommendation Rule #3 provides the option to FD users to choose between both user authentication types, in case the visitation history log of those users indicates a high frequency of system access through touch-based devices. (*Algorithm #1, line 10-11*).

RECOMMENDATION RULE #4. Overall results have shown that touch-based devices significantly decrease the task completion time for text-based password authentication (Finding C) due to the increased difficulty of text entry on virtual keyboards [von Zezschwitz et al., 2014; Findlater et al., 2011]. Thus, in case the cognitive characteristic of the user is not provided to the system, Recommendation Rule #4 recommends a graphical authentication mechanism for touch-based user interactions, irrespective of the cognitive factor. (*Algorithm #1, line 12-13*).

RECOMMENDATION RULE #5. Following the approach in Forget et al. [2014] that proposed a multi-type authentication framework in which users select the preferred user authentication type, Recommendation Rule #5 gives the option to the user to choose the preferred user authentication type in case the cognitive characteristic of the user is not provided to the system and the user interaction takes place on a desktop computer. (*Algorithm #1, line 14-15*).

From a practical perspective, the aforementioned recommendation rules could be applied through generic adaptation and personalization frameworks (*e.g.*, Fidas et al.

[2015]) which will personalize user authentication tasks (and other user-security tasks, *e.g.*, CAPTCHA tasks as in Belk et al. [2015b]), considering human cognitive differences and the interaction device type. Figure 12 illustrates a personalization paradigm in which the "best-fit" user authentication design (textual *vs.* graphical), framed by a given security policy, is provided to the users considering human factors (field dependence-independence) and technology factors (device type). As shown in Figure 12, such a personalization approach would support multiple security policies, aligned to multiple individual context models' groups. It would also allow service providers to fully customize usability aspects of user authentication towards the benefit of end-users.

In this context, future studies are necessary to incorporate these factors and recommendation rules under an operable personalization system aiming to increase the validity of this research and evaluate the users' acceptance, experience and efficacy of the proposed approach.

