THE EFFECTS OF COHERENCE OF THE IMAGE USED IN THE GRAPHICAL PASSWORD SCHEME IN TERMS OF USABILITY AND SECURITY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ÜLKÜ ARSLAN AYDIN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF SCIENCE
IN
DEPARTMENT OF COGNITIVE SCIENCE

AUGUST 2012

THE EFFECTS OF COHERENCE OF THE IMAGE USED IN THE GRAPHICAL PASSWORD SCHEME IN TERMS OF USABILITY AND SECURITY

Submitted by **Ülkü ARSLAN AYDIN** in partial fulfillment of the requirements for the degree of **Master of Science in the Department of Cognitive Science**,
Middle East Technical University by,

Prof. Dr. Nazife Baykal ____________________
Director, Informatics Institute

Prof. Dr. Cem Bozşahin ____________________
Head of Department, Cognitive Science

Assist. Prof. Dr. Cengiz Acartürk ____________________
Supervisor, Cognitive Science, METU

Prof. Dr. Kürşat Çağıltay ____________________
Co-Supervisor, Computer Education and Instructional Technology, METU

Examining Committee Members

Prof. Dr. Deniz Zeyrek Bozşahin ____________________
COGS, METU

Asist. Prof. Dr. Cengiz Acartürk ____________________
COGS, METU

Prof. Dr. Kürşat Çağıltay ____________________
CEIT, METU

Assoc. Prof. Dr. Tolga CAN ____________________
CENG, METU

Asist. Prof. Dr. Murat Perit Çakır ____________________
COGS, METU

Date: 28.08.2012

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name	**:**	**Ülkü Arslan Aydın**
Signature	**:**	

ABSTRACT

THE EFFECTS OF COHERENCE OF THE IMAGE USED IN THE GRAPHICAL PASSWORD SCHEME IN TERMS OF USABILITY AND SECURITY

Arslan Aydın, Ülkü
M.S., Department of Cognitive Science
Supervisor: Assist. Prof. Dr. Cengiz ACARTÜRK
Co-Supervisor: Prof. Dr. Kürşat ÇAĞILTAY

August 2012, 134 pages

There is a dilemma between security and usability, which are two fundamentally conflicting issues. From the usability perspective, authentication protocols should be easy to use and passwords generated from these protocols should be easy to remember. From the security perspective, passwords should be hard to guess and should not be written down or stored in a plain text. Instead of using text based passwords, graphical passwords have been proposed to increase both memorability and security. Biederman (1972) and Biederman, Glass, & Stacy (1973) reported that the objects in a coherent image were recognized and identified more efficiently and quickly than the objects in a jumbled image in which the jumbled image was created by dividing the coherent image into sections and changing the position of the sections without rotating them.

The study was designed to experimentally examine the differences in usability and security of the graphical password scheme by manipulating the coherence of the displayed image. Sixty-three volunteers participated in the main experiment. The participants were divided into groups according to the type of image they were presented in the password creation (either coherent-image or jumbled-image) task. Each participant created a graphical password and three days after the first session (i.e., second session) s/he tried to remember it in order to authenticate to the system. The results revealed that in the proposed graphical password scheme, using coherent image has more advantages over jumbled image in terms of usability and security.

Keywords: Usable Security, Graphical Password, Scene Memory

ÖZ

GRAFİK ŞİFRE YÖNTEMİNDE KULLANILAN RESİMDEKİ UYUMLULUĞUN (COHERENCE) KULLANILABİLİRLİK VE GÜVENLİK AÇISINDAN ETKİLERİ

Arslan Aydın, Ülkü
Yüksek Lisans, Bilişsel Bilimler Bölümü
Tez Yöneticisi: Yard. Doç. Dr. Cengiz ACARTÜRK
Ortak Tez Yöneticisi: Prof. Dr. Kürşat ÇAĞILTAY

Ağustos 2012, 134 sayfa

Temelde birbirleriyle çelişen iki konu olan güvenlik ve kullanılabilirlik arasında bir ikilem mevcuttur. Kullanılabilirlik açısından, kimlik doğrulama protokolleri kolay kullanılabilir ve ilgili protokollerden üretilen şifreler kolay hatırlanabilir olmalıdır. Güvenlik açısından, şifreler zor tahmin edilebilir olmalı ve düz metin halinde yazılarak saklanmamalıdır. Hem güvenlik hem de kullanılabilirliği arttırmak için, metin-tabanlı şifreleri kullanmak yerine grafik şifreler önerilmiştir. Biederman (1972) ve Biederman, Glass, & Stacy (1973), uyumlu resimlerdeki nesnelerin (coherent image), uyumlu resmi bölümlere ayırıp, bölümleri döndürmeden karıştırarak oluşturulan karışık resimdeki (jumbled image) nesnelerden daha verimli ve hızlı hatırlandığını ve belirlendiğini raporlamışlardır.

Bu çalışma, grafik şifrelerde görüntülenen resmin uyumluluğuna göre kullanılabilirlik ve güvenlikteki farklılıkları deneysel olarak incelemek için tasarlanmıştır. Çalışmaya 63 gönüllü katılmıştır. Katılımcılar, şifre oluşturma işlemi sırasında sunulan resmin türüne göre (uyumlu veya karışık resim) iki gruba bölünmüştür. Her bir katılımcı grafik şifre oluşturmuş ve ilk oturumdan 3 gün sonraki ikinci oturumda, sisteme giriş yapabilmek için ilgili şifreyi hatırlamaya çalışmıştır. Sonuçlar, önerilen grafik şifre yönteminde uyumlu resmi kullanmanın karışık resmi kullanmaya göre kullanılabilirlik ve güvenlik açısından daha avantajlı olduğunu göstermiştir.

Anahtar Kelimeler: Kullanılabilir Güvenlik, Grafik Şifreler, Görsel Ortam Belleği

This work is dedicated to;

Murat – a part of every page, every thought

&

My mother – source of encouragement and inspiration to me
throughout my life

ACKNOWLEDGMENTS

I want to express my sincere appreciation to my advisor Assist. Prof. Dr. Cengiz Acartürk, without him this thesis would not have been possible. I am very grateful for his encouragement, guidance, detailed comments and passion for research. I am also heartily thankful to my co-supervisor, Prof. Dr. Kürşat Çağıltay for his invaluable suggestions, friendly attitude and constructive feedback throughout the study.

I would like to thank Mehmet Ekiz, the head of Data Processing Statistics and Information department of Capital Markets Board of Turkey (CMB), for giving us permission to conduct empirical work in a dedicated room in CMB. Spec. Engin Gelibolu who is responsible from the photography workshop of Hacettepe University and Oğuz Saygılı assisted to take photos used in the experiment. Dr. Gijsbert Stoet gave valuable advice and permission to use set of stimuli developed by him for mental rotation task. Murat Aydın made very valuable contributions to the development of software applications. Özge Alaçam helped me a lot when I conducted pilot studies at METU HCI Research and Application Laboratory. Assist. Dr. Murat Perit Çakır contributed especially in the second pilot study and was always kind and supportive. Dr. Oya Can Mutan guided on statistical analyses.

My special thanks go to Dr. Bilge Say for her intellectual support and guidance and to Prof. Dr. Deniz Zeyrek Bozşahin. They both supported and encouraged me to apply for master program in Cognitive Science when I was a special student in the department.

Many thanks to all participants and my colleagues from CMB who contributed by providing their valuable time for this study. I would be remiss without mentioning my lifelong friends Aslı Bürhan and Filiz Şen for being by my side through all the good times and the bad.

Last but not least, I would also express my sincerest appreciation to my husband, my parents and brothers, Mustafa, Fatih and Emre. They have all inspired and helped me to remain focused and finished my studies.

TABLE OF CONTENTS

CHAPTER

APPENDICES

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVATIONS

AOI	Area of Interest
ATM	Automatic Teller Machines
BDAS	Background Draw a Secret
CCP	Cued Click Points
CHC	Convex Hull Click
CIA	Confidentiality, Integrity and Availability
DAS	Draw a Secret
FAT	False Accept Rate
FRR	False Reject Rate
ISO	The International Standard Organization
IT	Information Technology
LSD	Levenshtein String Distance
MIMA	Man In the Middle Attack
OTP	One Time Password
PCCP	Persuasive Cued Click Points
PIN	Personal Identification Number
QDAS	Qualitative Draw a Secret
USB	Universal Serial Bus
VIP	Visual Identification Protocol
YAGP	Yet Another Graphical Password

CHAPTER 1

INTRODUCTION

Letting authorized users in and keeping unauthorized users out are the two main principles of information security systems (henceforth, security systems). Security systems involve three phases: identification, authentication and authorization. In the identification phase, a user identifies himself usually via an identification string such as an account number. Then, in the authentication phase the user supplies proof of his identity by passwords, certificates or biometrics such as fingerprints. Finally in the authorization phase, the system specifies user's access rights, if the user is successfully authenticated (Renaud, 2005).

The authentication phase is the main concern of this thesis. Knowledge based authentication, which can be either text based or image based, is one of the frequently used authentication methods. Recent surveys show that text based authentication, also called as alphanumeric password authentication, is widely being used (Herley, Van Oorschot, & Patrick, 2009).

Compared to other techniques, text-based passwords seem to remain dominant because of obstacles in usability and security of other techniques (Brostoff, & Sasse, 2000). However, as Wiedenbeck, Waters, Birget, Brodskiy and Memon (2005) stated, there is a dilemma between security and usability, which are the two fundamentally conflicting issues. According to the usability perspective, authentication protocols should be easy to use and passwords generated from these protocols should be easy to remember. From the security perspective, passwords should be hard to guess and should not be written down or stored in a text file.[1] However, studies have shown that, to remember a password easily, users tend to choose short passwords or passwords that have some meaningful content related to their daily life (Adams, & Sasse, 1999). These tendencies result in insecure passwords. Conversely, if the password is hard to guess, in other words if it is more secure, it will also

[1] Text file is a computer file includes sequence of lines of electronic text that is human readable.

be hard to remember. In this case, users generally write them down on paper or store them in a text file (Angeli, Coventry, Johnson, & Coutts, 2003). So, Chiasson (2008) stated that "passwords are often either memorable-but-insecure or secure-but-difficult-to-remember" (p. 3). For these reasons, some alternative methods to increase both password memorability and security have been discussed for fifteen years. Instead of using text based passwords which are easily forgotten, graphical passwords have been offered as an alternative to remember the password easily (Wiedenbeck, Waters, Birget, et al., 2005). The background idea about the usage of graphical passwords instead of alphanumerical passwords is based on the classic cognitive psychology studies that have shown that images are usually "easier to remember and more secure than words" (e.g., Cranor, & Garfinkel, 2005; Kirkpatrick, 2002; Suo, Zhu, & Owen, 2005).

There are three types of possible memory-related cognitive activity required to remember a password: *recall*, *recognition* and *cued recall* (Raaijmakers, & Shiffrin 1992). The findings in the literature show that recall is usually more difficult than recognition (e.g., Kintsch, 1970). Thus, recognition-based graphical passwords may be advantageous over recall-based passwords in terms of memorability. In cued-recall systems, a cue is provided to the user by the security system. There have been numerous studies demonstrating that users tend to remember images for longer and far better than words (i.e., picture superiority effect) (e.g., Madigan, 1983; Paivio, 1971; Paivio, & Csapo, 1973). Paivio explained images' superior retention over text by means of the dual-code theory. According to the dual-code theory, verbal and imaginal coding systems, which are two separate systems, can operate in parallel independently and are interconnected with each other. When images are studied, verbal label of the image is automatically elicited. Thus, according to Paivio, images carry both a verbal code and an imaginal code in memory. It makes retrieval of images more probable than stimuli, just verbally coded.[2] Dual coding theory provides supporting evidence for the use of graphical passwords and implies that they may have advantages over alphanumerical passwords in remembering.

A click-based graphical password system is a kind of cued-recall system in which a pixel based image acts as a cue to help trigger the user's memory. During click-based graphical password creation, users select a sequence of four or five click points on the presented image. Then, to login the system, the user reselects the points on the image in the correct order by clicking on them (Chiasson, Biddle, & Van Oorschot, 2007; Chiasson, 2008). Blonder (as cited in Chiasson, Biddle, & Van Oorschot, 2007) gave the first sample of click-

[2] Debates on the Dual Coding Theory is beyond the scope of this thesis.

based graphical passwords. The proposed system allows users to pick password by clicking on the predefined items within the presented image. Wiedenbeck, Waters, Birget, et al. (2005) and Wiedenbeck, Waters and Brodskiy (2005) stated that restricting the clickable regions by predefining them may lead to security problems arising from the small password space. They proposed an alternative click-based graphical password scheme (i.e., PassPoint) that allows users to choose password by clicking on anywhere within the presented image. Users can successfully log in the PassPoint system by clicking on the exact points or by clicking on the near points with the acceptable tolerance around that point in the correct order. The studies (e.g., Dirik, Menon, & Birget, 2007; Salehi-Abari, Thorpe, & Oorschot, 2008; Chiasson, Forget, Biddle, & van Oorschot, 2009; van Oorschot, & Thorpe, 2011) investigated the usability and security of PassPoint and demonstrated that PassPoint suffers from well-known patterns (S, W, L) and hotspots resulting in security problems. Passpoint was suggested to make the system more robust against attacks by increasing the password space; on the other hand, increasing password space did not help making the system more secure. It reveals the importance of *effective password space* in representing a set of passwords that are likely to be chosen by the users. In another study, Wiedenbeck, Waters, Birget, et al. (2005) evaluates the effect of image choice on the usability and security of the click-based graphical passwords. It was reported that, passwords chosen on some images were weaker than the passwords chosen on the other. They offered that additional studies are needed to identify common criteria for the "good" image choices. The present study was conducted on the click-based graphical password scheme to investigate the effect of the properties of the provided image –in particular, coherence– on usability and security, which is a topic that has not been well-investigated to date. In the following, the concept of coherent image is introduced based on the previous work conducted in the domain of cognitive psychology.[3]

Biederman (1972) and Biederman, Glass and Stacy (1973) reported that the objects in a *coherent image* were recognized and identified more efficiently and quickly than the objects in a *jumbled image*, in which the jumbled image was created by dividing the coherent image into six sections and changing the position of the sections without rotating them. In addition, Hock, Romanski, Galie and Williams (1978) stated that there was a spatially coherent relationship between the arrangement of objects in the scene (i.e., inter-object relation) and studied the effect of inter-object relations in the recognition. They had evidence indicating

[3] Enlarging the password space is not the main concern of the present study. Users are allowed to select anywhere on the displayed image but the tolerance around the click points was 200 *200 pixels, which may be considered a relatively large value.

that the image presented with the familiar (i.e., compatible with the experience in the real world) inter-object relations facilitate object recognition rather than the image with the novel inter-object relations. Finally, Brockmole, Castelhano and Henderson (2006) conducted an experiment and demonstrated the effects of global and local context on remembering. Global context is a large number of spatially and semantically related objects, whereas local context is a subset of spatially and semantically related objects. They reported that changes in the global context have support in favor of *better retrieval cue* rather than the changes in the local context. In the light of these findings, two graphical password schemes which differ in the presented types of the images (either jumbled or coherent) have been comparatively investigated in this thesis. This comparison has been made according to the security perspective and the usability perspective which are two dilemmas of the password problem.

For this purpose, two pilot studies (with 20 and 4 participants) and one experiment (with 63 participants) were conducted. The experiment consisted of two sessions. In the first session, each participant created a graphical password and in the second session, which was performed 3 days after the first session, s/he tried to remember it in order to authenticate to the system. Moreover, in the password creation task, the participants were presented with either the coherent image or the jumbled image created by dividing the coherent image into equal-sections that includes at least one nameable object and shuffling the order of the sections. Thus, divided and shuffled image (i.e., the jumbled image) was devoid of the coherence schemata. Participants' eye movements were recorded and questionnaires were administered in the experiment. The advantages and disadvantages of coherent and jumbled image as graphical password interfaces were discussed under the framework of the dilemma between usability and security. In the rest of Chapter 1 the assumptions and the problem statement are presented. Then the significance and the purposes of the study will be discussed.

1.1 Assumptions and Problem Statement

Assumptions in the design of the experiment are the following. Firstly, in order to disrupt the scene, the method in which Biederman (1972) jumbled the coherent image to disrupt the spatial relations between objects in the natural scene image was adopted. Secondly, studies have shown that, color is one of the visual features that is responded by specialized visual receptors and involved in the early stages of vision. Therefore, usually users are more attracted to colorful objects than they are to grayscale ones (Therriault, Yaxley, & Zwaan, 2009; Dirik et al., 2007; Aglioti, Smania, Barbieri, & Corbetta, 1997). Thus, to reduce the visual salience effects, grayscale images were used in the experiment rather than the color images. Thirdly, Nelson and Kosslyn (1976) reported that adults have better recognition memory for previously labeled image than unlabeled image. Hence, in the jumbled image, there is at least one nameable object in each section. Finally, based on the Hollingworth study[4] (2006), in the login task, participants were asked to enter their passwords on the black grid on a gray background by supposing that the image containing the cells they picked their password was covered with a black blanket. It was used to understand participants' behavior in choosing their password, whether they choose their passwords by memorizing the spatial location of the sections without object identification or not. Furthermore, two further assumptions were made in the eye movement analysis. First, Just and Carpenter (1980) stated that "there is no appreciable lag between what is fixated and what is processed" (p. 331) (i.e., eye-mind hypothesis). The second is related to the fixation duration. According to Renshaw, Finlay, Tyfa and Ward (2004), users usually spend more time when they have difficulty in processing the display.

In the present study, the effect of the image-coherence in graphical password systems was investigated in terms of usability and security metrics by experimental and descriptive methods. In particular, differences between a jumbled-image-group, in which the spatial relation between the objects was reduced, and a coherent-image-group were examined. The experimental studies reported in the present study investigate the following research question:

What is the difference in the usability and/or security of the graphical password system when provided image is coherent or jumbled?

The research question leads in to the statement of following testable hypotheses;

H1: Coherence of the image used in the graphical password system affects the memorability (usability aspect).

H2: Coherence of the image used in the graphical password system affects the used strategy to create password (security aspect).

H3: Coherence of the image affects the time taken to login (usability aspect).

[4] Study was carried out across three groups. In the first group, which was called *target present preview,* subjects were shown a natural scene image involving the target object for 20 seconds. Then, in the next screen, the target object was displayed and subjects were requested to specify the location of the target object without seeing the natural scene image on the black screen.