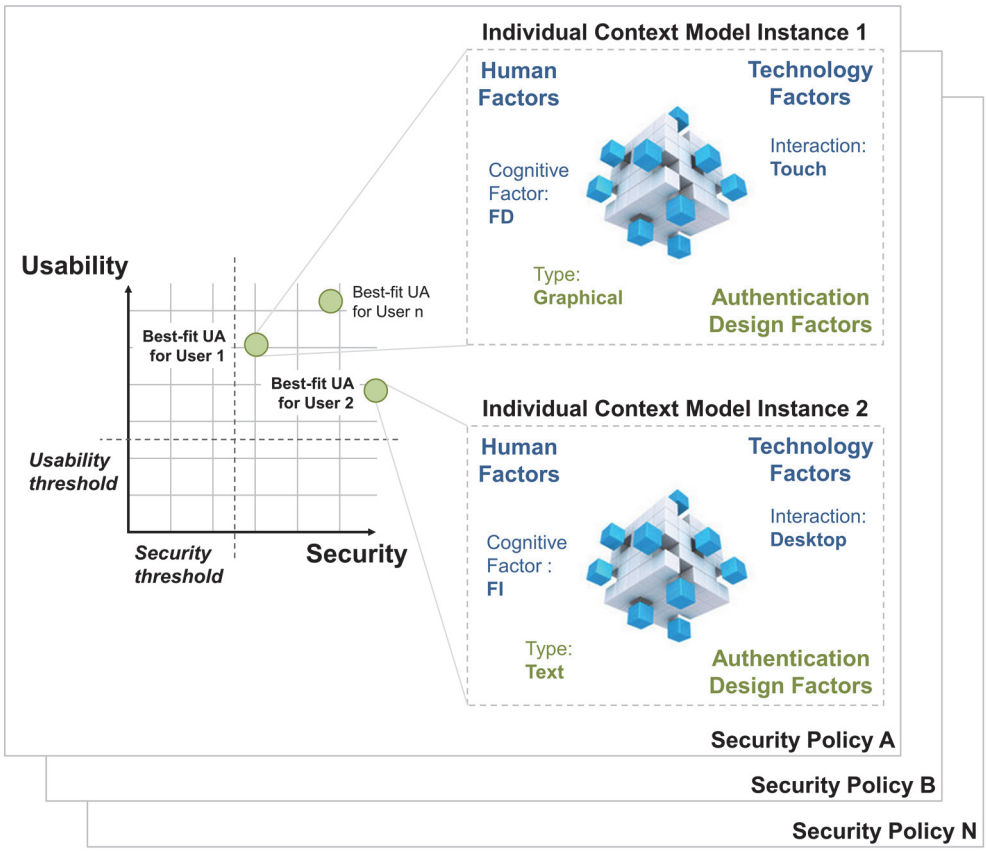


Fig. 12. "Best-fit" authentication type based on human and technology factors, framed by a security policy.

8. LIMITATIONS AND VALIDITY OF THE STUDY

This research work entails a number of limitations given the intrinsic complexity and multi-dimensional nature of the factors investigated. An important limitation is related to the rather limited number and non-varying user profiles of the sample since undergraduate students were recruited for conducting the study. Nonetheless, given that high-level cognitive styles are stable and do not significantly change throughout age [Peterson et al., 2009], we expect that replicating the experiment in a

different age population would probably reveal the same effects as presented in this paper. Apparently, the authentication performance of users would be different between age groups (younger *vs.* older users), however this would be mainly due to variations in elementary cognitive processing, and physical motor and perceptual abilities (as shown in the literature, *e.g.*, [Nicholson et al., 2013a; 2013b; Ma et al., 2013]), but not due to high-level cognitive factors which is the primary factor under investigation in our sample. Another limitation concerns the classification of users into FD and FI. Given that the GEFT highlights cognitive differences along a continuum scale, the use of a cut-off score [Hong et al., 2012; Altun and Cakan, 2006] might not accurately classify intermediate individuals that fall in between the two end points (*e.g.*, Field-mixed [Angeli et al., 2009]). Nevertheless, it is important to stress that the frequencies of user's scores on the GEFT in our sample is comparably similar to general public GEFT scores as shown in several studies which embraced individuals with different demographics [Rittschof and Chambers, 2005; Robinson et al., 2009; Khatib and Hosseinpur, 2011; Davis, 2006].

Furthermore, there has been an effort to increase the internal validity of the research since our sample involved rather experienced and average, than novice users with respect to user authentication and interaction with desktop computers and mobile touch-based devices. Inevitably, the sample included users who were more familiar with text-based passwords than with recognition-based graphical authentication, since text-based authentication mechanisms are currently the most popular and widely deployed [Herley and van Oorschot, 2012; Chiasson et al., 2009].

There was also an attempt to increase ecological and external validity. Hence, we embedded the user authentication mechanisms in a real Web-site of a university course's laboratory that was used by students in a real-life scenario for accessing and uploading course material throughout an academic semester. This way, students interacted in a physical environment as they would normally do, without the involvement of any experimental equipment or person. Another concern relates to the incentives participants had to select a "secure" authentication key rather than an easy one to remember and type in the context of an experimental study; prior research has shown that users select authentication keys differently for higher valued accounts [Florencio and Herley, 2007]. As such, the experimental setup aimed to create an environment that students would value and protect access to their accounts since the Web-site was used to perform important tasks in the context of their university course, *e.g.*, upload and manage their exercises, view lab grades, etc. This process was also strengthened during user enrolment, in which students were given incentives to create a secure authentication key and follow several rules for keeping their authentication key safe (*e.g.*, not share the key with other people, not write down the key, etc.). In this context, it is important to stress that the security analysis revealed that users valued their accounts. In text-based passwords, the analysis based on CMU PGS revealed that the majority of text-based passwords were hard to crack. For graphical authentication, the practical security strength of the generated keys was fairly similar to the theoretical strength.

However, there are limitations related to the ecological and external validity of the study. First, controlling the policy (*e.g.*, requiring a specific length of authentication key and specific length of upper-case characters, numbers and special characters) might not fully reflect the behavior of users in real-life situations. For example, prior work has shown that users tend to use more than the minimum length of required policies in order to meet complex requirements in a memorable way [Shay et al., 2010]. Furthermore, two particular authentication mechanisms

were investigated although numerous other mechanisms and features, especially in graphical authentication are worth of examination. In this context, we underpin that recognition-based graphical authentication mechanisms entail a number of features that need further investigation since these might affect differently the usability and security aspects of user authentication mechanisms, *e.g.*, the types of images used in the challenge (faces, abstract or single-object), the number of user-selected and decoy images used in the authentication process, the policy and the procedure for authentication (*e.g.*, allow the reselection of images in a single key, show decoy images in one screen or in multiple screens) [Ma et al., 2013].

The aforementioned decisions were taken under the following perspectives. First, a fixed authentication key policy was chosen as more appropriate aiming to control the security aspects of each mechanism and thus allowing us more accurately investigating the effects of human cognitive differences in user authentication task performance. Furthermore, we applied the specific authentication policies which embraced 8 characters for text-based and 5 images for graphical authentication having in mind that the former policy entails a common character length and password facets that have been used in practice to generate *"medium complexity internet passwords"* [von Zezschwitz et al., 2013; 2014], while the latter policy generates graphical authentication keys and key spaces similar to other authentication schemes used in practice [Dhamija and Perrig, 2000; PassFaces, 2009; Mihajlov and Jerman-Blazic, 2011; Ma et al., 2013]. In addition, a challenge of the reported study relates to the choice of user authentication mechanisms for investigation. Recognition-based graphical authentication was intentionally chosen since the task of entering the graphical authentication key primarily involves a visual search task in which users are required to recognize their selected images among other decoy images, and thus might be affected by human cognitive differences in field dependence-independence in the way they interact with the stimuli and locate pertinent information to their objective. For example, it might not be that straightforward for a particular user to recognize and recall a target object (*e.g.*, basketball) among a number of heterogeneous objects. Nonetheless, the design and results of the study cannot be generalized to all graphical authentication mechanisms that exist in the literature, since a number of other factors exist that need to be addressed in the design of such mechanisms.

With regards to the suggested recommendation rules, we stress that these embrace new challenges from the security and technology perspective that need closer attention. Given that the suggested recommendation rules depend on the factors of the individual context model, the main skepticism on the practical feasibility of such an approach is focused on the required prior knowledge of the system on the users' cognitive characteristics, which are necessary to personalize the user authentication task. In this context, implicit user data collection methods for field dependence-independence elicitation could be based on the users' interactions with the system. Such methods would increase user acceptance of the approach since the field dependence-independence characteristics could be transparently elicited based on the users' interactions with the system, without requiring to conduct any additional psychometric tests that would add a burden to the users. Accordingly, the users' field dependence-independence characteristics could be implicitly inferred by tracking their behavior with navigation tools [Chen and Liu, 2008], or based on the usage of search tools [Chan et al., 2014].

With regards to security, given that each user authentication scheme has different strengths and weaknesses [Renaud et al., 2013; Biddle et al., 2012], the

recommendation of a particular user authentication type would change the security metrics of the authentication mechanism. Accordingly, the recommendation rules could be further extended with several security factors defined by the system administrator depending on the application domain and custom requirements [Fidas et al., 2015; Belk et al., 2015b]. For example, current recognition-based graphical authentication schemes have a smaller theoretical key space than text-based passwords (assuming a set of images whose cardinality remains reasonable with respect to usability) [Biddle et al., 2012], making these more vulnerable to offline guessing attacks. Thus, depending on the application domain, additional mechanisms could be enabled in case a graphical authentication type would be recommended for preventing these attacks (*e.g.*, TwoStep Authentication based on a combination of text-based and graphical authentication [van Oorschot and Wan, 2009]). In case the service provider desires to use a single method of authentication but increase the theoretical space of the graphical authentication key, a different policy could be applied requiring users to choose a higher number of images and present a higher number of decoy images in the challenge set during authentication.