For the aim of finding the result of the question, following analyses were performed;

- Analysis of the elapsed time to create and confirm password and time taken to login
- Eye movement analyses were conducted
    - to investigate the effect of group type in fixation count on login screen for each session.
    - to investigate the difference in fixation count on the black grid between the pass-items and non pass-items
    - to investigate the effect of group type in total fixation duration on the black grid for each session.
    - to investigate the difference in total fixation duration on the black grid between the pass-items and non pass-items
    - *Levenshtein String Distance* algorithm was used to measure the similarity between strings created during password creation and login tasks, for each group
- Patterns of user choices in graphical passwords were investigated to model similar participant choices and to reveal the effect of given image types on these choices.
- Finally, preferred strategy for choosing a password and the users' perception of the usability of the given graphical password scheme were collected by questionnaire.

1.2 Significance of the Study

End users of information technologies usually focus on goal-oriented tasks and maintaining information security, most of the time, is not considered and is not part of the major goal. Information security is maintained mostly in the background, and users may remain less knowledgeable about security concerns in information access. Therefore, users do not take the time to behave securely.

In the PassPoint scheme, users generally select password in the specified pattern like W or S shape, select the centers of the objects or select more salient objects in the images (Chiasson et al., 2009; Lashkari, Zakaria, Salleh, & Farmand, 2009; Biddle, Chiasson, & Oorschot, 2010; Dirik et al. 2007; van Oorschot, & Thorpe, 2011). Accordingly, as reported in the previous studies (e.g., Chiasson et al., 2009; Florencio, & Herley, 2007), those systems would allow automated attack by determining likely hotspots. There exist studies (e.g., Dirik et al. 2007; Salehi-Abari et al. 2008; Chiasson et al., 2009; van Oorschot, & Thorpe, 2011) that investigated the usability and security of the click-based graphical passwords. However, the studies that systematically investigate the properties (in particular, coherence) of the provided image in the proposed graphical password scheme in terms of usability and security have been lacking to date. The present study investigates the effect of coherence in the provided image on which users choose a password.

Relevant to the present study, Hollingworth (2006) presented participants with natural scene images including a set of target objects. In the first group, the target object remained at the same location with the initial scene, while in the second group it was presented at a different location. The finding supports the *spatial binding hypothesis.*[5] Participants remembered target objects more accurately when it was presented in the same location with the initial. Hollingworth also reported that spatial binding was not observed only in natural scene. Similar results were obtained when the same test was repeated for an array of individual objects (i.e. Gauthier, & Tarr, 1997; Jiang, Olson, & Chun, 2000). However, to the best of our knowledge, comparative studies on the spatial binding effect between the objects in natural scene and an array of individual objects have been scarce. This thesis examines this difference in the context of graphical passwords.

Furthermore, as mentioned above, within the domain of cognitive psychology, there exist studies reporting a better memory performance for objects in the natural scene than the objects in non-scene images. The studies investigated the scene memory were generally conducted on the functions, such as change blindness and visual search that take place in interaction with scene memory (Hollingworth, 2009; Desimone, & Duncan, 1995; Hollingworth, Richard, & Luck, 2008; Rensink, 2000; Simons, 1997). However, password notion is different than both visual search and change blindness. Users generally choose pass-items by themselves instead of the given target objects as in the study of visual search. In the login task, depending on the presented password scheme, recognizing the location of the pass-item is sufficient enough to successfully log in to the system without knowing the identity. Therefore, HCI studies are needed to clarify the difference between natural scene (i.e., coherent) and non-scene (i.e., jumbled) images used in graphical passwords. For this purpose, an interdisciplinary study was conducted by employing experimental investigations. The research question has been investigated by combining insights from cognitive psychology, computer science and human computer interaction.

The results of the study will be helpful to explain the effect of natural scene image usage in graphical passwords. Dirik et al.'s study (2007) has shown that in a graphical password

[5] Initial support for this hypotheses comes from Reises (1989) study in which participants shows better performance in recognition of location and identitiy of objects in the natural scene.

system, the presented image has an impact on both security and usability. Presented image directly affects the security of the system. The results of the study will be helpful to determine the image type for the new proposed graphical password schemes. In addition, in the experiment, during the first phase of login, all participants, regardless of their group membership, were asked to enter their password on the black-grid-screen by supposing that the image containing the cells they picked their password was covered with a black blanket. Black-grid-screen is designed to understand participants' strategy of creating password that is whether they choose their passwords by memorizing the spatial location of the sections without object identification or not.

In the experiment, eye movements of the participants were recorded. As stated in a previous study (Henderson, Weeks, & Hollingworth, 1999), eye movements give additional insight into the nature of the processing of cognitive and visual information during the complex visual cognitive task such as scene processing. Also, by analyzing the eye movement data, the difficulty that users had in processing the display information between the coherent-image group and the jumbled-image group, and the difference in total fixation duration/count between the pass-items and non pass-items will be obtained.

1.3 Organization of the Thesis

This chapter has described the problem addressed by this thesis according to the context in which it arises. The assumptions and the investigated research questions have been introduced. Finally, contributions of the study were presented.

Chapter 2 examines the authentication methods and explores the pros and cons of each method. Then security-usability dilemma is stated by describing general security and usability metrics separately and in combination.

Chapter 3 examines the graphical passwords proposed as an alternative to text-based authentication methods. Firstly, a literature review about the categories of the graphical passwords is presented. Then, graphical password schemes are evaluated in terms of usability and security. In the second part of the chapter 3, there is a literature review about scene memory and methods for disrupting scene used in the literature were discussed.

In chapter 4, two pilot studies are reported. After that, information about materials, design, procedure and measurements of the main experiment are given. Then the results of the analyses are presented.

Chapter 5 starts with the summary of the present study. Following, the results of the pilot studies and the main experiment are interpreted and discussed in terms of usability and security. Finally, possible future work topics are proposed while discussing the limitations of the present study.

CHAPTER 2

AUTHENTICATION METHODS AND USABLE SECURITY

2.1 Security Systems

Information security systems involve three phases; identification, authentication and authorization. In the identification phase, user identifies himself usually via an identification string such as an account number; then, in the authentication phase user supplies proof of his identity by passwords, certificates or biometrics such as fingerprints and, finally, in the authorization phase, the system specifies his access rights if the user is successfully authenticated (Renaud, 2005). Authentication is within the concern of this thesis. In the next section, methods of authentication are presented.

2.2 Authentication Methods

As Suo et al. (2005) stated and as shown in Figure 2-1, authentication methods could be categorized into three main groups; token based, biometric based and knowledge based.

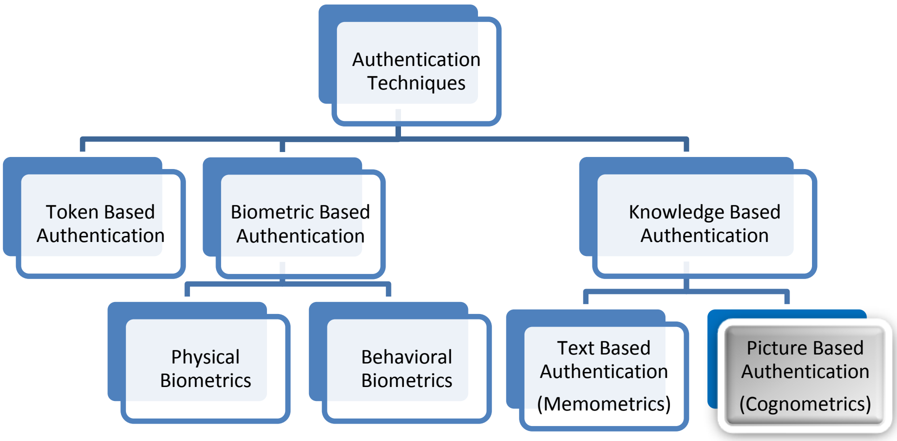

Authentication Techniques
Token Based Authentication
Biometric Based Authentication
Knowledge Based Authentication
Physical Biometrics
Behavioral Biometrics
Text Based Authentication (Memometrics)
Picture Based Authentication (Cognometrics)

Figure 2-1 Categories of authentication methods

2.2.1 Token Based Authentication

In token based authentication, a hardware or software token is assigned to an individual. Hardware tokens are generally produced as small and easy to carry devices. In many cases, tokens have a screen and display a Personal Identification Number (PIN) while others come with integrated biometric data and/or cryptographic keys. Software tokens run on computers and they provide user with a PIN changing in time. Software of the tokens consists of *one time password* (OTP) algorithms that are used to generate PIN. OTP algorithms should generate an irreversible, unpredictable and random PIN. Otherwise, unauthorized users would be able to impersonate the legitimate user (Aloul, Zahidi, & El-Hajj, 2009).

Many commercial and non-commercial solutions are available for the aim of token based authentication. In the recent state of the technology, a widely used application of hardware token is Universal Serial Bus (USB) tokens. In these systems, there exists an integrated smart card on USB devices. Users must both plug the USB tokens and enter the accurate password in order to log in to the specific system or be given access to a logical location such as a laptop.[6] That is why unauthorized users must have both a USB token and a user account password to pretend as an authorized user. Next, tokens with a display constituted the other kinds of applications. In some of them, there is a button which is used to switch them on and to produce a new PIN.[7, 8] In a third type of application, there is an integrated clock that is synchronized with the server. It generates a new PIN in a certain period of time (30-60 sec).[9] And finally, there is another one on which a PIN is displayed after user types the shared secret key and presses the button.[10] In all the applications employing tokens with a display, the same OTP algorithms and processes being used on tokens are repeated on the server to generate a PIN that user must enter. If the user account, password and the PIN match with the stored password and the generated PIN on the server, then, the user is authenticated.

[6] Goleden Security Token. (2012). Retrieved July 16, 2012, from http://www.goldkey.com/products/

[7] OTP Token. (2012). Retrieved July 16, 2012 from, http://en.digilion.com/goods/otp_token.html

[8] Strong Authentication Solution with ZyWALL OTPv2. (2012). Retrieved July 16, 2012 from, http://us.zyxel.com/Products/details.aspx?PC1IndexFlag=20040908175941&CategoryGroupNo=96C9CDE6-F2AA-4D84-9D62-311A7CCD996C

[9] UniToken Familiy | USB Authentication. (2012). Retrieved July 16, 2012 from, http://www.esecutech.com/usb-authentication/unitoken-family/unitoken-family-%7C-usb-authentication.html

[10] Security Tokens Enabling Totally Mobile Security. (2012). Retrieved July 16, 2012 from, http://www.authenex.com/site_en/Product-AKeyToken.html

Considering the fact that purchasing and maintenance of hardware tokens will be very costly for organizations; moreover due to the necessity of keeping and carrying a token, mobile phones were suggested as tokens instead (Aloul et al., 2009). A PIN can be obtained either by running the application installed on the mobile phone or by an SMS sent to the mobile phone.

The most popular samples of token based authentication techniques are bank cards, key cards and smart cards. Since providing authentication by just this kind of cards would not be enough for much of the security systems, knowledge based techniques are generally used with token based authentication (i.e., two factor authentication). For instance, Automatic Teller Machine (ATM) cards are just tokens and generally used with a PIN number (Suo et al., 2005). Previously, two factor authentications were suggested to access systems where higher level of assurance is required such as accessing online banking services (O'Gorman, 2003). However, recently, popular social networking and e-mail websites such as Facebook and Gmail of Google have also made two factor authentications available.

Yet, concerning token based authentication, there are certain disadvantages of token based authentication. The tokens can be lost, be stolen, be copied, or may not work properly. Also, they are vulnerable to the *man in the middle attack* (MIMA) and *mobile number porting attack*s. In MIMA, the attacker makes the legitimate user believe that s/he has a secure connection; whereas, in fact, the attacker controls the conversation by making an independent connection. Actually, this type of attack can also be seen in many other authentication methods. On the contrary, mobile number porting attacks have started to occur more often since the mobile phone has been adopted as a security token. The attacker calls the victim's mobile phone provider and demands to port the victim's phone number in to another one which is actually controlled by the attacker. In that way, the attacker succeeds in redirecting the verification codes to his new account.

In conclusion, the mentioned disadvantages of token-based authentication cause deficiencies both in the usability and the security of the system. This leads to exploration of new ways or methods of authentication to make systems more usable and secure.

Figure 2-2 USB-token http://www.windowsecurity.com/articles/Multifactor-authentication-Windows-Part1.html, retrieved on July 16, 2012

Figure 2-3 Tokens with a display http://en.wikipedia.org/wiki/File:Token_Verisign.JPG, retrieved on July 16, 2012

2.2.2 Biometric Based Authentication

The other type of authentication method is biometric based authentication such as fingerprints, iris scan or facial recognition. In some biometric based authentication systems, user both identifies himself (identification) and supplies proof of his identity (authentication) just via his biometric properties.

Biometric based technology has been examined in two groups (Barral, 2010).

- Behavioral biometrics: It is also called “something we do”. It depends on the behavioral traits of the individual such as hand written signature and speech patterns.
- Physiological biometrics: It is also called “something we are”. It depends on the personal physical characteristics. Finger print recognition, face recognition, hand geometry, iris/retina recognition and brainwave patterns are some of the well known physiological biometrics.

There are various biometric authentication techniques including DNA, vein pattern, ear shape, lips dynamics, face-iris recognition, fingerprint, voice and so forth (Barral, 2010; Bhattacharyya, Ranjan, A, & Choi, 2009). Each biometric technique has its own set of features, advantages and disadvantages. In addition to their features, advantages and disadvantages, applicability to the given situation and cost in that situation are all considered to determine which biometric technique can be used in a specific domain. For instance, techniques of voice recognition instead of face recognition may be used in order to identify the caller in a phone call.

There is an increased government and industry adoption of biometric based authentication due to technological advancements that contribute in quality improvements and cost reductions (Barral, 2010; Coventry, 2005; Derawi, Nickel, Bours, & Busch, 2010). As stated in the articles (i.e., Barral, 2010; Coventry, 2005), biometric based authentication is grouped according to usage area as follow*ing below:*

Physical and Logical Access Control:

In these security systems, users are given access to physical or logical location, if s/he is authorized to enter. Hand geometry, face-iris recognition and fingerprint are the most widely used techniques. These systems are generally sold as standalone and/or they're integrated to the hardware like in mobile devices and laptops. In these systems, it is not required to check and enroll biometrics of users on remote servers. These mentioned features facilitate the commercialization of biometric security systems concerning physical and logical access. That is why this is the most commonly used application of biometrics. These systems can be used at general workplaces, military areas, prisons and secure spaces where valuables, records, confidential documents or money are stored. It is also used to confirm season passes for access to hotel rooms or schools.

Transaction Security:

In the transaction security systems, users are remotely authenticated to server or network via ATMs, internet or phone. Online banking, website's membership and customer services by phone are recently used applications of the usage of biometrics in the domains concerning transaction and security. This kind of biometrics does not have widespread usage because of some of its disadvantages. Firstly, false accept rate (FAT) representing the possibility of unauthorized user access and false reject rate (FRR)representing the possibility of denying legitimate user access may have lead to problems in usage. Secondly, there is no upper limit on the number of users on the system. And lastly, since it works remotely, it requires remote access to servers or network.

Forensic identification:

Forensic identification systems are used for the aim of Immigration and Border control such as e-passports, e-visas. Fingerprint, face and iris recognition are the most widely used techniques.

Government Application (Law and Order):

In some countries driver licenses and ID cards are given with the embedded biometrics. Thus, citizens securely carry out their operations with the public sector.

Figure 2-4 Physical and Logical Access Control http://www.personelkontrolsistemi.com/index.php?p=1_29, retrieved on July 16, 2012

Figure 2-5 Transaction Security*:* http://www2.ljworld.com/photos/2005/oct/13/66221/, retrieved on July 16, 2012

Figure 2-6 Forensic identification http://crisisboom.com/2012/02/21/biometrics-are-here-to-stay/, retrieved on July 16, 2012

Figure 2-7 Government Application (Law and Order) http://fi2w.org/2009/05/28/goal-of-visiting-the-u-s-without-a-visa-still-eludes-poles/, retrieved on July 16, 2012

Biometrics is unique to each individual, so it proves identity efficiently. In the usage of biometrics, there is no physical token to carry and password to memorize. Such systems are more resistant to shoulder surfing and brute force attacks.[11] These are the important advantages of biometric systems. However biometric based security systems have some major disadvantages preventing the spread of usage. Some of the biometrics such as fingerprints can be copied. In such cases, it is not easy or possible to change copied

[11] In the brute force attack, an attacker program pretends to be a real user and tries to automatically generate exact password by trying possibilities. In the shoulder surfing attack, an attacker tries to observe user's passwords directly by watching his/her password entry (Lashkari et al., 2009; Biddle et al., 2010).

biometrics with the new ones. Also, the usage of biometrics has raised concerns relating to privacy that need to be addressed. Some of the biometric based authentication systems may not be applicable for visually, auditory of physically impaired people. At the same time, race, age, gender, occupation and environmental conditions (e.g., temperature and air pressure) affect the accuracy of some types of biometrics. In contrast to passwords, some biometrics such as fingerprints is public. It is easy to leave a trace of fingerprints. Furthermore, enrollment and access to the biometric system require much more computational resources and special hardware and software equipment (Barral, 2010; Coventry, 2005; *Matyá˘s, & Ríha, 2002*). In addition to the above mentioned issues, users of some type of the biometric systems must be trained in its use.

Heckle, Patrick and Ozok (2007) reported that, users' perspective on the usability and the security of the biometric based system varies depending on the domain in which the biometric based authentication is supplied. Even if users prefer biometric authentication, their choice is due to the usability instead of security. Because, users believe that the risks of using biometric authentication has not yet been fully understood and the biometric authentication will become the social norm will be used in most of the security system in the future.

In summary, there are still some constraints in terms of usability and security of the biometric systems (Barral, 2010; Heckle et al., 2007). Thus, until biometrics evolve into more robust authentication methods, knowledge based authentication techniques will continue to be one of the most widely used methods of authentication (De Angeli, Coventry, Johnson, & Renaud, 2005). Below, knowledge based authentication, which is the major focus of this thesis, is introduced.

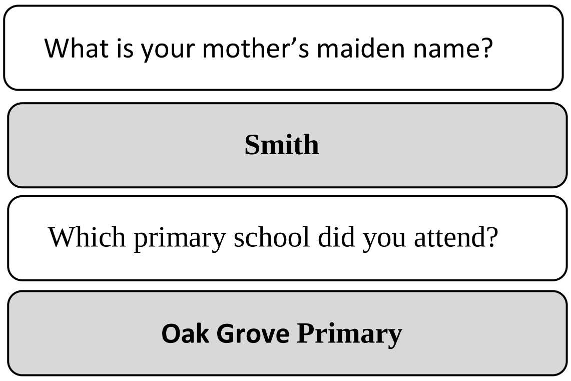
2.2.3 Knowledge Based Authentication

Knowledge based authentication can be either text based (memometrics) or image based (cognometrics). Recent surveys have shown that text based authentication, also called alphanumeric passwords, are widely being used in security systems (Herley et al., 2009).