9. CONCLUSIONS

The purpose of this paper was to investigate the interplay among human cognitive processing differences (field dependence *vs.* field independence), alternative interaction device types (desktop *vs.* touch) and user authentication schemes (textual *vs.* graphical) towards task completion efficiency and effectiveness. In this context, we have designed a three-phase experimental study which entailed a credible psychometric-based survey for eliciting users' cognitive characteristics based on the field dependence-independence theory. A total of 164 participants were classified as field dependent (FD) and field independent (FI) based on their performance on the Group Embedded Figures Test (GEFT). Following a between-subjects study design, we split randomly FD and FI user groups, and formed four sub-groups which interacted with two types of user authentication mechanisms (textual or graphical) that were deployed on standard desktop computers or mobile touch-based devices. The study took place for a period of three months and a total of 3674 user authentication sessions have been recorded. Analysis of results, that were based on objective and subjective measures, revealed significant interaction effects among several human, technology and user authentication design factors.

In this context, this paper contributes in terms of *theory* and *application*. Regarding *theory*, findings of the study indicates that socio-cognitive theories, like the field dependence-independence theory, can be considered as applicable analysis frameworks in further understanding user authentication tasks. Such frameworks are necessary within nowadays complex computation realms in which user authentication takes place. In particular, results of the study can be interpreted under the light of the field dependence-independence theory as they revealed a main effect of human cognitive differences on task performance of different user authentication mechanisms and interaction devices.

Regarding *application*, the findings of the study underpinned the added value of adaptivity in user authentication and accordingly, we have identified several context-based recommendation rules for delivering personalized user authentication tasks based on a combination of human, technology and design factors. We envision that building such personalized user authentication mechanisms would have many positive implications from the users' point of view. Providing user authentication tasks, personalized to the users' cognitive characteristics would support the users'

efficiency of processing information cognitively as well as task execution performance, and eventually improve the user experience and user acceptance of such tasks.

Finally, we stress that an interesting implication of our work is related to the ecumenical character of user authentication tasks. Scholars have provided evidence that the differences in cognitive processing styles and abilities exist not only within a certain nation, but as well across diverse nations around the globe, as they are affected by the cultural background in which they are developed [Cui et al., 2013; Varnum et al., 2010; Engelbrecht et al., 1997]. From this perspective, given the globalization of Information Technology applications and services, studies like the one reported herein could be replicated on a multinational scale aiming the design and development of globalized user authentication schemes, which will have an impact on a high number of individuals. In this context, bearing in mind that prior research has shown cross-cultural differences in field dependence-independence, *e.g.* between Eastern societies and Western societies [Cui et al., 2013; Varnum et al., 2010], and African Americans and South Africans [Engelbrecht et al., 1997], future work entails exploring the effects of cultural differences on user authentication across different nations.

REFERENCES

Alshehri, M., & Crawford, H. (2016). Using image saliency and regions of interest to encourage stronger graphical passwords. In *Proceedings of the 32nd Annual Conference on Computer Security Applications (ACSAC 2016)*, ACM Press, 127-138.

Altun, A., & Cakan, M. (2006). Undergraduate students' academic achievement, field dependent/independent cognitive styles and attitude toward computers. *Educational Technology & Society, 9*(1), 289-297.

Angeli, C. (2013). Examining the effects of field dependence-independence on learners' problem-solving performance and interaction with a computer modeling tool: Implications for the design of joint cognitive systems. *Computers & Education, 62*, 221-230.

Angeli, C., & Valanides, N. (2004). Examining the effects of text-only and text-and-visual instructional materials on the achievement of field-dependent and field independent learners during problem-solving with modeling software. *Educational Technology Research and Development, 52*(4), 23-36.

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.

Baddeley, A. (1992). Working memory. *Science, 255* (5044), 556-559.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*, 1-29.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Belk, M, Fidas, C., Germanakos, P., & Samaras, G. (2015a). A personalised user authentication approach based on individual differences in information processing. *Interacting with Computers*, 27(6), 706-723.

Belk, M., Fidas, C., Germanakos, P., & Samaras, G. (2015b). Do human cognitive differences in information processing affect preference and performance of captcha?. *International Journal of Human - Computer Studies*, 84, 1-18.

Belk, M., Fidas, C., Katsini, C., Avouris, N., & Samaras, G. (2017a). Effects of human cognitive differences on interaction and visual behavior in graphical user authentication. In *Proceedings of the IFIP TC13 Conference on Human-Computer Interaction (INTERACT 2017)*, Springer-Verlag (to appear)

Belk, M., Pamboris, A., Fidas, C., Katsini, C., Avouris, N., & Samaras, G. (2017b). Sweet-spotting security and usability for intelligent graphical authentication mechanisms. In *Proceedings of the ACM Conference on Web Intelligence (WI 2017)*, ACM Press (to appear)

Belk, M., Papatheocharous, E., Germanakos, P., & Samaras, G. (2013). Modeling users on the world wide web based on cognitive factors, navigation behaviour and clustering techniques. *Journal of Systems and Software, 86*(12), 2995-3012.

Biddle, R., Chiasson, S., & van Oorschot, P. (2012). Graphical passwords: Learning from the first twelve years. *ACM Computing Surveys, 44*(4), 41 pages.

Brostoff, S, Sasse, M.A. (2000). Are passfaces more usable than passwords? A field trial investigation. In *Proceedings of People and Computers - Usability or Else (HCI 2000)*, Springer-Verlag, 405-424.

Cao, K., & Jain, A. (2016). *Hacking mobile phones using 2D printed fingerprints*. MSU Technical Report