There are two types of text based authentication, namely, random and cultural. In the random method, memometrics that contains only alphanumeric characters, is selected or generated randomly by the system. It has three subgroups: *PIN* containing only digits; *password* containing alphanumeric characters and *passphrase* a combination of words. The most commonly used memometric is the passwords which is a subgroup of random-memometric

(Herley et al., 2009). The term password is also used for all of the other authentication techniques (Renaud, 2005).

There is a general guideline for creating secure passwords. A strong password should include numbers, letters (capital and/or small) and special symbols in a total of no less than eight characters (John, 2010). In the cultural method, a password is the answer given to a series of questions in a deductive process. System allows the user to enter the shared social reality, established fact or opinion as his/her password. The user is required to remember that reality, which is also the answer of the asked question acting as a cue to recall the password. However, malicious users can easily guess the cultural-password by giving the correct answer to the question (Renaud, 2005).



What is your mother’s maiden name?

Smith

Which primary school did you attend?

Oak Grove Primary

Figure 2-8 Sample of cultural method authentication (adapted from Renaud, 2005, p. 111)

Attack types on knowledge-based authentication are divided into two groups, namely, guessing and capture. In guessing attack, the malicious user tries to guess the user password by social engineering techniques, software algorithms or predefined dictionaries. Password defined in a small password space and those based on common patterns are vulnerable to guessing attacks. Shoulder surfing where hackers observe user behaviors while s/he logging to the system, phishing, which is usually made through a fake email asking for users’ personal information, and malware a software to help hackers disrupt users computer operation generally by capturing keyboard, mouse or screen output in order to gather sensitive information are well known forms of capture attack (Biddle et al., 2010).

The underlying causes of the password attacks arising from a user weakness in choosing and managing their passwords are the following. Firstly, in order to remember them easily, users prefer passwords which are words easily found in dictionaries and/or related with their daily life, which are short in length and/or which do not include special characters. Next, instead of memorizing them, users write down their passwords in some easily accessible locations such as a notepad in their computer or a mobile device or a post-it note placed near their computer or stuck on it. Then, users generally re-use their passwords for multiple authentication systems. In this way, a malicious user, who managed to crack the password used for a low-security system easily, may access to a higher security system. Besides, users are not willing to change a password unless they are forced to do so. Finally, users generally do not hesitate to share their password (Inglesant, & Sasse, 2010).

In order to overcome these problems, password manager software tools that help users to organize and manage their password, have been proposed. Desktop, mobile, token based and web based options are available. The software generally has a local database where the encrypted password is stored. Once passwords are stored, password manager remembers and fills the password of the concerned site the next time it is visited. Chiasson, Van Oorschot and Biddle (2006) discussed the risks of password manager software and stated the reason why it is not a good alternative to the password problem as follows. They observed usability problems that resulted in insecure practices on the use of password managers. These usability problems generally arise from inaccurately perceiving the intended use of password managers. For instance, even if the users do not use the password manager system appropriately, they tend to believe that secure connection is established without any problem. The other important problems are the following; users do not believe in the security of such a method of protection and are reluctant to use password managers.

Due to the usability and security problems mentioned above, graphical passwords are proposed as an alternative method to text based passwords. Detailed information about the graphical password is presented in the next chapter.

2.3 Interim Summary: Authentication Systems

Information security systems include a three-step process namely, identification, authentication and authorization. This thesis is mainly related with the authentication phase in which the user supplies proof of his identity by passwords, certificates or biometrics such as fingerprints. There exist three main techniques for user authentication; token based, biometric based and knowledge based.

Bank cards, key cards and smart cards are the popular samples of token based authentication. Token based authentication requires user to keep and carry the token. In addition, from the organization perspective, it will be very costly to distribute and maintain the token to users.

Unlike token-based authentication, in the case of biometrics, there is no physical token to carry around and password to memorize. However, some major disadvantages prevent the spread of usage of the biometric-based authentication. In some situations, biometric system fails to detect matches (i.e., false reject rate) or makes incorrect matches (i.e., false accept rate) between the predicted and obtained data. Also, it may not be applicable for every person using the system (i.e., universality). For instance, visually, auditory or physically impaired people cannot use the system due to lack of those biometrics. In addition, higher computational resources and training requirements, performance problems and privacy concerns of users increase the necessity of evolving methods and/or finding alternatives to authenticate securely and easily (Barral, 2010; Coventry, 2005; De Angeli et al., 2005; Heckle et al., 2007; Matyá˘s et al., 2002).

Text based (memometrics) and image based (cognometrics) authentication are two types of knowledge-based authentication. Recent surveys have shown that text based authentication, also called as alphanumeric passwords, are widely being used (Herley et al., 2009). In order to increase the security, knowledge based techniques are generally used with token based authentication (i.e., two factor authentication). It was observed that text-based authentication systems are vulnerable to attacks mainly because of the limited capacity of humans. Graphical password systems, in which passwords are created on images instead of alphanumeric characters, have been proposed as an alternative authentication method to provide users with usable yet secure systems. Detailed information about the graphical password is presented in the third chapter. In the next section, security and usability issues are examined and security-usability dilemma is discussed.

2.4 Usable Security

2.4.1 Security

Usage of computers and software to gain and manage information of individuals, organizations, companies and also government has increasingly become an essential part of daily life, and this is called Information Technology (IT). IT plays a vital role in a wide range of areas including education, banking, financial transactions, shopping, communication, entertainment and socialization. It is essential that all of the IT application data be protected and secured. However, the importance given to the security increases in parallel with the increase in number and type of attacks (Steering Committee on the Usability, Security, and Privacy of Computer Systems, National Research Council, 2001).

Information security, also called Computer Security, protects data from unauthorized access, use, observation, copy, deletion, modification and disruption. Parker (1998) claimed that existing CIA (confidentiality, integrity and availability) may not be adequate to describe the security issues completely, comprehensively and statistically. Therefore he proposed a new information security framework (Parker, 1998; Parker, 2002). In this framework, there are six essential foundation elements in terms of the scenarios concerning information losses; availability, utility, integrity, authenticity, confidentiality and possession.

Loss of *availability* occurs when a user cannot access information for instance a file, the name or location that is changed even though it remains stored in the file system. In the worst case scenario, if there is not any chance of recovery and the information that is unavailable is very critical, the system may become inoperable. In the case that information requires a high degree of confidentiality, it is stored as encrypted and as a single copy. A private key is required to decrypt the encrypted information and make it readable again. If that key is lost or stolen, it can render the information useless even though the file is still reachable. Next, *integrity* is explained with the most common situation seen in computer games. If the publisher's name is removed after the content of the computer games are copied and replicated without any permission of the publisher, it will lack integrity. In the preceding scenario concerning illegal copy of the computer games, if the distributor, who made the copies of the game, changes the publisher name instead of just removing it, copies of the game will contain incorrect information and that cause the lack of authenticity. Next, *confidentiality* would be provided by preventing the disclosure or observation of information from the unauthorized user. And finally, in order to not to lose *possession* of the security system and to prevent theft, there must be a special mechanism to provide protection of the information concerning possession (Hintzbergen, Hintzbergen, Smulders, & Baars, 2010; Parker, 2002). Making regular backup, maintaining system availability, following innovations in technology are some of the most common strategies to overcome security issues caused by hardware or software problems or by malicious users. Those six elements proposed by Parker are important in two ways. First, testing and designing the security systems are facilitated by the use of common concepts. Second, those six elements give the opportunity to compare the security of different systems.

There are a number of software and hardware solutions available that aim to address the growing need for security. Some of the most common recent solutions are anti-virus and anti-spyware software, network traffic monitoring tools, firewalls and cryptographic algorithms. Lampson (2004) stated that as the internet has become a substantial part of our daily lives, it results in new types of security attacks. Also, he proposed that due to the lack of the metrics, it is not possible to measure the costs of getting security in terms of money and time. Therefore, companies and individuals have no encouragement to make their systems more secure. There are laws that govern the security in real world and that punish those who violate the security. However, since it is more difficult to find out the criminals (i.e., lack of accountability) who violate computer security, penalties cannot be applied to crimes. He proposed two isolated machines called green and red to maintain accountability while giving the people the right to roam freely over the internet. The green machine requires accountability and the red one that does not. On the other hand, Norman (2009) claimed that even the system is designed to be as secure as possible, after a time it becomes increasingly insecure, because new types of attacks emerge constantly. Therefore, design and development of completely secure systems is probably impossible.

2.4.2 Usability

Usefulness is the question of whether the system can be used to achieve the intended goal. Grudin (1992) divided usefulness in two parts, namely, utility and usability. Utility is the issue of whether the system satisfied the requirements specified before. On the other hand, usability explains how it fulfills the specified requirements (Nielsen, 1994a). Miller (1971) described usability as ease-of-use.

In the process of acceptability, usability is concerned together with practical issues such as cost, compatibility and reliability. In fact, usability is necessary at every stage of the system. Nielsen (1994a) proposed that usability consists of the following features; *learnability*, *efficiency*, *memorability*, *errors* and *satisfaction*. A system must be easy to learn (i.e., *learnability*) to quickly begin using the system. Once users learn how to use the system, it must be efficient enough (i.e., *efficiency*) to provide a high level of productivity. The use of the system must be easy to remember to allow s/he to use the system without re-learning everything from scratch (i.e., *memorability),* after s/he does not use the system for a period. In order not to make user angry with the system, it should have a low error rate (i.e., *errors)*. Lastly, satisfaction is a measure of how pleasant it is for the user to use the system.

The International Standard Organization (ISO) (1998) described usability as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use "(p. 2). Effectiveness is the capability of the system to achieve a desired goal. Efficiency represents the required time and effort to achieve the intended goal. Satisfaction indicates the opinion of the user with respect to gain satisfaction.

Partala (2009) summarized the evaluation methods of usability in three categories; *usability testing*, *usability inspection* and *inquiry*. In usability testing, a user-centered approach is adopted. The usability of the product is evaluated by testing it on end-users. Usability tests are one of the important activities in the HCI studies. Usability inspection is the other usability evaluation method in which system is tested before it is released. Thus, it will be easier to apply changes, before implementing the system. Two common examples of usability inspection are heuristic methods (Nielsen, 1994b) and cognitive walkthrough (Wharton, Rieman, Lewis, & Polson, 1994). Lastly, in the inquiry method, evaluator asks questions to the user or observes the user while s/he carries out the tasks on the system (e.g. surveys, questionnaires).

Usability is a kind of non-functional requirement (Ben-Asher, Meyer, Möller, & Englert, 2009). It cannot be measured directly, but it can be quantified by measuring the features of usability in terms of operationalized metrics (Nielsen, & Levy, 2003). Table 2-1 shows the usability factor and the measurable metrics of each factor (Kainda, Fléchais, & Roscoe, 2010).

Table 2-1 Usability factors and measurable metrics (from Kainda et al., 2010, p. *5)*

Usability	
Factor	**Measurable Metrics**
Effectiveness	Task success
Satisfaction	Satisfaction
Accuracy	Success rates
Efficiency	Completion times, number of clicks/buttons pressed
Memorability	Recall
Knowledge/skill	Task success, errors, mental models

2.4.3 Security-Usability Dilemma

"When security procedures impede users' end goals, users bypass security" (Chiasson, et al., 2006, p.1)

As mentioned in the above security section, many hardware and software solutions are available in order to provide security (Ben-Asher et al., 2009). All of these proposed solutions have performed well only when they are used effectively by users. That is why users are a fundamental factor to provide security. Using the existing tools properly is just as important as developing new technologies for the aim of providing security. Although the company ensures that they get the most secure products, if they do not have enough employees with the skills to use those products effectively, they will face with security problems. In short, users are the weakest link in any security system (Schneier, 2004; Sasse, Brostoff, & Weirich, 2001).

End-users' interaction with the systems, in certain ways, may result in information security issues. First of all, users may not be aware of the importance of data as an asset for the enterprise organization. Next, users may not realize that their goods are under risk and may not consider protecting them until they are under attack. Then, they may not believe that their behavior can put organizational assets at risk (Sasse, & Flechai, 2005). Furthermore, they may have no knowledge of, or interest in, what they are expected or required to do for the aim of providing security (Tognazzini, 2005). On the other hand, from the perspective of system designers, security systems are generally not usually designed by taking into account user factors in spite of the fact that most of the threats on information security focus on user factor-related aspects which increase chances of a successful attack (Simon, 2002).

With the increasing prevalence of information technology in many aspects of our lives, most people have to deal with a lot of knowledge based authentication items in order to access secure systems (Sasse et al., 2005). According to Sasse et al. (2001), the average number of passwords is estimated to be around six per person. In addition to this, a survey study conducted by Dunphy, Nicholson and Olivier (2008) with people working in the field of information technology showed that users could have up to 16 passwords. Also, the number of passwords per user increase with the improvements in the effectiveness of information system security. In terms of the usability of passwords as an authentication method, memorability of passwords is an important parameter, as well as perceived ease of user interaction with the system. On the contrary, from the security perspective, passwords should not be easy to guess by others and not be easy to detect with malicious password detection tools (Luca, Denzel, & Hussmann, 2009; Wiedenbeck, Waters, Birget, et al., 2005). Therefore, it may be difficult to maintain usability while increasing security. This indicates the dilemma between security and usability. In other words, strong passwords tend to be hard to remember whereas memorable passwords are usually weak (Chiasson, 2008).

In the above paragraph, only the impact of the user factor on the security system is evaluated. Actually, there is a bi-directional relation between the security system and the user . The way that the users interact with the system affects the security of the system while the design of the system influences the usage of the system by the user (Adams et al., 1999). For instance, in order to provide security, security systems require long password and/or regular change which direct the user to choose their password from the first letters of alphabet, sequential number series or the personal information such as first name, last name, birth date, name of children, phone number etc. On the other hand, because of their limited cognitive capacity, individuals put the system in jeopardy by writing their passwords down instead of memorizing them or using the same password for different accounts (Adams et al., 1999; Dhamija, & Perrig, 2000; De Angeli et al., 2005).

In 2005, in a big company, a network-based password detection tool was run to assess security of users' passwords. The results showed that 80% of the user passwords were identified in 30 seconds. The major outcome was that the users preferred easy-to-remember passwords (Suo et al., 2005). This finding supports the dilemma between usability and security, in other words ensuring usable security is difficult to achieve.

The problems of providing usable security might be due to handling of security and usability from different points of view. Security studies generally focus on preventing malicious users from identifying passwords, which also adversely affects the compromise between the authorized user and the system. However, usability studies generally focus on letting authorized users access to the system easily, which also facilitates the unauthorized access to the systems (Kainda et. al., 2010). Security and usability are considered non-functional requirements. Therefore, security and usability are handled separately by different experts after completing design, development and testing of the system. Usability experts deal with usability issues whereas experts in security aim at meeting the security requirements (Ben-Asher et al., 2009). Riley (2006) argued that the problem of providing usable security arises from the belief that the problem could be solved by designing theoretically secure systems instead of effectively secure ones.

There are many considerations about how to provide usable security. Almost every one of us uses the internet regularly or is connected to the networks. Johnston, Eloff and Labuschagne (2003) believes providing security by using complex tools and bad interfaces is difficult; because most of the users are not a security expert. Thus according to him, improvements in the design of the interfaces are crucial for providing usable security.

Gaw and Felten (2006) and Riley (2006) claimed that it would not be sufficient to ensure that users choose more secure passwords just by keeping the user informed about the security issues and about what needs to be done for security. Birge (2009) claimed that security and usability are complementary rather than conflicting goals. Sytems should be designed to prevent users making mistakes.

Also, password managers proposed to provide usable security. The study (i.e., Gaw et al., 2006) showed that password management systems increase a tendency to choose poor passwords or re-use the same passwords across multiple account/services. They suggested that password management systems should encourage users to choose strong password instead of just helping them.

Usable security has become a major topic of debate in information technologies since the past several decades. There are interdisciplinary studies in usable security, including computer security, human computer interaction, psychology and even neuroscience (Steering Committee on the Usability, Security, and Privacy of Computer Systems, National Research Council, 2001). The first studies in this field were published from the late 1980s and 1990s. *Users are not enemy* (Adams et al., 1999) and *User-centered security* (Zurko, & Simon, 1996) are two of them. Following, in 2003, HCI-SEC, which is the first workshop about usable security, has held. With the increasing interest in the field, Symposium on usable Privacy, which was first held in 2005 and is still ongoing, began to be realized. Moreover, a few universities, like Carnegie Mellon and University College London, has programs and research groups in the field of usable security (Theofanos, & Pfleeger, 2011).

2.5 Interim Summary: Usable Security

Miller (1971) described usability as ease-of-use. According to Nielsen (1994a) usability consists of the following features; *learnability*, *efficiency*, *memorability*, *errors* and *satisfaction*. System should be easy to learn (i.e., *learnability*), be efficient enough (i.e., *efficiency*) to provide high productivity, be easy to remember to allow users the opportunity to be able to use the system without relearning how to use it (after not using the system for a while) (i.e., *memorability)*, have a low error rate (i.e., *errors)* and make the user feel good when a need or desire is fulfilled (i.e., satisfaction).

The information security focuses on protecting data from unauthorized access, use, observation, copying, deleting, modification and disruption. In the information security framework proposed by Parker (1998), there are six essential foundation elements in terms of the scenarios concerning information losses; availability, utility, integrity, authenticity, confidentiality and possession. There are a number of software and hardware solutions available that address the growing need for security. Anti-virus and anti-spyware software, network traffic monitoring tools, firewalls and cryptographic algorithms are among the technologies available for protecting the system and provide security to the systems. Lampson (2004) reported that the increasing presence of the internet in everyday lives has contributed to the emergence of new types of security attacks. Also, he argued that companies and individuals have no encouragement to make their systems more secure because of the lack of the metrics to measure the costs of providing security in terms of money and time.

Companies are faced with security problems unless security products are used effectively, even if they have the most secure ones. In short, users comprise the weakest link in security systems (Sasse et al., 2001; Schneier, 2004). The design of the system influence users' behaviours and behaviours of users can potentially affect the security of the system (Adams et al., 1999). Chiasson (2008) stated that "passwords are often either memorable- but-insecure or secure-but-difficult-to-remember" (p. 3). Ensuring usable security is difficult to achieve because of the conflict between usability and security.