MSU-CSE-16-2.
Chan, C., Hsieh, C., & Chen, S. (2014). Cognitive styles and the use of electronic journals in a mobile context. *Documentation, 70*(6), 997-1014.
Chen, S., & Liu, X. (2008). An integrated approach for modeling learning patterns of students in web-based instruction: A cognitive style perspective. *ACM Transactions on Computer-Human Interaction, 15*(1), Article 1, 28 pages.
Chiasson, S., Forget, A., Stobert, E., van Oorschot, P., & Biddle, R. (2009). Multiple password interference in text passwords and click-based graphical passwords. In *Proceedings of the ACM Conference on Computer and Communications Security (CCS 2009)*, ACM Press, 500-511.
Chowdhury, S., Poet, R., Mackenzie, L. (2013). A comprehensive study of the usability of multiple graphical passwords. In *Proceedings of the IFIP TC13 Conference on Human-Computer Interaction (INTERACT 2013)*, Springer-Verlag, 424-441.
Clewley, N., Chen, S.Y. & Liu, X. (2011). Mining learning preferences in web-based instruction: holists vs serialists", *Educational Technology & Society, 14*(4), 266-277.
Clewley, N., Chen, S.Y., & Liu, X. (2010). Cognitive styles and search engine preferences: Field dependence/independence vs. holism/serialism. *Journal of Documentation. 66*(4), 585-603.
Cui, G., Liu, H., Yang, X. & Wang, H. (2013). Culture, cognitive style and consumer response to informational vs. transformational advertising among East Asians: Evidence from the PRC. *Asia Pacific Business Review, 19*(1), 16-31.
Davis, D., Monrose, F., & Reiter, M. (2004). On user choice in graphical password schemes. In *Proceedings of the USENIX Conference on Security Symposium (SSYM 2004)*, USENIX Association, 11-11.
Davis, G. A. (2006). Learning style and personality type preferences of community development extension educators. *Journal of Agricultural Education, 47*(1), 90-99
Davis, J. K. (1991). Educational implications of field dependence–independence. In S. Wapner & J. Demick (Eds.), *Field dependence–independence: Cognitive style across the lifespan* (pp. 149–175). Hillsdale, NJ: Lawrence Erlbaum Associates.
Dhamija, R., & Perrig, A. (2000). DejaVu: A user study using images for authentication. In *Proceedings of the USENIX Security Symposium*, USENIX Association.
Dinno, A. (2015). dunn.test: Dunn's Test of Multiple Comparisons Using Rank Sums. R package version 1.3.1. http://CRAN.R-project.org/package=dunn.test
Engelbrecht, P., Engelbrecht, P., Natzel, S., & Natzel, S. (1997). Cultural variations in cognitive style: Field dependence vs field independence. *School Psychology International, 18*(2), 155-164.
Fidas, C., Hussmann, H., Belk, M., Samaras, G. (2015). iHIP: Towards a user centric individual human interaction proof framework. In *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems (CHI 2015)*, ACM Press, 2235-2240.
Findlater, L., Wobbrock, J., & Wigdor, D. (2011). Typing on flat glass: Examining ten-finger expert typing patterns on touch surfaces. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2011)*, ACM Press, 2453-2462.
Florencio, D., & Herley, C.A. (2007). Large-scale study of web password habits. In *Proceedings of the ACM Conference on World Wide Web (WWW 2007)*, ACM Press, 657-666.
Forget, A., Chiasson, S., & Biddle, R. (2014). Towards supporting a diverse ecosystem of authentication schemes. In Proceedings of the Who are you?! Adventures in Authentication Workshop (WAY 2014) at the Symposium on Usable Privacy and Security (SOUPS 2014), USENIX Association.
Goodenough, D. R., & Karp, S. A. (1961). Field dependence and intellectual functioning. *Abnormal and Social Psychology, 63*, 241–246.
Hayashi, E., Hong, J., & Christin, N. (2011). Security through a different kind of obscurity: Evaluating distortion in graphical authentication schemes. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2011)*, ACM Press, 2055-2064.
Herley, C., & van Oorschot, P. (2012). A research agenda acknowledging the persistence of passwords. *Security and Privacy, 10*(1), 28-36.
Hong, J., Hwang, M., Tam, K., Lai, Y., & Liu, L. (2012). Effects of cognitive style on digital jigsaw puzzle performance: A GridWare analysis. *Computers in Human Behavior, 28*(3), 920-928.
Hudson, L. (1966). Contrary imaginations: a psychological study of the English school boy. London: Methuen.
Jonassen, D.H., & Grabowski, B.L. (1993). Field dependence and field independence (Global vs. Articulated Style). Handbook of individual differences, learning, and instruction. London, England: Lawrence Erlbaum Associates.
Katsini, C., Fidas, C., Belk, M., Avouris, N., & Samaras, G. (2017). Influences of users' cognitive strategies on graphical password composition. In *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems (CHI 2017)*, ACM Press, 2698-2705.
Khatib, M., & Hosseinpur, R.M. (2011). On the validity of the group embedded figure test (geft). *Journal of Language Teaching and Research, 2*(3), 640-648.
Kim, K.J., Sundar, S.S., & Park, E. (2011). The effect of screen-size and communication modality on

psychology of mobile device users. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI 2011)*, ACM Press, 1207-1212.

Kinley, K., Tjondronegoro, D.W., Partridge, H.L., & Edwards, S.L. (2014). Modeling users' web search behavior and their cognitive styles. *American Society for Information Science and Technology, 65*(6), 1107-1123.

Komanduri, S., Shay, R., Kelley, P., Mazurek, M., Bauer, L., Christin, N., Cranor, L., & Egelman, S. (2011). Of passwords and people: Measuring the effect of password-composition policies. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI 2011)*, ACM Press, 2595-2604.

Koved, L., & Stobert, E. (2016). *Who are you?! Adventures in authentication (WAY).* Workshop at the Symposium on Usable Privacy and Security (SOUPS 2016), USENIX Association.

Kozhevnikov, M. (2007). Cognitive styles in the context of modern psychology: Toward an integrated framework of cognitive style. *Psychological Bulletin, 133*(3), 464-481.

Ling, C., & Salvendy, G. (2009). Effect of evaluators' cognitive style on heuristic evaluation: Field dependent and field independent evaluators. *Human-Computer Studies, 67*(4), 382-393.

Ma, Y., Feng, J., Kumin, L., & Lazar, J. (2013). Investigating user behavior for authentication methods: A comparison between individuals with Down syndrome and neurotypical users. *ACM Transactions on Accessible Computing, 4*(4), Article 15, 27 pages.

Mare, S., Baker, M., & Gummeson, J. (2016). A study of authentication in daily life. In *Proceedings of the USENIX Symposium on Usable Privacy and Security (SOUPS 2016)*, USENIX Association, 189-206.

Melicher, W., Kurilova, D., Segreti, S., Kalvani, P., Shay, R., Ur, B., Bauer, L., Christin, L., Cranor, L., & Mazurek, M. (2016). Usability and security of text passwords on mobile devices. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2016)*, ACM Press, 527-539.

Messick, S. (1993). The matter of style: Manifestations of personality in cognition, learning, and teaching. Princeton, NJ: Educational Testing Service.

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

Miyake, A., Witzki, A.H., & Emerson, M.J. (2001). Field dependence-independence from a working memory perspective: A dual-task investigation of the hidden figures test. *Memory, 9*, 445-457.

Morgan, H. (1997). *Cognitive styles and classroom learning*. Westport, CT: Praeger.

Nelson, D., & Vu, K. (2010). Effectiveness of image-based mnemonic techniques for enhancing the memorability and security of user-generated passwords. *Computers in Human Behavior, 26*(4), 705-715.

Nicholson, J., Coventry, L. & Briggs, P. (2013a). Age-related performance issues for PIN and face-based authentication systems. In *Proceedings of ACM Conference on Human Factors in Computing Systems (CHI 2013)*, ACM Press, 323-332.

Nicholson, J., Coventry, L., Briggs, P. (2013b). Faces and pictures: Understanding age differences in two types of graphical authentications. *International Journal of Human-Computer Studies, 71*(10), 958-966.

Pask, G., & Scott, B.C.E. (1972). Learning strategies and individual competence. *Man-Machine Studies, 4*, 217-253.

Passfaces Corporation (2009). The Science Behind Passfaces. White paper, http://www.passfaces.com/enterprise/resources/white_papers.htm.

Peterson, E. & Deary, I. (2006). Examining wholistic-analytic style using preferences in early information processing. *Personality and Individual Differences, 41*(1), 3-14.

Peterson, E., Rayner, S., & Armstrong, S. (2009). Researching the psychology of cognitive style and learning style: Is there really a future?. *Learning and Individual Differences, 19*(4), 518-523.

Pinheiro, J. C., Bates, D. M. (2000). *Mixed-effects models in s and s-plus*. Statistics and Computing, Springer-Verlag, New York, NY.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Raptis, D., Tselios, N., Kjeldskov, J., & Skov M.B. (2013). Does size matter? Investigating the impact of mobile phone screen size on users' perceived usability, effectiveness and efficiency. In *Proceedings of the ACM Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI 2013)*, ACM Press, 127-136.

Raptis, G., Fidas, C., & Avouris, N. (2016). Do field dependence-independence differences of game players affect performance and behaviour in cultural heritage games?. In *Proceedings of the ACM Symposium on Computer-Human Interaction in Play (CHI PLAY 2016)*, ACM Press, 38-43.

Reardon, L. B., & Moore, D. M. (1988). The effect of organization strategy and cognitive styles on learning from complex instructional visuals. *Instructional Media, 15*, 353–363.

Renaud, K. (2005). Evaluating authentication mechanisms. In: Cranor, L., Garfinkel, S. (eds.), *Security and usability: Designing secure systems that people can use*, chapter 6, 103-128. O'Reilly Media.