As mentioned in this chapter, providing both usability and security is a difficult task because of the dilemma between usability and security. In this thesis, security-usability dilemma is investigated under the graphical password domain. For this purpose, existing and proposed graphical password schemes are discussed in terms of usability and security. Furthermore, how users recall or recognize graphical passwords has been investigated within the scope of cognitive science and related disciplines. In the next chapter, firstly, types of graphical passwords are examined in terms of usability and security. Then, scene memory is explained and methods for scene-disruption is given.

CHAPTER 3

GRAPHICAL PASSWORDS AND SCENE MEMORY

Graphical password is one of the knowledge-based authentication methods proposed as an alternative to text-based passwords. In graphical password systems, in order to authenticate to the system, users click on the provided image rather than enter an alphanumeric character (Wiedenbeck, Waters, & Brodskiy, 2005). It is one of the human computer interaction topics such that usability is important as well as security. This part begins with the categories of the graphical passwords then continues with the usability and security aspect of them. In the security and usability aspects sections, the attributes that are used in evaluating security and usability of passwords are summarized. Scene memory is one of the influencing factors of both usability and security. Therefore, at the end of this chapter, there is a literature review about scene memory that summarizes relevant research studies.

3.1 Graphical Passwords

Graphical password schemes have been proposed as an alternative to text-based passwords. Graphical passwords are supposed to be more secure, because they would have a larger password space than alphanumeric passwords. For instance, number of possible passwords (i.e., password space) of the graphical password scheme in which a password consists of 5 click-points on the provided image size of 1024*752 with a tolerance value of 20*20 pixels will be $2.6*10^{16}$. On the other hand, password space of the alphanumeric password scheme in which a password consists of 8 elements over a 64 character alphabet will be $64^{8}=2.8*10^{14}$ (Wiedenbeck, Waters, Birget, et al., 2005). Also, they are supposed to be more usable, because they are easily memorable than alphanumeric passwords. Graphical passwords have another advantage over alphanumeric passwords in that it is more difficult to write down, store and share them. Graphical password schemes having been worked in the past two decades propose a solution, both usable and secure, to the password problem (Dunphy et al., 2008).

Studies have demonstrated that users tend to remember images for longer and far better than words (i.e., picture superiority effect) (e.g., Madigan, 1983; Paivio, 1971; Paivio, & Csapo, 1973). Paivio stated that learned information is encoded in two ways; verbal and visual. In addition to this, learned information is stored in visual and verbal memories which are two functionally separated but linked memory traces. Moreover, both visual and verbal information can be used while recalling or retrieving information for subsequent use (Sternberg, 2003). For instance, when a word is recalled, it would retrieve just as verbal, as visual or both simultaneously. Storing the desired learning material in two separate functional locations instead of just one will increase the likelihood of remembering. These findings support the memorability of the graphical passwords.

There are several areas of use for graphical passwords. Some of these include login to websites, workstation applications and automatic teller machines (ATM). Also, after the increase in the usage of mobile touch screen devices, it is especially used to unlock the screens of these devices. In addition, there exist browser extensions that give the possibility to login to the system by graphical passwords without changing the interfaces of the applications (Bicakci, Yuceel, Erdeniz, & Gurbaslar, 2009).

It is useful to describe the terms and concepts frequently encountered during the rest of the study. Since several types of graphical passwords have been proposed, in order to distinguish them, *scheme* is used for the name of the proposed system. The other one is *theoretical password space* which is the total number of the possible combinations of all available elements in the given scheme. On the other hand, *effective password space* is a subset of the theoretical password space. Studies that examined the user selected passwords yields that users have a tendency to choose predictable and similar passwords. That is why in most schemes passwords do not have equal probability to be chosen by users. Effective password space represents a set of passwords that are likely to be chosen by the users.

3.2 Categorization of Graphical Passwords

Graphical passwords can be categorized into three groups based on the type of the cognitive activities materialized during password remembrance; recall, recognition and cued-recall (Biddle et al., 2010; Chiasson, 2008; Lashkari et al., 2009; Suo et al., 2005). Recalling is the most difficult task (Craik, & McDowd, 1987).

3.2.1 Recall-Based Algorithms

In the recall based algorithms, the user and the system share a secret. Users must recall this secret in order to get access to the system. In fact, this kind of graphical password is the most similar to text based passwords. Because, as in the case of text passwords, in the recall based algorithms, users must remember the password and reproduce it (Chiasson, 2008). Most popular recall-based algorithm schemes are presented below.

Draw a Secret (DAS): In their paper, Jermyn, Mayer, Monrose, Reiter and Rubin (1999) proposed DAS system as an alternative to the text based passwords. In this method a 2D grid is presented to the user. S/he uses a mouse to draw password. Then, in order to login to the system, user must repeat the same path. Some studies have lessened the effectiveness of DAS. For instance, Goldberg, Hagman and Sazawal (2002) showed that memorability of text passwords is better than DAS.

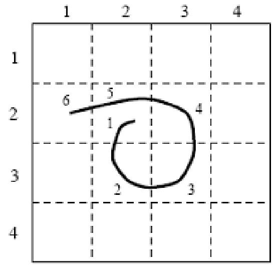


Figure 3-1 DAS
(taken from Jermyn et al., 1999)

Passdoodle: Goldberg and his colleagues proposed a Passdoodle algorithm (Goldberg et al., 2002). Passdoodle is similar to the DAS. This method allows users to create their own handwritten design or text, usually onto a touch sensitive screen. Some studies have shown the negative aspects of Passdoodle. For instance, Varenhorst (2004) reported that users could remember their doodle as accurately as alphanumeric passwords, but they made mistakes in recalling.

Figure 3-2 Passdoodle
(taken from Varenhorst, 2004)

Pass-Go: It's designed was inspired by a Chinese board game called GO. Similar to DAS, it is a grid based scheme. Unlike DAS, users click or touch on the intersections of the cells instead of the cell itself in order to choose their password. Thus, diagonal lines would also be chosen during password creation which results in an increase in the theoretical password space. Also, in later versions of Pass-Go, it allows users to select the color of the pen to increase the password space (Tao, & Adams, 2008).Pass-Go is vulnerable to exploit-based attacks (Chiasson et al., 2009).

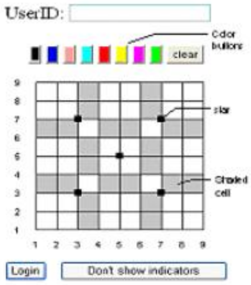


Figure 3-3 Pass-Go (taken from Tao, & Adams, 2008)

Other recall based schemes

Studies have shown that users have a tendency to select symmetrical passwords in DAS scheme. Thus, effective password space of DAS scheme is much less than the theoretical password space. Dunphy and Yan (2007) proposed *Background Draw a Secret* (BDAS) to overcome this problem. They supposed that background images would reduce the possibility of selection of symmetrical passwords. There have not been many studies to evaluate the BDAS in terms of usability and security.

Another scheme that proposed to increase the theoretical password space of the DAS scheme is *Yet Another Graphical Password* (YAGP). Direction of pen strokes are calculated by the Levenshtein String Distance (LSD) algorithms and matched with the stored password. Unlike the DAS scheme, YAGP makes possible to use finer grids (Gao, Guo, Chen, Wang, & Liu, 2008).

Qualitative Draw a Secret (QDAS) is proposed by Lin, Dunphy, Olivier and Yan (2007) with the aim of preventing shoulder surfing attacks take place while user drawing their passwords. In QDAS, haptic input device measuring pen pressure used for that purpose. However, the study (i.e., Orozco, Malek, Eid, & El Saddik, 2006) revealed that the usage of the haptic device makes a very little difference in preventing shoulder surfing attacks.

GrIDsure (2009), which is a commercial product, is invented by Jonathan Craymer and Stephen Howes. In this scheme, users generate their one time passwords without carrying hardware tokens. Users select the cells in the 5*5 grid during password creation. Then to login to the system, users are required to enter the corresponding digits inside the cells selected during the password creation phase, by using their keyboard. Since numbers placed in the grid cells are selected randomly, each time users enter a different password. Therefore, GrIDsure is resistant to the shoulder surfing attack.

As well as security, meeting the usability concerns is the other issue to improve the existing schemes. Por, Lim, Su and Kianoush (2008) suggested using background in order to increase the memorability of the Pass-Go scheme.

Recall based graphical password schemes are also available in mobile touch screen devices to unlock the screen. For this purpose, Patternlock (2010) running on the BlackBerry devices and another scheme similar with the Pass-Go running on android devices are available.

3.2.2 Recognition-Based Algorithms

There have been several theories investigating the differences between recall and recognition memory. According to one perspective, recall and recognition are unique tasks, while others proposed that they are similar but differ only in their retrieval difficulty. Graphical passwords, when based on recognition based algorithms, have two advantages over recall algorithms. Firstly, there is a general acceptance of the idea that recall is difficult than recognition (Kintsch, 1970). Secondly, users have the ability to recognize images even though they see them very briefly (Nelson, & Kosslyn, 1976). Thus, recognition based graphical passwords could be preferable than recall based passwords. Three of the most popular recall-based graphical password schemes are presented below.

Déjà vu: Dhamija and Perrig (2000) proposed the Déjà vu scheme. In Déjà vu users select images from a larger sample to create their passwords. In order to authenticate to the system s/he will be required to recognize images which s/he select during the password creation process, from a set of images including decoy images. Chiasson (2008) reported that in *Déjà vu* theoretical password space is less than text based passwords.

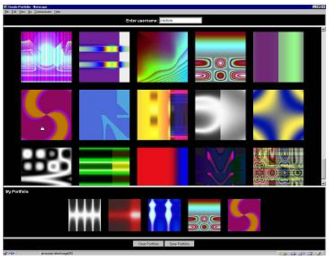


Figure 3-4 Déjà vu (taken from Dhamija, & Perrig, 2000)

Pass Faces: Valentine (1999) reported that users could remember their PassFaces password with the login success rate between 72% to 100%. In PassFaces, system presents face images to the user. In the password creation phase, it is asked user to select a set of faces from a larger set. Login process consists of 4 rounds. In each round, a set of face images including the face that is part of his/her password are presented. User has to pick a preselected face. This process will be repeated 4 times until all pieces of the password are collected. Chiasson (2008) reported that PassFacess is significantly less vulnerable to shoulder-surfing than text based passwords.

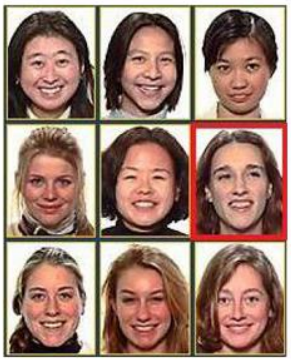


Figure 3-5 Pass Faces
(taken from Valentine, 1999)

Weinshall (2006) proposed a graphical password scheme as a solution to the network security problems. It is resistant to spyware and shoulder surfing attacks. System generates and assigns a random password to the user account for the aim of increasing effective password space. In the learning phase, super set of images including both the pass-items and decoy images is presented, and user is asked to select the pass-items among the given set.

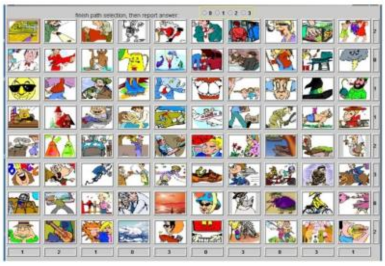


Figure 3-6 Query panel
(taken from Weinshall, 2006)

In the login phase, user is asked a multiple choice question. To give the answer to the question, user is required to draw a mental path on the presented super set of images by starting from the left corner and until reaching the right most or the bottom end image. While drawing this path, if user comes across with the pass-item, s/he should direct the path to the one below image. Otherwise, s/he should follow the next right image. Finally, s/he is asked to choose the number written near the right most or bottom end image reached by the mental path. This process is repeated specific number of times. System calculates the probability of the user's answers could occur by chance. If it met the pre-defined threshold value, the user will be authenticated.

Other recognition based schemes

Akula and Devisetty (2004) proposed an alternative scheme to the Déjà vu. It aims to reduce the need for memory and also provide security. To this end, in this proposed scheme, passwords are run through a SHA-1, which is a hash function, before being stored. There are some deficiencies in the proposed scheme. As such, it is prone to shoulder surfing attacks. It still needs future improvements.

Studies have shown that users' choices in the Pass Faces scheme are affected by gender and race. Davis, Monrose and Reiter (2004) had decided that their primary goal was to measure effects of user choice in the security and their secondary goal was memorability. They came up with the story scheme. In the proposed scheme, users select their password by choosing a story from a single set of images, each derived from a distinct category such as faces, sports, animals, cars or everyday objects.

Convex Hull Click (CHC), a multiple round authentication scheme, aims to prevent shoulder surfing attack. In this scheme, system presents a portfolio consisting of several hundred icons, and the user is asked to select a pre-defined number of icons from the given portfolio. In each round, a large number of icons including at least three pass-items are presented from the entire portfolio. The user is asked to find the pass-items (icons) chosen and to mentally create a convex hull formed by these icons and then click anywhere inside this convex hull instead of clicking directly on pass-items. Users are not expected to click on the icons. Thus, it is very difficult for the attacker to guess the password just by observing the user. That is why it is secure against the shoulder surfing attack. Since it takes long time to login, it is not very usable even users found it fun (Wiedenbeck, Waters, Sobrado, & Birget, 2006).

In the other scheme proposed to prevent the shoulder surfing attack, the system presents a large portfolio of images to the user. Users have to select specific number of images from the presented images in order to set their passwords. In the portfolio, there also exists a set of pertubated images which have small changes in the appearance of the original image from scene to scene. A different code is assigned for each image (Man, Hong, & Mathews, 2003). To log in to the system, the user has to find the pass-items from the presented portfolio and generate a code including both the pre-assigned one and the location information relative to the screen. There are multiple rounds of challenges in this scheme. In this scheme, it takes longer to train user to memorize the pre-assigned code and to learn how to use the system. Hong, Man, Hawes and Mathews (2004) suggested a modified version of this scheme, in which users allowed to assign the codes to the images. Nevertheless, both versions make the users memorize code numbers and it is the drawback of text-based passwords.

In the *Visual Identification Protocol* (VIP) scheme, the system presents a large portfolio of images to the user. Users asked to select a specific number of images from the presented portfolio to set their password. In order to authenticate to the system, users are required to recognize images selected during the password creation process, from a set of images that includes decoy images (De Angeli, Coventry, Johnson, & Renaud, 2010). Pering, Sundar, Light and Want's (2003) Photographic scheme is very similar to VIP. The only difference is that it lets users to upload their own images and allows choosing their password from the uploaded images. To log in to the system, user has to find the pass-items from the presented portfolio including decoy sets consist of the images uploaded by the users.

Jansen (2003) described a general-purpose scheme to authenticate users on mobile devices. Thumbnail images, which are either individually recognizable or a part of a larger mosaic formed by their combinations, were placed in a 5 by 6 grid. System allows users to choose the theme on which they select their password. To log in to the system, users have to click the pass-items in the correct sequence. Takada and Koike (2003) proposed a very similar scheme to authenticate users on mobile devices. The only difference is that it lets user to upload their own images and allows choosing their password from the uploaded images.

Hayashi, Dhamija, Christin and Perrig (2008) claimed that, the usability of the proposed graphical passwords for authenticating on the mobile devices is depended on the screen size of the concerning device. They proposed a graphical password scheme which is usable regardless of the screen size. Like some other schemes, it lets users to upload their own images. The difference is that uploaded images become distorted and in the login process, system displays the distorted images. The idea behind this approach is that legitimate users would recognize the image even it is distorted and it would be difficult for illegitimate users to crack the password by guessing it or observing the user. Thus, this scheme is secure against shoulder surfing and social engineering attacks.

3.2.3 Cued-Recall Algorithms

A cued-recall system incorporates recall and recognition. Tulving and Pearlstone (1966) stated that some of the items in human memory could be inaccessible for retrieval but still can be available. Their results indicate that users are able to access previously inaccessible items with the help of cues. In cued-recall systems, the user firstly recognizes the cue, and then s/he will recall the shared secret with the help of this cue (Chiasson, 2008). Three of the most popular recall-based algorithm schemes are presented below.

Blonder: This approach is designed by Greg E. Blonder (1996). This is the first click-based graphical password method. In this method, system presents a pre-defined image which has predefined clickable points. In order to create his/her password, user selects one or more predefined points. S/he must reselect those positions in the same order for the aim of getting access to the system. In this method, since the system allows users to click only on the predefined areas, the password space is relatively small (Lashkari et al., 2009).

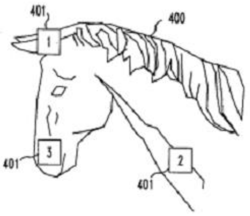


Figure 3-7 Blonder (taken from Blonder, 1996)

PassPoint: As it is mentioned above, Blonder method has some constraints about the password space. Wiedenbeck, Waters, Birget, et al. (2005) proposed PassPoint which is an extension of Blonder's method as an alternative to Blonder. Passpoint overcomes the drawback of Blonder by enlarging the password space. On the other hand, learning and login time of PassPoints is longer than alphanumeric passwords.

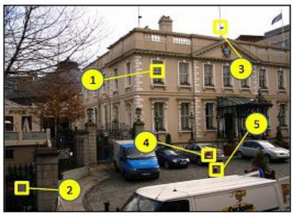


Figure 3-8 Passpoint (taken from Wiedenbeck, Waters, Birget, et al., 2005)

Passlogix v-Go: This scheme is proposed by the Passlogix Inc. Security company which has been acquired by Oracle in 2010[12]. It is very similar to the Blonder scheme. System allows users to click on a set of predefined regions. Unlike the Blonder scheme, in Passlogix v-Go, there is a nameable individual item in each predefined region, and its background could be room images such as kitchen, bedroom or bathroom. Users choose their password in the chronological order of situations by clicking on or dragging items, and to authenticate to the system, they have to repeat the actions in the same order. Theoretical password space is small (Lashkari et al., 2009; Thorpe, & Oorschot, 2004).