Renaud, K., Mayer, P., Volkamer, M., & Maguire, J. (2013). Are graphical authentication mechanisms as strong as passwords?. In *Proceedings of the Federated Conference on Computer Science and Information Systems (FedCSIS 2013)*, IEEE Computer Society, 837-844.

Rezaei, A.R., & Katz, L. (2004). Evaluation of the reliability and validity of the cognitive styles analysis. *Personality and Individual Differences, 26*, 1317-1327.

Riding, R., & Cheema, I. (1991). Cognitive styles - An overview and integration. *Educational Psychology, 11(*3-4), 193-215.

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.

Rittschof, K.A. & Chambers, W.L. (2005). Instructional favoritism? Field dependence-independence does not consistently predict self perceptions or teaching preferences. *Presented at the Annual Conference of the American Psychological Society*, Los Angeles, California.

Robinson, J. S., Kitchel, T., & Garton, B. L. (2009). Using agricultural education graduates' geft scores to assess their level of job satisfaction. *Journal of Southern Agricultural Education Research, 59*, 28-43.

Schlöglhofer, R., & Sametinger, J. (2012). Secure and usable authentication on mobile devices. In *Proceedings of the ACM Conference on Advances in Mobile Computing & Multimedia (MoMM 2012)*, ACM Press, 257-262.

Shay, R., Komanduri, S., Kelley, P., Leon, P., Mazurek, M., Bauer, L., Christin, N., & Cranor, L. (2010). Encountering stronger password requirements: user attitudes and behaviors. In *Proceedings of the ACM Symposium on Usable Privacy and Security (SOUPS 2010)*, ACM Press, article 2, 20 pages.

Thorpe, J., Al-Badawi, M., MacRae, B., & Salehi-Abari, A. (2014). The presentation effect on graphical passwords. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2014)*. ACM Press, 2947-2950.

Ur, B., Segreti, S., Bauer, L., Christin, N., Cranor, L., Komanduri, S., Kurilova, D., Mazurek, M., Melicher, W., & Shay, R (2015). Measuring real-world accuracies and biases in modeling password guessability. In *Proceedings of the USENIX Security Symposium (SEC 2015)*, USENIX Association, 463-481.

van Oorschot, P.C. & Wan, T. (2009). TwoStep: An authentication method combining text and graphical passwords. In *Proceedings of the International MCETECH Conference on eTechnologies (MCETECH 2009)*, Springer-Verlag, 233-239.

Varnum, M., Grossmann, I., Kitayama, S., & Nisbett, R. (2010). The origin of cultural differences in cognition: The social orientation hypothesis. *Current Directions in Psychological Science, 19*(1) 9-13.

von Zezschwitz, E., De Luca, A., & Hussmann, H. (2013). Survival of the shortest: A retrospective analysis of influencing factors on password composition. In *Proceedings of the of the IFIP TC13 Conference on Human-Computer Interaction (INTERACT 2013)*, Springer-Verlag, 460-467.

von Zezschwitz, E., De Luca, A., & Hussmann, H. (2014). Honey, I shrunk the keys: Influences of mobile devices on password composition and authentication performance. In *Proceedings of the Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (NordiCHI 2014)*, ACM Press, 461-470.

Wang, J., & Katabi, D. (2013). Dude, where's my card?: RFID positioning that works with multipath and non-line of sight. In *Proceedings of the ACM SIGCOMM 2013 Conference*, ACM Press, 51-62.

Winter, B., & Grawunder, S. (2012). The phonetic profile of Korean formality. *Journal of Phonetics, 40*, 808-815.

Witkin, H.A., Moore, C.A., Goodenough, D.R., & Cox, P.W. (1977). Field-dependent and field-independent cognitive styles and their educational implications. *Educational Research, 47*(1), 1-64.

Witkin, H.A., Oltman, P., Raskin, E., & Karp, S. (1971). *A manual for the embedded figures test*. Palo Alto, CA: Consulting Psychologists Press.