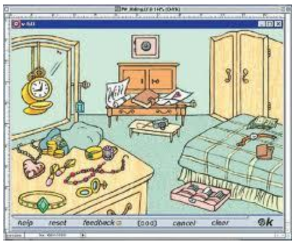


Figure 3-9 Passlogix v-Go (taken from Lashkari et al., 2009, 2005)

[12] Oracle buys Passlogix. (2010). Retrieved July 18, 2012, from http://www.oracle.com/us/corporate/press/176326

Other cued-recall schemes

Studies have shown that Passpoint scheme suffer from patterns and hotspot problem(e.g., Dirik et al., 2007; Salehi-Abari et al., 2008; Chiasson, et al. 2009; van Oorschot & Thorpe, 2011). Chiasson, Van Oorschot and Biddle (2007) proposed Cued Click Points (CCP) scheme in order to overcome this problem. In CCP, users choose their passwords from five different background images by clicking one region on each image, instead of just one background image. Thus, there is a one-to-one relation between pass-item and background. Persuasive Cued Click Points (PCCP) is the other cued recall scheme which is very similar to CCP. In PCCP, system offers suggestions for pass-items to the user for the aim of increasing effective password space. A small rectangular area is highlighted while the other parts of the image are disabled for clicking. User can either chose it by clicking on the highlighted area or skip it by clicking the shuffle button. In order to authenticate to the system, it is required to select the same areas in the same order (Chiasson, Forget, Biddle, & Van Oorschot, 2008).

Alsulaiman and Saddik (2006) proposed 3D scheme. Like in the 2D schemes, environment serves as a cue. It allows user to interact with the system not only by clicking the items or areas in the environment but also by writing or drawing on the surface or by moving objects inside the virtual environment. In order to login to the system, user has to repeat same type and sequence of actions in the environment. Theoretical password space of this scheme is relatively large because of the variety of actions that can be done. But there have not been enough studies on usability and security of the proposed scheme because it is still just a prototype.

Inkblot is proposed by Microsoft, and it is not a strict graphical password scheme (Chiasson, 2008; Stubblefield, & Simon, 2004). In this scheme, system displays inkblots generated by computer and ask user to enter the first and the last letter of the word or phrase describing the inkblot on the screen. Inkblot serves as a cue to remember the entered letters. In order to authenticate to the system, users have to enter the letters corresponding to the displayed inkblot.

Viskey SFR is proposed especially for mobile devices by the German company SFR. It is very similar to the Blonder and Passlogix v-Go. In order to login to the system, users have to click the pass- items in the same order. It accepts all input within a certain tolerance that would result in incorrect authentications (Abdullah, M., Abdullah, A., Ithnin, & Mammi, 2008; Lashkari et al., 2009).

3.3 Evaluation of Graphical Passwords in terms of usability and security

3.3.1 Usability Aspect

In this section, usability reports of the existing graphical password studies will be examined. Since there is not any accepted standard about the usability criteria of passwords yet, it is possible to see different metrics related to the type of the compared passwords schemes and the authors who are making this comparison. It makes difficult to evaluate and compare the different usability studies' results together.

Lashkari et al. (2009) offers a complete table of usability attributes based on the models developed by the International Organization for Standardization (ISO). They combine the ISO 9241, which is a standard about ergonomics, ISO 9126, which is about software quality, and ISO 13407, which provides standards for the life cycles of user centered approaches. Effectiveness, efficiency and satisfaction are determined as the major attributes of the usability aspects. Effectiveness is the level of the achievement of the goals using the system. Efficiency refers to the resources necessary to achieve goals. Lastly, satisfaction refers to the feelings of users while using the system. De Angeli et al. (2005) measured effectiveness by calculating the number of correct entries and examining the error types of wrong entries. They proposed error types as the followings, erroneous in selection, erroneous in the sequence and composite error which is composed of the first two types. They measured efficiency by the entry time and measured user satisfaction by means of the questionnaires and pre-tests.

Biddle et al. (2010) examined usability aspect under the target users, tasks and domain headings. In the target users, they addressed the importance of the characteristics of the intended users while designing or selecting graphical passwords schemes. In the first place, in fact all of the well known graphical passwords schemes has an implicit assumption that indented users of given scheme have a good vision and good motor skills (i.e., can click on an image). The proposed graphical password schemes should address the limitations of the given scheme or suggest alternative methods for these kinds of disabilities.

They also specified that there are different tasks in the system besides the login process. Password initialization is the starting process of the authentication systems. In the previous graphical password studies, passwords are either assigned to user or selected by the user. Assigning password to user by the system as randomly can increase security but suffer from memorability. On the other hand, allowing users to select their passwords increase memorability since password would include personal meaning. But this time it would suffer from security because of its high predictable level. Login is the second and the most common step of the system. Thus, simplicity and the velocity are more important for this task than the others. Lastly, resetting password and changing password are often required in practice when the user forget his/her password.

In the domain, they noticed the difference in the domain should cause difference in the level of the security and usability concerns. For instance, the risk level of online banking should be high which requires focusing on the security more than usability. Therefore, the target environment should be clearly explained for the given password scheme.

Table 3-1 Usability attributes from ISO standards (Lashkari et al., 2009) with the perspective of Biddle et al. (2010) in usability headings

Usability features	**Attributes**	**Attributes especially for graphical user authentication**	**Abbreviation**
Effectiveness	**(*)**Reliability& Accuracy	Reliability& Accuracy	R&A
Efficiency	**(*)**The utilization in real world	Applicable	Applicable
Satisfaction	**(*)**Easy to use	Use the input device easily (would be mouse or eye tracker)	
	(*)Easy to create	Select simple way to create the password	
	(*)Easy to memorize	Meaningful	Meaningful
		User assign image	Assignable image
		Freedom of choice	
	(*)Easy to execute	Selects simple steps of related task	Simple steps
	(*)Good view	Select good interface	Nice interface
	(*)Easy to understand	Simple training session	Training simply
	(*)Pleasant	Pleasant image	Pleasant image

(*) evaluate for each tasks: password initialization, login, reset and change of the password for intended users in the predefined domain

Usability of the existing graphical passwords based on their categories is as followings. In effectiveness, especially in the recall based graphical passwords, determining the error tolerance acceptable for each data is a key task. Error tolerance is a user defined integer and higher error tolerance can cause the outcome to be inaccurate and unreliable in quality. In efficiency, especially in the recognition based graphical passwords, it requires the selection of different images in each round. Moreover, the requirement of high speed network and storing tens of thousands of images in database are not applicable to the real world. Lastly, authenticating to the system and creating a password especially in some of the recognition based graphical passwords requiring a number of rounds of verification take a long time. In terms of satisfaction, they are not much usable (Bicakci, Atalay, Yuceel, Gurbaslar, & Erdeniz, 2009; Thorpe, & Oorschot, 2004).

Results of the main experiment are discussed for the following usability metrics given in the Table 3-1; reliability & accuracy, utilization in real world, easy to use, easy to create, easy to memorize and pleasant. In the next section, studies about security aspect of graphical passwords are summarized.

3.3.2 Security Aspect

This section discusses standard attacks to graphical password systems. Attack types are summarized based on the previous studies (Biddle, Chiasson, & Oorschot, 2010; Lashkari et al., 2009). Attacks can be carried out by malicious users or programs.

Brute Force: In the brute force, an attacker program pretend to be a real user and tries to automatically generate exact mouse motions to login the system by generating the password correctly. It is believed that graphical passwords are more protected to brute force attacks than the text-based passwords (Lashkari et al., 2009; Biddle et al., 2010).

Dictionary Attacks: In the dictionary attacks, an attacker program tries to find out the password by the trial and the error method. In this case, tested passwords are chosen from the pre-defined dictionary. Studies about DAS scheme have shown that there is a user's tendency to choose mirror symmetric passwords. Thus, the effective password space of DAS is significantly smaller than the theoretical DAS password space. So, DAS is vulnerable to dictionary attacks (Thorpe, & Oorschot, 2004). On the other hand, in recognition based and cued recall methods (e.g. Déjà vu, Blonder), passwords are created by clicking on items with a mouse. Hence, it is relatively difficult to make a dictionary attacks on this types of graphical passwords. However, there are exceptions. For instance, in PassFaces better looking faces and same racial group as themselves and in PassPoint bright colors, nameable objects and center of images are more likely be to chosen (Dirik et al., 2007; Thorpe, & Oorschot, 2004). Thus, they are vulnerable to dictionary attacks.

In summary, it has been stated that graphical passwords are more protected to dictionary attacks than text-based passwords. Because automated dictionary attack for the graphical passwords should be much more complex than text-based passwords (Biddle et al., 2010; Lashkari et al., 2009).

Guessing: In the guessing attack, an attacker tries to guess the password either online or offline. Guessing attack is based on the assumption that users tends to select their password related to their personal information like their pet's name, family name, birth date and so on. The studies shown that in graphical passwords users generally select their password in the specified pattern like W or S shape (Biddle et al., 2010; Chiasson et al., 2009; Lashkari et al., 2009). Like text-based passwords, graphical passwords are not secure against guessing attack (Davis et al., 2004; Nali, & Thorpe, 2004).

Spyware: In the spyware attack, an attacker tries to collect password information of users by using a key logger or key listeners which are installed to computer without the user knowledge. Since graphical passwords are generally selected by clicking on the specified point of the image and this point could change related to the resolution of screen, it is believed that gathering graphical password by spyware is hard than gathering passwords on text based schemes (Biddle et al., 2010; Lashkari et al., 2009).

Shoulder Surfing: In the shoulder surfing attack, an attacker tries to observe user's passwords directly. Crowded places are the most appropriate environment for attackers. Because it's very easy to be near the user's screen and watch his/her password entry. Both texts based and graphical passwords are vulnerable to this attack (Biddle et al., 2010; Lashkari et al., 2009). It is possible to make some of the recognition based and cued-recall methods resistant to this attack (Suo et al., 2005).

Social Engineering: In the social engineering attack, an attacker tries to collect information about an organization or a computer system by using human relations. Attacker can connect different sources and pieces together until s/he reaches the information what s/he needs (Biddle et al., 2010; Lashkari et al., 2009). Since it is difficult to share graphical passwords than text-based passwords, graphical passwords are more secure against this attack (Suo et al., 2005).

Smudge Attack: The touch screen is widely used in smart phones. In this attack, attacker tries to determine the user's password by finger smudges on the screen (Aviv, Gibson, Mossop, Blaze, & Smith, 2010).

In summary, graphical passwords, in particular recognition based and cued-recall methods, are more resistant to social engineering, spyware, brute force and dictionary attacks than text-based passwords (Biddle et al., 2010; Dirik et al., 2007; Lashkari et al., 2009;Suo et al., 2005; Thorpe, & Oorschot, 2004). Further studies are needed in this area, to investigate security attacks on graphical passwords.

3.4 Interim Summary: Graphical Passwords

Graphical password is proposed as an alternative to text-based passwords in order to provide both usability and security. Users click on the provided image to authenticate the system (Wiedenbeck, Waters, & Brodskiy, 2005).

From the usability perspective, graphical passwords are easily memorable than text-based passwords. From the security perspective, graphical passwords should have a larger password space than text-based passwords and storing, sharing and writing them down is more difficult than the text-based passwords. Therefore, graphical passwords propose a solution to the password problem that is both usable and secure (Dunphy et al., 2008). There are three types of graphical passwords, namely, recall, recognition and cued-recall (Chiasson, 2008; Biddle et al., 2010). Studies (e.g. Craik, & McDowd, 1987; Kintsch, 1970) have shown that recognition is difficult than the two others. Thus, recognition based and cued-recall graphical password schemes are more memorable than the recall based password schemes. Cued-recall systems cooperate recall and recognition. Tulving and Pearlstone (1966) reported that humans are able to access previously inaccessible items by the help of cues. In cued-recall systems, user firstly recognizes the cue, and then s/he will recall the shared secret by the help of this cue (Chiasson, 2008).

Effectiveness of the recall based graphical passwords is affected by the error tolerance, a user defined integer, of the proposed scheme. The requirement of storing tens of thousands of images in database, high speed network and number of rounds of verification affects efficiency and satisfaction of the recognition-based graphical passwords (Bicakci, Atalay et al., 2009; Biddle et al., 2010).

The recognition-based and cued-recall graphical passwords are more resistant to social engineering, spyware, brute force and dictionary attacks than text-based passwords (Biddle et al., 2010; Dirik et al., 2007; Lashkari et al., 2009; Suo et al., 2005; Thorpe, & Oorschot, 2004). In the graphical password domain, inherent features of graphical passwords (e.g. type of the presented image) lead interaction between security and usability. The presented image in the graphical password scheme is related both usable security and human cognition and has been rarely investigated until now. In the next section, human cognition aspect (i.e., scene memory) of this issue is examined through relevant studies in the area.

3.5 Scene Memory

Studies have shown that visual context, which consists of objects that are semantically related with each other; facilitates detecting and identifying the objects in the context (Biederman, 1972; Biederman, Mezzanotte, & Rabinowitz, 1982). Hollingworth (2005) stated that identifying objects and knowing their locations are essential to enable intelligent interactions with the environment. By vision, humans can perceive visual features of the object such as shape and color, also detect positions of objects relative to each other and relative to our positions, and then classify them according to these features. However, perceiving objects individually is not enough to achieve intelligent behavior within an environment, but one also has to perceive the relationships between objects. Many studies (e.g. Biederman, 1981; Biederman, Mezzanotte, & Rabinowitz, 1982; Biederman, Rabinowitz, Glass, & Stacy, 1974; Green, & Hummel, 2006) have shown the significance of the object interaction for visual cognition. If this interaction represents meaningful and functional context, it also facilitates object identification (Green, & Hummel, 2004).

Biederman (1981) proposed five classes of features that represents the relation between an object and its surroundings, which are used to distinguish between *well-formed scene* and *array of unrelated objects*, namely; *support, interposition, probability, positions* and *size*. Support is the necessity for objects to stand on a ground. For instance a table cannot fly in the air. As for Interposition, if the object is not transparent, behind of it cannot be seen. Probability is the likelihood of presence of an object in the scene. For instance, a basinet is not expected to be in the service station. As for position, objects should be placed in a specific location in the scene. For instance, in a petrol station, a gas pump is not expected to locate on a car. Lastly, there is a certain relationship in the sizes of objects relative to each other. Biederman claimed that, unlike probability, position and size, support and interpositions are the physical constraints that can be perceived without identifying the concerned object. He reported that semantic relation is perceived efficiently and quickly as well as physical relations.

There exist other studies (e.g., Minsky, 1974; Palmer, 1977) that discuss features of a well-formed scene by the differences between local and global information (see Antes, Penland & Metzger, 1981, for a review). Physical properties of a chair such as color and line angle are local information. On the other hand, information concerning the entire chair is global but information of a dining room is more global than the information of a single chair. Dining room supports contextual information.

Henderson (1992) stated that some of the studies in the scene processing were intended to clarify the *scene identification*. When one looks out the window, s/he can notice that it is a city scene but not a kitchen or a bathroom. How the class of the observed scene is determined? Friedman (1979) argued that scene identification is provided by identifying one or more characteristic objects in the scene. De Graef, Christiaens and d'Ydewalle (1990) claimed that scene identification is provided by the spatial relations of the characteristic objects. On the other hand, Biederman (1981) and Biederman et al. (1973) suggested that scene identification occurs before identifying the objects in the scene.

Schema hypothesis is a predominant view that explains the relation between scene identification and the identification of objects in the concerned scene. According to this view, firstly, the individual recognizes the class of the observed scene and then, prototypical information of the scene is accessed from memory. After that, according to the prototypical information, an expectation for the objects that are more likely to be present in the concerning scene is created.

Many studies in this area have been conducted because of the importance of scene memory in our daily lives. There exist studies that examine whether the functionally specialized brain regions are recruited for scene memory. Kim and Biederman (2011) investigated in which phase of the visual cognition, interaction between objects is occurred and which parts of the brain is activated during this process. They showed that a lateral occipital complex (LOC) activated during the identification of individual objects (as cited in Vuilleumier, Henson, Driver, & Dolan, 2002) is also activated during the occurrence of interaction between objects. We have a robust visual memory which preserves information about the low level visual features and spatial location of objects in the natural scene (Hollingworth, 2009). In addition, cognitive psychology studies to investigate the scene memory are generally conducted on the functions, such as change blindness and visual search that take place in interaction with scene memory (Desimone, & Duncan, 1995; Hollingworth, 2009; Hollingworth et al., 2008; Rensink, 2000; Simons, & Levin, 1997). In the following section, main functions that interact with scene memory are given.

3.5.1 Visual Search

Hollingworth (2009) stated that the visual search is one of the important functions that interact with scene memory. Land and Hayhoe (2001) pointed out the significance of visual search in our lives through the example of cooking. Before starting cooking, one must prepare all ingredients, which requires visual search. After that, ingredients are mixed in certain proportions which also require visual search in its inner processes. Similar to the prior example, many of our daily life tasks require visual search.

Some studies have proposed that low level visual features, such as color and dimensions, and spatial locations of objects facilitate the visual search (Desimone, & Duncan, 1995; Hollingworth, 2009; Hollingworth et al., 2008). For instance, while one tries to find the book among the other books, if s/he knows that the book is red covered, since s/he redirects her/his attention to the red ones, s/he finds the book wanted much quicker. Similarly, if s/he knows the spatial location of it instead the visual feature, since s/he redirects her/his attention to the specific regions, s/he again finds the book wanted much quicker.

Over the last 15 years, there have been studies that investigated the effects of spatial information on visual search. Hollingworth (2005) investigated the spatial memory which is a memory for the location of the individual objects in the natural scene. He carried out the study across three groups. In the first group, namely *target present preview,* subjects were shown a natural scene image involving the target object for 20 seconds. Then, in the next screen, the target object was displayed and subjects were requested to click on the black grid to specify the location of the target object without seeing the natural scene image. In the second group, namely *target absent preview*, subjects were shown a natural scene image that did not involve the target object for 20 seconds. Then, in the next screen, the target object was displayed and subjects were requested to click the probable location of the target object if the scene had contained it. In the last group, namely *no preview*, without showing any image, the target object was displayed, and subjects were wanted to imagine a natural scene image which would involve that target object, and click the most likely location where the target object would place. The obtained results revealed that the spatial memory of the objects in the natural scene plays an important role in the visual search task.

Furthermore, Hollingworth (2009) summarized the other studies as followings. Firstly, it has been observed that if an individual wants to find out a target object in a natural scene, s/he focuses on the most likely location of the target object, even though s/he has never seen that scene before (Henderson et al., 1999; Neider, & Zelinsky, 2006). For instance, one looks at the walls of the room firstly, when s/he is prompted to find the painting in the room. Secondly, Castelhano and Henderson (2007) shown that if one is allowed to take a look at the scene in a very short period of time, it makes visual search on that scene easier. Lastly, repeated search on the same scene increases the level of success in finding the target object as long as the positions of objects in the scene remain constant (Henderson et al., 1999; Neider, & Zelinsky, 2006).

Lastly, in their study (Foulsham, Alan, & Kingstone, 2011) created 4 version of images, namely, coarse-normal (divided into 4*4 grid), coarse scrambled (divided into 4*4 grid and jumbled), fine normal (divided into 8*16 grid), fine scrambled (divided into 8*16 grid), fine scrambled (divided into 8*16 grid and jumbled) and defined one region, stayed in fixed position, of each grid as a target presented 3 second on the center of a gray screen. After that one of four versions of the images was displayed and participants were asked to response whether shown image included target by keyboard entry. If the participant accepted the presence of target, s/he wanted to click the position of the target on a black grid on the white background. Results have shown that, on the fine-scrambled version, less accurate responses were given, it took more time to focus on the target, more fixation count and shorter mean fixation durations were allocated.

3.5.2 Change blindness

Rensink (2001) mentioned the phases of vision and explained the natural scene memory via change blindness. According to Rensink, vision takes place in three levels; low, middle and high. Geometrical shape, color and orientation of the object are all gathered in the low level. The high level involves issues of physical and semantic knowledge concerning object types representing the meaning of the object. The middle level involves everything that the two other do not cover. In fact, the middle level aims to close the gap between the visual features obtained in the low level and the knowledge of the object’s identity used to perceive our environment. There are different opinions on the definition and the functions of the middle level. Rensink assumed that the natural scene features are perceived and interpreted in the middle level. Thus, understanding the middle level helps us to understand scene perception. Rensink benefitted from change blindness, which is the inability to encounter the changes between the scenes being viewed in a short exposure, to examine scene perception. In change blindness, while low level visual features of the object are perceived, person cannot access the knowledge of meaning obtained in the high level. In other words, the gap between the low and high level cannot be closed which is supposed as the middle level function.

Sensitivity to changes in scenes depends on the focus of attention. As Pashler (1988) stated, humans are able to focus their attention only a few objects at the same time because of their limited capacity. Thus, most of the objects in the scene could be located outside of the observer’s attention and that cause the change blindness. Rensink (2001) illustrated change blindness and scene perception by the *coherence* and *visual representation theories*. According to coherence theory, firstly, low level volatile *proto objects* are formed and they exist as long as the scene continues to be viewed. Then, coherence is provided by focusing attention to a set of proto objects. In fact, it functions as a link and is created through the feedback comes from the nexus[13] and proto objects. Rensink claimed that change blindness occurs, because enough coherence to detect the changes is not provided.

Both visual search and change blindness are two main functions that interact with scene memory. Studies in visual search and change blindness have revealed that coherence in the scene is important from the point of view of visual scene memory. Scene disruption and contextual cueing are two other issues examining the role of coherence in the scene memory. Scene disruption and contextual cueing are discussed below.

3.5.3 Scene disruption

Biederman (1972) and Biederman et al. (1973) investigated one of the first methods of disrupting scene. They jumbled the scene to disrupt the spatial relations between objects in the natural scene image. The coherent image was divided into six sections by one vertical and two horizontal lines. Then, jumbled image was created by changing the order of sections without having to rotate them. They have reported that objects in the coherent image was recognized and identified more efficiently and quickly than the jumbled one. In addition, in some of his studies, in order to disrupt the scene, one or more features such as position and size, of the well-formed scene defined in the previous study (i.e., Biederman, 1981) were eliminated. Furthermore, Boyce, Pollatsek and Rayner (1989) made studies by comparing the original scene image with the modified one in which the target image was presented without the background.

Yokosawa and Mitsumatsu (2003) investigated the effect of the scene disruption in change blindness. He provided scene disruption by splitting an image into 6 or 24 sections and then shuffling section or by making a modification in one of the object in the image. These modifications can occur either on color, location or presence of the object. Thus, semantic and global context of the scene was changed. Unlike Biederman et al. (1973), in these studies, Yokosawa and Mitsumatsu investigated change blindness in which focusing on the color and the position of the objects, instead semantic information, had been sufficient to detect change. He reported that scene context does not facilitate the change detection. In

[13] Nexus create the basis of high level decisions by collecting informations from proto-objects (Rensink, 2001).

addition, splitting an image into 24 sections and making some parts invisible by replacing particular sections with the black ones is the other method used by Yokosawa in order to disrupt scene. In the present study, methods of Biederman (1972) and Biederman et al. (1973) was adopted for disrupting scene.

3.5.4 Contextual Cueing

Natural scenes contain more than one object related with each other in different class of features. When one looks at a scene, even s/he does not remember every detail; s/he will have expectations of the presence of possible objects in the scene and their spatial layouts. Spatial layout of the objects in the natural scene image facilitates objects recognition in the scene (Bar, & Ullman, 1996). Object recognition and visual search are different tasks. When an individual searches for a target object in the scene, s/he already knows the model of the object. Knowing only the location of the target object could be sufficient. On the other hand, in object recognition, unlike visual search, an individual has to determine the model of object among the others, and in some situations knowing the location of the object is also necessary (Bar & Ullman, 1996).

Biederman (1972) and Biederman et al. (1973) conducted one of the first studies to investigate how spatial context aids visual recognition by breaking the spatial relations between the objects in the context. They have reported that objects in the coherent image was recognized and identified more efficiently and quickly than the jumbled one.

In their study, Hollingworth (2006) presented participants with natural scene images including target object. In the first group, target object remain at the same location with the initial scene while in the second group, it is presented in different location. The finding supports the *spatial binding hypothesis*[14]. Participants remember target objects more accurately when it is presented in the same location with the initial.

Furthermore Hock et al. (1978) reported that inter-objects relations facilitate the object recognition. In another study (i.e., Brockmole, Castelhano, & Henderson, 2006), the effects of the global and local context were examined in three groups. In the first group, changes were made to the global context while local context remain constant. In the second group, changes were made to the local context while global context remained constant. And in the last group, both local and global context remained constant in each repetition. According to

[14] Initial support for this hypotheses comes from Reises study in which participants shows better performance in recognition of location and identitiy of objects in the natural scene (Rieser, 1989).

the results, although contextual cue observed in all three groups, the strongest influence is when the global context remains constant. Brockmole and Henderson (2006) attributed this result to the following factors. At first, objects in the scene were arranged in "spatially licensed manner" (Henderson, & Hollingworth, 1999). Thus, objects have specific physical and semantic constraints depending on the identity of the scene in which it is placed. So, as mentioned in the previous studies (e.g. Henderson, & Ferreira, 2004), even identities of local objects are determined, global context ensures expectations about the spatial layout of objects. Individual focuses attention on the target object faster when it is presented with the global information than it is presented only with local information. Secondly, changing the identity and/or spatial layout of objects has more impact on the identity of the scene rather than the changes in the local context. Thirdly, studies (e.g. Epstein, Harris, Stanley, & Kanwisher, 1999) have shown that there are several functionally specialized brain regions processing global scene information.

3.6 Interim Summary: Scene Memory

One has to perceive the relationships between objects in addition to perceiving them individually, in order to achieve intelligent behavior within an environment. For visual cognition, object interaction is a significant process (Biederman, 1981; Biederman et al., 1982; Biederman et al., 1974; Green, & Hummel, 2006). According to Biederman (1982), *support, interposition, probability, positions* and *size* represents the relation between an object and its surroundings. *Well-formed scene* and *array of unrelated objects* can be distinguished by those features. Scene memory is generally investigated through tests in visual search and change blindness which are made easier by the existence of scene memory (Desimone, & Duncan, 1995; Hollingworth, 2009; Hollingworth et al., 2008; Rensink, 2000; Simons, & Levin, 1997).

There exist studies that examine whether the functionally specialized brain regions are recruited for scene memory (Kim, & Biederman, 2011). In addition, psychological studies to investigate the scene memory are generally conducted on the functions, such as change blindness and visual search (Desimone, & Duncan, 1995; Hollingworth, 2009; Hollingworth et al., 2008; Rensink, 2000; Simons, & Levin, 1997).

There exist some similarities and differences between the mental operation used for authentication and visual search and/or change blindness. In order to authenticate, users look at the image to find the target items but they generally choose pass-items themselves instead of the given target objects and they are generally authenticated just by recognizing the location of the pass-item rather than knowing the identity of each item. Therefore, studies are needed to clarify the effect of schema in the graphical passwords.

In this thesis, scene disruption was provided by disrupting the spatial relations between objects (i.e., jumbled image) in the natural scene image. It has been adopted from the previous studies (e.g. Biederman, 1972 and Biederman et al., 1973). Studies have shown that (e.g. Henderson, & Ferreira, 2004), global context ensures expectations about the spatial layout of objects and one focuses attention on the target object faster when it is presented with the global information than it is presented only with local information. This is one of the main motivations that guide research questions.

CHAPTER 4

EXPERIMENTAL INVESTIGATION

In the methodology chapter, details about two pilot studies and one follow-up study which were conducted before the main experiment will be reported. Then, participants, experiment environment and materials used in the present study, dependent and independent variables and design issues will be explained. After that, important design issues of the experiment, procedures and the data collection will be presented respectively. Finally, analysis procedure will be examined.

4.1 Pilot Studies

Two pilot studies and one follow-up studies were performed before the main experiment. The results of each study were evaluated and methodological problems were detected based on these results. Then, the necessary changes were applied to the methodological issues of the follow-up study in order to reduce the effects of the detected problems.

4.1.1 First Pilot Study

Research Questions

As stated in the previous chapter, according to the *picture superiority effect* pictures have superiority in the encoding, storing and retrieval processes over alpha numeric representations. Furthermore, according to schema theory, memory performance is affected by schema based expectations (Spiro, 1977). These two are the assumptions of the first pilot study. In particular, the following research question was investigated.

- What is the difference in the usability and/or security of the graphical password system when provided image is coherent or separated?

Methodology

Participants

Twenty participants (10 females and 10 males) took part in the first pilot study, all of whom were either professionals or university students. The age range was 23-41 (M=28.1, SE=0.84). All participants had computer skills. Furthermore, since they used web services such as e-mail or on-line banking, by using their own passwords regularly, all of them were aware of the security systems and password notion. Age and education distributions of participants are given in the Appendix A.2.

Materials and design

A desktop software application was developed for the first pilot study. It was written in Microsoft C# programming language and build with .Net Framework 3.5. Application was installed on the participants' computers. So, they conducted the study on their own computers. The participants were presented either a bedroom image or separated items in the bedroom (Figure 4-1). Pre-defined 12 items in the bedroom image were cropped out by using Photoshop and these 12 objects were arranged in two rows side by side (Figure 4-2), to create a separated image. Since the number of clickable items was equal in both groups; password space for each group was also equal.

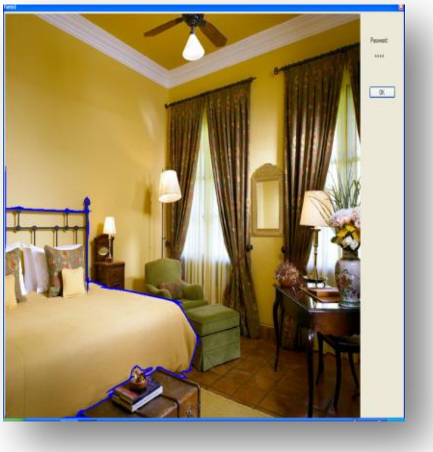


Figure 4-1 Bedroom image (1024 *754)

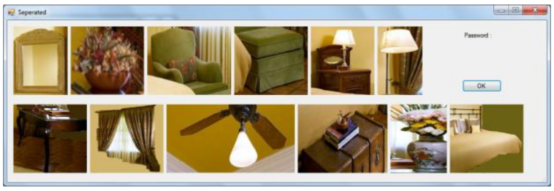


Figure 4-2 Separated image ([86-246]*131)

By means of the application and questionnaires, demographic and task specific information were collected from the participants. The main independent variable in the study was the image type (coherent or separated) as a between-subject variable. The dependent variables were the followings: successful login (yes or no), error type (sequence error, wrong choice error, both of the sequence and wrong choice errors, error free) and the number of login attempts.

Procedure

The study consisted of two sessions. Ten participants (5 female and 5 male) created a password with the coherent image and the remaining ten participants used separated images to create a password.

The first session was divided into three phases: *practice, password generation* and *retention*, as shown in Table 4-1. The practice phase was used to explain the task and to familiarize the participants with the user interface of the developed application. At the beginning of the practice phase, participants were informed about the task. Then the experimenter introduced the interface through creation, confirmation and login processes by using a practice password. Finally, at the end of the practice phase, participants filled out the demographic questionnaire shown in the Appendix A.1.1 In the password generation phase, participants were asked to choose realistic passwords that they could remember but that would be difficult for others to guess. Participants picked their passwords by clicking four click-points on the displayed image. A selected item was framed with a blue line, for giving feedback to the participant. Finally, in the retention phase, the last phase of the first session, participants were asked to log in to the system ten times by using their usernames and passwords.[15] After then, the participants answered a printed questionnaire (Appendix A.1.2) aimed at gathering their perception on the usability and security of the provided system.

The second session was conducted after one week and it was completed by all the 20 participants. The participants tried to log in to the system with their user names and passwords. They were allowed to try three times until they recalled their password. Otherwise, they were evaluated as failed to access. At the end of the session, participants were demanded to answer a short printed questionnaire in order to gather participants' strategies used for remembering.

[15] The selection of the trial number is based on the previous studies (e.g. De Angeli et al., 2005).

Table 4-1 Methodology of the first pilot study

Phase	Number Of Trials	Steps
I. Practice	One trial These data were not used in the analysis.	Explain System Create Login Questionnaire (Appendix A.1.1)
II. Password Generation	One trial These data were recorded as user passwords.	Create
III Retention	Ten trials These data were used in the analyses.	Login Questionnaire (Appendix A.1.2)
IV. One week Retention	Maximum three trials These data were used in the analyses.	Login Questionnaire (Appendix A.1.3)

Results

As shown in Table 4-2, all the participants (i.e., ten of ten participants) who performed the study with the separated image remembered their passwords whereas six of the ten participants who performed the study with the coherent image were able to remember their passwords.

Since expected frequencies of participants in some groups were smaller than 5, Fisher's exact test was performed. There was not any effect of gender (P=1, FET) and education level (P=1, FET) on the accuracy of remembering. Also, there was not any effect of the group type (separated vs. coherent images) on the elapsed time (P=.87, FET) during the login process in the second session.

Table 4-2 Password retention rates of the participants

		Remember	**Not remember**	**Remember after trials**
Group type	Separated Image	10	0	1 of 10
	Coherent Image	6	4	0 of 6

The error of failed login attempts was classified into three types;[16]

- Type 1: Sequence Error: It arises when pass-items were retrieved but the arrangement of items is not appropriate.[17]
- Type 2: Wrong Choice Error: It occurs when the password does not include the retrieved item.
- Type 3: Both of the Sequence and Wrong Choice Errors

30% of the participants made at least one error. 50% of these errors belong to type 1 and the remaining errors belong to type 3. These results demonstrate that second type error was not observed alone.

The participants' answers to the questionnaires showed that 85% (17 participants) created the graphical password easily. The remaining, 15% (3 participants) reported that it was not easy. These three participants (2 performed the study with separated images) remembered their password. Nine participants reported that usability of an alphanumerical password is better than presented graphical schemes. On the other hand, 55% (11 participants) noted that they would prefer graphical passwords rather than alphanumerical ones.

According to the answers of Appendix A.1.3 questions, all of the participants, who performed the study with the coherent image and remembered their password correctly (60 %), were aware of each items of their password. However, 50% (5 participants) performed the study with the separated images did not aware of every item. They said that they remember their password by the help of the location cues.

[16] This classification is made based on the features presented in De Angeli et al. (2005).

[17] Pass-item refers to each pieces that make up the password.

Discussion

The results obtained in the first pilot study demonstrate that coherence in images may not contribute to memorability of graphical passwords. 50% of the separated-image-participants, who remembered the password correctly, also remembered the pass-items themselves. The remaining separated-image-participants, who remembered the password correctly, remembered only the locations of the pass-items such as four corners or the first four items but not the identity (i.e., the content) of the pass-items. This type of recognition is known as *spatial coding* which can be described as recognizing the physical position of an image rather than the image itself.[18]

As stated in the previous usable security chapter, the aim of the security systems is making a compromise between security and memorability. When a user pick his/her password just based on the location of the items, it will be easy for attackers to guess the user's password which cause a security problem.

In order to reduce or eliminate this kind of coding, in the second session, images could be located randomly. Thus, participants who could not remember the identity of each pass-item also were not able to make a successful login. This suggestion is based on the assumption of the effect of spatial coding that is in disfavor of password security. In the first pilot study, there are design issues that introduced confounding factors in fully interpreting the results. Firstly, in the separated-image group, although the same items with the coherent image were used, the sizes of them were different from their coherent-image counterparts and from each other. Secondly, the coherent image was rectangular in shape, whereas separated images were displayed as a non-rectangular shape. And thirdly, the participants in the separated group were allowed to click on any object they want. On the other hand, the participants in the coherent group were allowed to click on the pre-defined items in the image. In the main experiment, these issues were eliminated by making the necessary changes in the design.

4.1.2 Second Pilot Study

The goal of the second pilot study was to explore design issues that are relevant to data collection in different modalities (e.g., eye tracking) and to resolve technical issues before the main experiment. According to Biederman (1972) and Biederman et al. (1973**)** studies

[18] Studies have shown that the spatial position of the stimuli among local features in an image is encoded (e.g. Nicoletti, & Umilfft, 1989; Stoffer, 1991).

coherent image was recognized and identified more efficiently and quickly than the jumbled image.[19] This is the basic assumption of this pilot study.

In the second pilot study, in addition to the obtained data from application and questionnaires, eye movement data (fixation count, total fixation duration and area of interests) were obtained by eye-tracker. The research question investigated depends on these data are as followings:

- What is the difference in eye movement data when the presented image is coherent vs. jumbled?

Methodology

Participants

Four participants attended in the study. Two participants were assigned to the jumbled image group and the other two to the coherent image group. There were one male and one female in each group. Participants completed their sessions individually. They did not have any vision problems. Also, in order to ensure that participants were aware of the password notion, at the beginning of the pilot study, they were informed about graphical passwords and the general aim of the study. Between-subjects design was used for this pilot study.

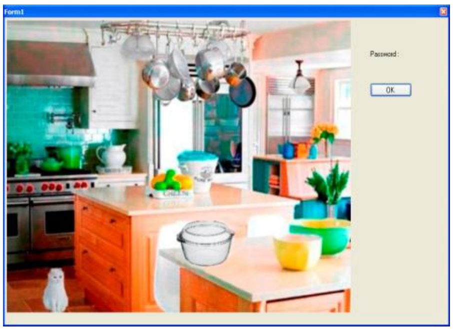
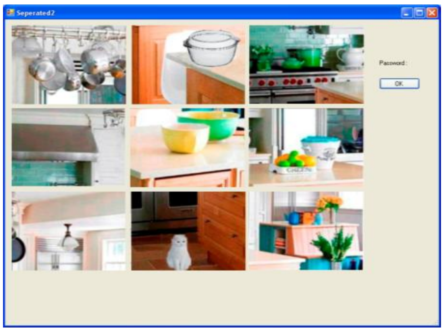
Materials and design

Eye movements of participants collected by a 60 Hz Tobii 1750 Eye Tracker and analyzed with Tobii Studio software. The developed application for the first study was updated for the second pilot study. Time taken to generate password and login, user name-password pair, user selections on the images were collected and stored by this application.

A kitchen image (Figure 4-3) was presented to the coherent-image group and the jumbled version of the image (Figure 4-4) was presented to the jumbled-image group. Jumbled version of the image was produced by the help of the free web application, which included

[19] Biederman (1972) and Biederman et al. (1973) conducted one of the first methods of disrupting scene. He used jumbled and coherent terms. In the present study, the same method is used to disrupt scene. In the following parts of the thesis, the word “jumbled” will be used instead of the “separated”.

some intermediate processes that led to the small differences in the sizes of each cells ([209-220]*[158-162]).[20]

 *Figure 4-3* Kitchen image (636*478) It was presented to the coherent-image group	 *Figure 4-4* Jumbled kitchen image ([209-220]*[158-162]) It was presented to the jumbled-image group

In the separated image method there is at least one meaningful object in each cell. To support this condition, cat and glass bowl added to the coherent kitchen image by using the Photoshop before making up the jumbled version. In addition, the following design issues were considered to make a more robust experiment design and to clear the confounding variables (cf. the first pilot study). In the login processes of the jumbled-image group, the shuffled version of the objects was presented. In both groups, the participants were allowed to click anywhere on the image. Password space (i.e., the number of clickable points) in both groups was the same.

Procedure

In the second pilot study, eye tracker was also used to obtain eye movement data. Participants were seated at a distance of approximately 60 cm from the eye tracker monitor. A calibration with 9 dots was made to calibrate the eye tracker. The remaining procedures were the same as the first pilot study (Table 4-1).

[20] Firstly; by using the “make your own puzzle” feature of the AllStarPuzzles website freely, coherent kitchen picture was uploaded to the site (Available November 20, 2010, from http://www.allstarpuzzles.com/usr/493/p.html?f=493e1fe729a02aa0dada). Then 3*3 shuffled version of the coherent picture was created by changing the row and column number as 3. Shuffled picture was obtained by cropping out the irrelevant items on the screenshot of the entire page. After that, since the developed graphical password application required to get each cells individually, each cells were cropped out individually from the shuffled image by using Photoshop.

Results

In the second pilot study, data were collected in three modalities; eye-tracker, software application recordings and questionnaires. For the eye movement analysis, each 9 items in both images were defined separately as an area of interest (AOI). The two participants in the coherent image group remembered their password whereas one of two participants in the jumbled-image group did not remember the password. The elapsed time during login is less in the jumbled-image group participants (M=9.62; SE=1.46) than the coherent-image group participants (M=10.86; SD=1.43). Total fixation count on the pass-items (M=32.5; SE=2.5) is bigger than the fixation count on non pass-items (M=19; SE=10). The most fixated AOI is the separated image of fruits. The probable reason is that the colors of fruits attract much more attentions than the others. Coherent-image group participants' cognitive load (M=4.57, SE=0.83) is more than the two jumbled-image group participants (M=17.93, SE=0.93).

Questionnaires' results showed that, both of the participants in the coherent-image group used the well known patterns (e.g., S, W shape) during creating the password whereas just one participant in the jumbled-image group used patterns. On the other hand, in the jumbled-image group, both of the participants try to memorize the identity of the object in the selected cell. Three of 4 participants reported that they create their passwords easily. All of the participants reported that they would remember their password in the second session.

Discussion

Only four participants attended the second pilot study, therefore results have the low potential to generalize. On the other hand, the following lessons were learned from the second pilot study. Login process generally completed within a few seconds. It was observed that the participants experienced difficulty in using the graphical password application, especially in the second session. The problems were identified and proposals were developed for the interface to be used in the main experiment.[21] It was also decided that in the new experiment design, each pieces of the coherent-image should have an individual object which can be named so that. This way, the confounding factors of object naming are minimized. In addition, to reduce the influences due to visual saliency, grayscale images

[21] Firstly, interfaces of the application should be changed with the new ones that participants are more familiar with using. Secondly, in the second session to login to the system participant must remember both of his/her username and password. However, since the aim of the study is to investigate the password problem, it does not matter if the participant remembers his/her username. The necessary changes should be carried out on the application for this issue.

seem more appropriate than color images.[22] Lastly, as mentioned above, jumbled image was obtained by using AllStarPuzzles website and making some intermediate processes. This intermediate processes led to the small differences in the size of each pieces. Using application that gets coherent image and produces pieces separately for the given row and column number could be a solution to this problem.

In addition to the above ones, two changes related to the questionnaires were observed to be necessary. In the former one, some of the participants could forget the strategy they used after one week, therefore it should be asked right after they create their password, in the first session instead of the second session. In the latter, questions of the questionnaire given in the Appendix B will be scaled rated from 1 to 5 to increase the interval.

Follow-up Study

The follow up study was conducted by 6 participants. All of the changes stated in discussion part of the second pilot study have been made. This follow up study was performed in order to test the developed applications, provided images, questionnaires and to measure the average completion time of the tasks before starting to conduct the main experiment. After that, necessary changes have been made to get rid of all errors of the applications and improve the user interface. The rest of the methodology parts of the follow up study were same with the main experiment.

Then the main experiment was started. In the following, information about the sample, material and apparatus, procedure, variables, data collection and data analysis of the main study will be presented.

4.2 The Experiment

4.2.1 Participants

Sixty-three participants (29 females, 34 males M= 32.05; SE= .73) participated in the experiment. All participants were the employee of the Capital Markets Board of Turkey and were able to use computers and volunteered to participate in the study.[23]

[22] Studies have shown that, usually users are more attracted to colorful objects than they are to grayscale ones (Dirik, Menon, & Birget, 2007; Therriault, Yaxley, & Zwaan, 2009).

[23] More specifically, all of them work on the Microsoft Windows 7 installed computers at work.

The experiment consisted of two sessions. The second session was performed 3 days after the first session. In the first session each participant created a graphical password and in the second session s/he tried to remember it in order to authenticate to the system. The participants were divided into four groups according to which type of image they are presented in the password creation and login tasks as shown in Table 4-3.

Table 4-3 The summary of the groups

Abbreviation for the group name	Shown image during password creation	Shown image during login
G1	Jumbled image	Jumbled image
G2	Jumbled image	The pieces of the jumbled image were shuffled
G3	Coherent image	Coherent image
G4	Coherent image	The pieces of the coherent image were shuffled. Shuffled version of the coherent image is same with the shuffled version of the jumbled image.

In the rest of the study, *coherent-image group* (includes both G3 and G4) and *jumbled-image group* (includes both G1 and G2) terms are used to differentiate participants with respect to the image they are presented during password creation.

Fifty-nine of 63 participants participated in the second session. There were approximately equal numbers of men and women between groups. The distribution of the participants according to gender, group and sessions is presented in Appendix B.1 in the Table B-1. The age distribution is given in the Figure B-1 (under Appendix B.1). And the other demographic data of participants collected via a questionnaire is given in the Table B-2 (under Appendix B.1).

All participants had normal or corrected to normal vision. Participation in the experiment was contingent on reliable eye tracking calibration and basic computer usage skills.

4.2.2 Materials and Design

Eye movements of participants were recorded by a Tobii 1750 Eye Tracker and analyzed with Tobii Studio software. The participants were made to sit at a distance of approximately 60 cm from eye tracker.

Software

The application that was developed for the first pilot study was updated with a new interface. The process flow of the application was changed according to the outcomes obtained in the pilot studies. It was written in C# programming language. Time taken to create password and login, user name- password pairs, user selections on the images during login , participants' responses to the questionnaire, participants' responses to mental rotation task and mouse movements were recorded by this application.

Since Windows 7 was installed on the computers of the employees of the capital Markets board of Turkey; employees are familiar to Windows 7 logon screen. In order to simulate the Windows 7 logon screen interface, the application was designed by using Windows 7 logon screen background images, buttons and icons. To increase participants' familiarity with the application, participants were directed by the instructions and the application was started with the test phase included password creation and login parts through the test images.

The application interface includes four main steps: password creation, login, questionnaire and mental rotation task.

1. *Password Creation:*

   In the password creation screen, there was a user name assigned automatically by the system, password and confirm password boxes as shown in the Figure 4-5. These boxes were not editable.

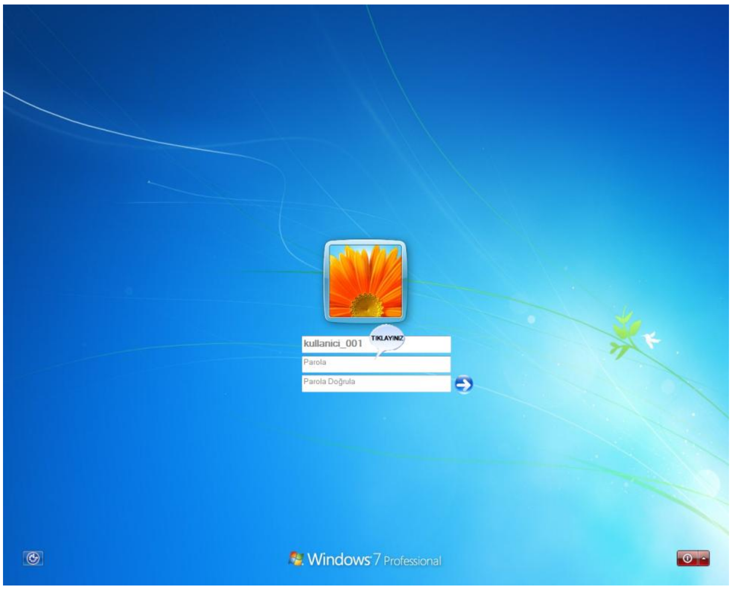

kullanici_001

Figure 4-5 Password creation screen of the application

When participants clicked on the password box an instruction screen appeared which explained how to create graphical password. After that, a password screen appeared and displayed the image according to the participants' group in the experiment. Password selection was completed after choosing 4 pieces from the presented image. If the participant accidentally selected unintended pieces, s/he could use clear button located just below the image to erase incorrect choices. Following, after participant confirmed his/her password by selecting same pieces in the same order, password creation process was completed. If the passwords that the participant typed did not match, the application prompted him/her to re-enter the password.

2. *Login:*

As in the password creation screen, in the login screen, there was a user name assigned automatically by the system and password box as shown in the Appendix B.2. This box was not editable.

When participants clicked on the password box, password screen appeared that displayed the image according to the participants' group in the experiment. Password selection was completed after choosing four pieces from the provided image. If the participant accidentally selected unintended pieces, s/he could use clear button located just below the image to erase the incorrect choices. Following, after participant clicked on the "OK" button in the login screen, login process was completed. When the participant successfully logged in, a success message was displayed and he/she was redirected to the welcome page. On the other hand, application provided the participant with a warning message, if he/she made an unsuccessful login attempt.

3. *Questionnaire:*

   Questionnaire screen was displayed, after password creation was completed. It is given in the Appendix B.2.2
   Participants filled out the questionnaire before they get access to the next part of the application. The questionnaire contained questions related to the following topics:

   - Password creation strategy
   - Demographic information
   - Usability of the given graphical password schemes

4. *Mental Rotation Task:*
   After the participant filled out the questionnaire, for removing the password from their visual working memory by simulating the effect of the passage of time, 2D mental rotation task (MRT) [24] was given as a visual distracter for at least 30 seconds.[25] A sample screenshot is given in the Appendix B.2.3.

Images

A 2362*2362 resolution image was taken in the photography workshop of Hacettepe University. A professional support was received in order to have the photos taken under sufficient light conditions and ensure that there is at least one object which can be named in each cell. Then, it was converted in to gray scale to reduce the effect of visual saliency may cause different effects between groups by color contrast.

Before the experiment, three people who saw the images for the first time were asked to name the objects in each cell. It was observed that they had difficulty in naming the lemon squeezer. The lemon squeezer was replaced with the beater by using Photoshop. In addition to that, carrots and cup were added to increase the coherence of each cell between them. The

[24] Shapes used in the MRT were retrieved on January 19, 2012 from the http://psytoolkit.leeds.ac.uk/lessons/mentalrotation.html by the permission of Dr. Gijsbert Stoet.

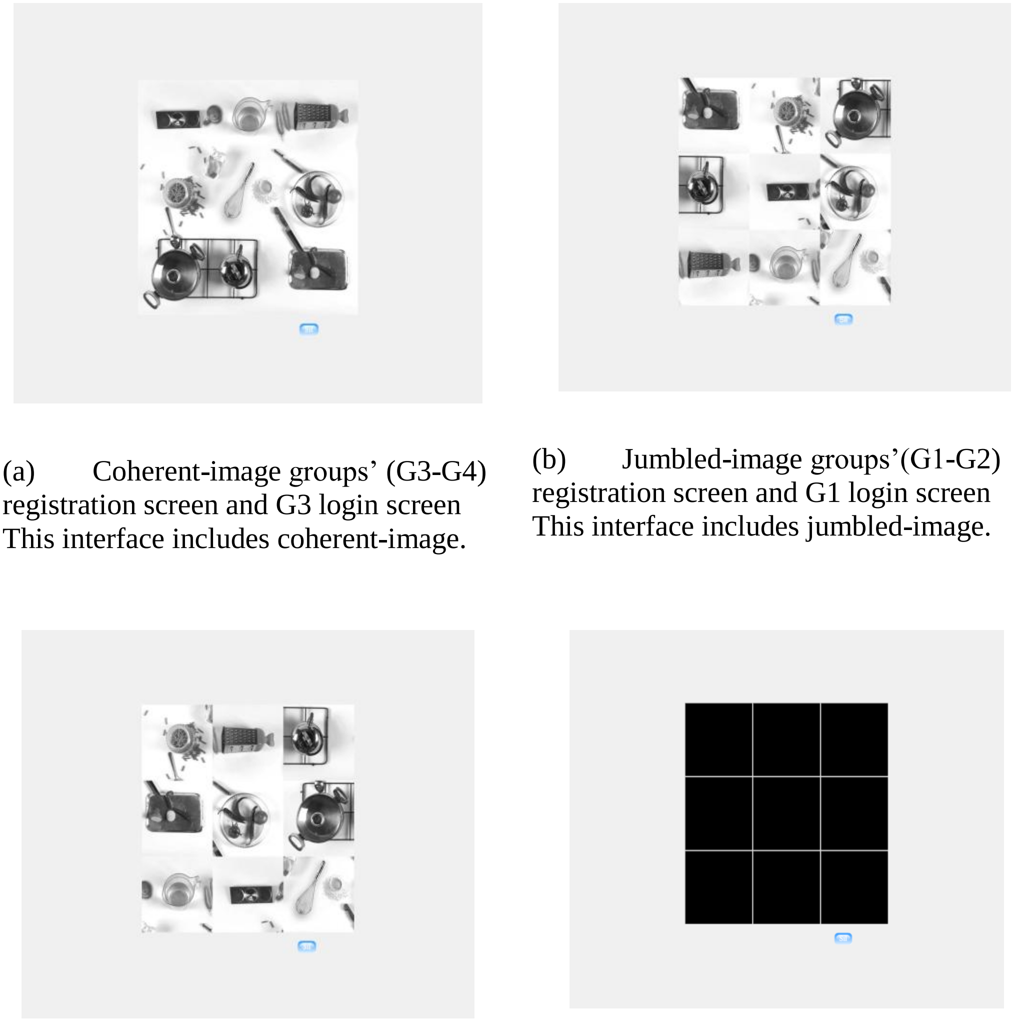
[25] In the previous studies (e.g. Chiasson, Van Oorschot, & Biddle, 2007; Thorpe, & Oorschot, 2007; Van Oorschot, & Thorpe, 2010), MRT used as a visual distracter (30 seconds).

original and the manipulated images can be seen in the Figure B-4 and Figure B-5 (under Appendix B.3).

Lastly, another application was developed to produce jumbled image out of coherent one. It was written in Windows Presentation Foundation (WPF). The coherent image, the desired size and the number of pieces were passed to the WPF application as an input and it returns pieces separately as an output. Also, WPF application was able to produce C# code of the coherent and jumbled screens concerning graphical passwords. The main screen of the WPF application was shown in the Appendix B.2.4.

During password creation, while participants in the G1 and G2 groups were presented the jumbled image, participants in the G3 and G4 groups were given the coherent image. On the other hand, during the first phase of login, all participants, regardless of their group membership, were asked to enter their password on the black grid on a gray background by supposing that the image containing the cells they picked their password was covered with a black blanket. The black-grid-screen[26] is designed to understand participants' behavior in choosing their password, whether they choose their passwords by memorizing the spatial location of the pieces without object identification or not. If participants do not use spatial coding even they recognize the object identity, it will be expected that they cannot log in to the system by selecting the password on the black grid. On the other hand, if participants use spatial coding, they will authenticate successfully, even they do not see the object itself. After three unsuccessful login attempts on the black-grid-screen, the participants in G1 and G3 were provided with the same image they used while creating their password. On the contrary, G2 and G4 were provided with the shuffled version of the related image (Figure 4-6.)

[26] This method is originated from the Hollingworth (2005) study. He carried out the study across three groups. In the first group, which he called *target present preview,* subjects were shown a natural scene image involving the target object for 20 seconds. Then, in the next screen, the target object was displayed and subjects were requested to click on the black screen that would be displayed to specify the location of the target object without seeing the natural scene image.



(a) Coherent-image groups' (G3-G4) registration screen and G3 login screen This interface includes coherent-image.

(b) Jumbled-image groups'(G1-G2) registration screen and G1 login screen This interface includes jumbled-image.

(c) G2 and G4 groups' login screen presented after three unsuccessful login attempts on the black-grid-screen. In this page, 3 login attempts are allowed.
This interface includes shuffled image.

(d) Black-grid-screen on a gray background. At first, all groups performed login task on the black-grid-screen. Three login attempts are allowed on the black-grid-screen.

Figure 4-6 Password creation and login screens of each group

An initial testing phase was conducted shortly before the main experiment. Free images found on internet used for practice are shown in the Figure B-2 and Figure B-3 (under Appendix B.3).

There were three types of independent variables in the experiment. The first one is the group (G1, G2, G3 or G4) based on the login screen image shown after black-grid-screen, which is a between group variable. The second one is the presentation method (coherent-image (G3-G4) or jumbled-image (G1-G2)) based on the registration screen image, which is again a between group variable. The last independent variable is the session (Session 1 or Session 2), which is a within-subject variable. Login success rate, time taken to generate password, time taken to login, fixation count, total fixation duration and Levenshtein String Distance were measured as dependent variables.

At the beginning of the application, participant's username was assigned automatically by the system (i.e., "Kullanici001"). Then in the second session, the experimenter (not the participant) entered the username assigned in the first session, so that participants were not expected to memorize and recall their username. The application started with the test phase and included information screens to increase the participants' familiarity. In the test phase, in order to eliminate the possible confounding effects of test phase image type, each group provided with the same type of image (coherent or separated) that they would saw while creating their main graphical password.

The questionnaire and Mental Rotation Task were also given before login process (i.e., in the first session). Since participants might forget in time the strategy they adopted during password creation, a relevant question was asked as the first question of the questionnaire, right after they create their password.

4.2.3 Procedure

The participants were informed about the task and asked to choose realistic passwords which they could remember but that would be difficult for others to guess. They were told that they would perform study alone and guided through instruction screens so it was important to read instructions carefully.

A 9-point calibration was used on the eye tracker.[27] Then application started automatically. The experiment consisted of two sessions. The second session was performed three days after the former one.

[27] A 50 Hz. non-intrusive eye tracker recorded eye- tracking data. The eye tracker was integrated into a 17" TFT monitor with a resolution of 1280 X 1024 pixels. Subjects were seated in front of the screen at an approximate viewing distance of 60 cm. Spatial resolution and accuracy of the eye tracker was about 0.25° and 0.50° degrees respectively.

Table 4-4 Methodology of the main experiment

Phase	Number Of Trials	Steps
I. Practice	One trial These data were not used in the analysis.	Explain System Create Test Password Login (Test Password)
II. Password Generation	Until password successfully created These data were recorded as user passwords.	Create
III. Questionnaire	These data were used in the analyses.	Questionnaire (Appendix B.4)
IV. Mental Rotation Task	It took 30 seconds (Visual distracter)	
V. Login	Maximum three trials on the black-grid-screen. Maximum three trials after the black-grid-screen, on the displayed image. These data was used in the analyses.	Login
VI. Three days Retention	Maximum three trials on the black-grid-screen. Maximum three trials after the black-grid-screen, on the displayed image. These data was used in the analyses.	Login

The first session was divided into five phases: practice, password generation, questionnaire, mental rotation task and login, as shown in Table 4-4. The practice phase was used to explain the task and familiarize participants with the user interface of the developed application. It consisted of password generation and login processes. Participants were guided by the on-screen instructions about how to create a password. At the end of the practice phase, an instruction screen appeared informing the participant about the end of the practice phase. The next phase was the password generation. Participants picked their passwords by clicking 4 click-points on the provided image. The selected item was framed with a blue line, for the aim of giving feedback to the participant. After that, in order to obtain demographic data, participants filled out the questionnaire that aimed at gathering their perception concerning usability of the system and their attitudes towards password creation (Appendix B.4). In the fourth phase, a 30 second mental rotation task was administered to disrupt visual memory. The questionnaire and the mental rotation task serve as filler tasks between the password generation step and the login step. Finally, in last phase of the first session, the participant was asked to re-select his/her password on the black grid by supposing that the image containing the cells they picked as their password was covered with a black blanket. After three unsuccessful login attempts in the black-grid-screen, the participant was provided with an image depending on his/her group type. Structure and screen flow of the application is shown in the Figure 4-7.

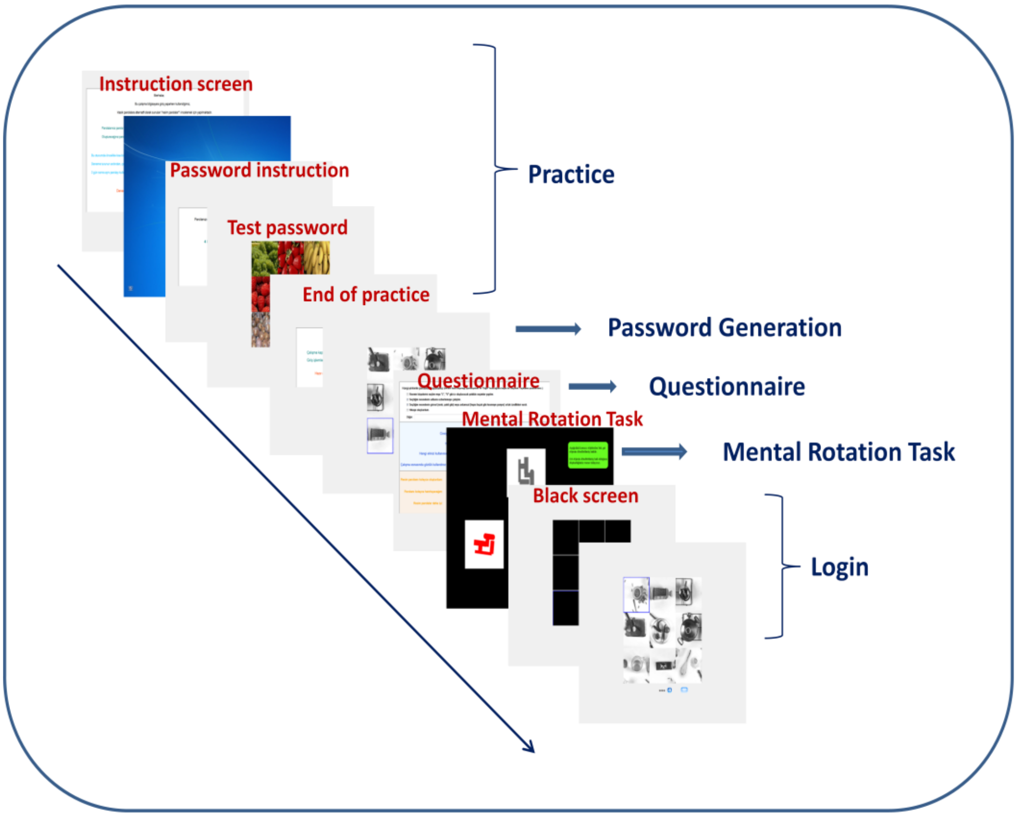


Figure 4-7 Screen flow of the application

The second session was conducted three days after the first one. Participants tried to log in to the system by re-selecting their password. It was the same as the login phase of the first session. After three unsuccessful login attempts in the black-grid-screen, the participants were provided with an image depending on their group type. They could re-try three times to recall their password on the image displayed screen. Otherwise, they were assumed that they forgot their password.

4.2.4 Data Collection and Methodological Assumptions for the Analyses

Summary of the collected data in the present study is given below;

- Age, gender, usage of eye glasses during the study and right or left handed were collected by means of the questionnaire.
- Password creation strategy and the participants' perception on the usability of the system were collected by means of the questionnaire.
- Passwords, login success, elapsed time to create and confirm password and time taken to login to the system were collected by means of the application and eye tracker.
- Eye movements (number of fixations, total fixation duration and scan path) were collected by the eye tracker.
- Test password, answers given to mental rotation task and mouse movements on the password screens were also collected by the application. However, these data would not be used in the analyses.

An overall description of data analyses are presented below.

During password creation, participants were divided into two groups: coherent-image and jumbled-image. On the other hand, during the first stage of login, all participants, regardless of their group membership, were asked to re-select their password on the black grid. When they made three unsuccessful login attempts, they were provided with an image on the password selection screen. Participants were divided into four groups depending on the displayed image type after the black-grid-screen: coherent, shuffled coherent, jumbled and shuffled jumbled. The vast majority of participants were able to login successfully on the black-grid-screen; therefore they did not see even the image. So, in most cases, there were not enough data points to make statistical analyses on the image displayed after the black-grid-screen. Therefore, data analysis was conducted between the coherent-image groups (consist of G3 and G4) and the jumbled-image groups (consist of G1 and G2) instead of the G1, G2, G3 and G4 separately.

In the login phase, after the participant selected all four pieces of his/her password, s/he was returned to the login screen. Then, the participant clicked on the "OK" button in the login screen to complete the login process.[28]

Another C# desktop software application that included functions to parse data files exported by using data export tool in Tobii Studio and to create project file in the acceptable format of the eyePatterns software (Version 0.91) was developed for the aim of automating some processes of obtaining required data to use in the analysis.

Kolmogorov-Smirnov test was conducted to test whether the distribution was normal. If there were small deviations from normality, Shapiro-Wilk test was also conducted to make an informed decision about the extent of non-normality. For the aim of meeting normality and the homogeneity of variances assumptions, Logarithmic or Reciprocal transformations were applied to data which violated parametric assumptions. After that, if there were outliers, in order to reduce the impact of them, those values were changed with the next highest or mean plus two standard deviations score. When it comes to reporting transformed data, although statistical tests were performed on the transformed values, means were reported in the unit of the untransformed scale. Tests that were conducted to assess the normality of the data distributions and transformations that were applied to meet normality and the homogeneity of variances assumptions are presented for each analysis under the result section.

Analysis of the login success

As mentioned above, a few participants re-selected passwords more than three times. (11 of 120 attempts) Even though, those participants did not click on the "OK" button in the login screen, if they selected the correct password within their first three choices, they were assumed to be successfully logged in.

In this part, four analyses were conducted. Two three-way Loglinear analysis were conducted, in order to investigate the login success of the participants in their first and all attempts on the black-grid-screen (Login Success * Session * Group Type). In the third one, Chi-square test was performed to reveal whether there is a relationship between participants' gender and their login success. Lastly, since frequencies of participants that fail to log in to

[28] A few participants (11 of 120 login attempts) re-select their passwords by clicking on the password box without clicking the "OK" button in the login screen. To avoid possible confounding effects resulting from differences in the number of trials, only the first three selections of those participants were taken into account.

the system was smaller than 5, Fisher's exact test was performed to see whether there is a relationship between participants' gender and their overall, in and after the black-grid-screen, login success.

Analysis of the elapsed time to create and confirm password and time taken to login

Event data of the participants in each session were exported by using Tobii Studio. Those files were passed to the developed application as an input and it generated and returned separate text files for each participant in each session. Text files contain separate lines for start and end times of password creation, confirmation and login processes. Start time is the time when participant clicked the related button or box which opens the password selection screen (black-grid-screen or screen that contain password image).[29] End time is the time when participant clicked the related button which closes the password selection screen.

As mentioned above, a few participants were selected passwords more than three times. If it was a password creation or confirmation process, elapsed time of the first attempt was taken in to account (3 of 122 attempts). On the other hand, if it was a login process, average elapsed time of the attempts (up to three trials) was calculated.

In this part, three analyses were performed. Firstly a t-test was performed to single out whether there was a difference in the time taken to create password between groups. Secondly, since data concerning the time taken to confirm password did not show normal distribution, Mann-Whitney test was performed to see if there is difference in the elapsed time during password confirmation between groups. Lastly, two-way mixed ANOVA was performed to investigate the effect of group type in the time taken to login in each session (Group type* Time taken to login*Session).

Analysis of the eye movement data

Eye movement data analyses were conducted on data collected from login attempts on the black-grid-screen and from password creation screen. In this part, fixation counts, total fixation durations and Levenshtein String Distances were calculated. And for this purpose, password creation screens and login attempts (up to three) on the black-grid-screens were defined as scenes. Those scenes separated into 9 Areas of Interests (AOI) (Appendix C.3).

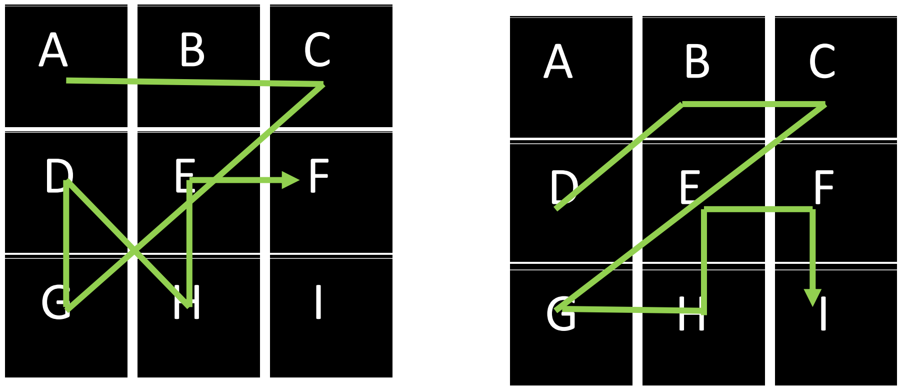
[29] In a few attempts (7/242), participants hold down the mouse button after clicking on the related button or box. In these situations, Tobii gets the first clicking time as an event time. But in fact, password selection screen will not open as long as participant hold down the mouse button. Therefore, accurate start time of those events was obtained by watching the related video records.

Three statistical tests were performed about the fixation counts of login attempts on the black-grid-screen. Firstly, two-way mixed ANOVA in which session factor is used as repeated-measure variable and group type is used as between group variable was performed to investigate the effect of group type in average fixation count (includes all AOIs) obtained from the login attempts (up to three) on the black-grid-screen for each session. Secondly, average fixation counts on 9 AOIs of each participant's login attempts (up to three) on the black-grid-screen, in the first session separated into two categories by whether this AOI is part of the related participant's password. And, fixation counts in each category were averaged. Then, two-way mixed ANOVA was performed to investigate the effect of group type in fixation count on the black grid in the first session between the area of interests which were parts of their password and not parts of it. Lastly, another two-way mixed ANOVA was performed to investigate the effect of group type in fixation count on the black grid in the second session between the area of interests which were parts of their password and not parts of it, after similar operations in the second one.

As the previous one, 3 statistical tests were performed about the total fixation durations concerning login attempts on the black grid. Firstly, two-way mixed ANOVA in which session factor is used as repeated-measure variable and group type is used as between group variable was performed to investigate the effect of group type in average total fixation duration(includes all AOIs) obtained from the login attempts (up to three) on the black grid for each session. Secondly, average total fixation durations on 9 AOIs of each participant's login attempts (up to three) on the black grid, in the first session separated into two categories by whether this AOI is part of the related participant's password. And, total fixation durations in each category were averaged. Then, two-way mixed ANOVA was performed to investigate the effect of group type in total fixation duration on the black grid in the first session between the area of interests which were parts of their password and not parts of it. Lastly, another two-way mixed ANOVA was conducted to investigate the effect of group type in total fixation duration on the black grid in the second session between the area of interests which were parts of their password and not parts of it, after similar operations in the second one.

Finally, Levenshtein String Distance (LSD) analyses were performed on the black-grid-screens and password creation screens. LSD is named by Vladimir Levenshtein. It represents the minimum number of operations (i.e., insertion, deletion or substitution of a single character) required to transform one string into the other. The example below illustrated how

LSD between two strings which are sequences of fixations represented by the identifier (i.e., letter code) of AOI (Figure 4-8) on the black grid is calculated.



(a) S_0= ABCEGDHEF (b) S_1=DBCEGHEFI

Figure 4-8 Sample sequences of fixations on the black grid

LSD between S_0 and S_1 is 3 and each step is shown below.

1. ABCEGDHEF → DBCEGDHEF (substitution of ‘D’ for ‘A’)
2. DBCEGDHEF → DBCEGHEF (deletion of ‘D’)
3. DBCEGDHEF → DBCEGHEFI (insertion of ‘I’ at the end)

Firstly, AOIs of predefined black-grid-scenes and password-creation-scenes were exported by using Tobii Studio. Those files were passed to the developed application as an input and it generated and returned file contains AOI string on the related scene for each participant in the separate lines. Ultimately, for each groups, 3 files(one for each login attempts) for login attempts on the black-grid-screen in the first session, 1 file for the password creation scenes and the other 3 files (one for each login attempts) for login attempts on the black-grid-screen in the second session were generated. After that, those files and file that contains each participant’s id and password pairs in the separate lines were passed to the developed application as an input and it generated and returned project file in an appropriate format for the eyePatterns application. Then, similarity between participant’s password and his/her AOIs string on the related scene was evaluated for each participant by using eyePatterns application. LSD of black-grid-screens (up to three) of each participant was averaged. Two statistical analyses were performed. Firstly, two-way mixed ANOVA in which session factor is used as repeated-measure variable and group type is used as between group variable was performed to investigate the effect of group type in LSD concerning login attempts on the black grid for each session. Secondly, t-test was conducted to investigate the effect of group type in LSD obtained during password creation.

Discovering Patterns

In this part, participant's passwords were investigated in order to model similar participant choices and the effect of given image types on these choices.

In addition to that, each 9 item was examined in terms of how many times and in which order it was used as participant's password. Results are given in clustered and stacked column charts and on the hierarchical tree.

Analysis of the collected data by means of the questionnaire

Demographic information, password creation strategies and the participant's perception on the usability of the given graphical password schemes were collected. In this part, difference between the groups in password creation strategy and perception on the asked usability metrics are given in clustered and stacked column charts.

4.2.5 Interim Summary: Main Experiment

Sixty-three volunteered participants (29 females, 34 males M= 32.05; SE=.73) took part in the experiment. The experiment consisted of two sessions. The second session was performed 3 days after the first session. In the first session the participants created a graphical password and in the second session they tried to remember it in order to authenticate to the system. The participants were divided into four groups according to which type of image they are presented in the password creation and login tasks.

Microsoft C# desktop application was developed for the study. The application interface included four main steps: password creation, login, questionnaire and mental rotation task. A 2362*2362 high resolution image which was taken in the photography workshop of Hacettepe University, converted to grayscale and modified via Photoshop

Participants in the coherent-image group chose their graphical passwords on that image. On the other hand, participants in the jumbled-image group chose their graphical passwords on the jumbled version of the image. Demographic information and questions related to usability were collected by means of the questionnaire. Password, login success, elapsed time to create and confirm password and time taken to login to the system were collected by means of the application and eye tracker. And eye movements (number of fixations, total fixation duration and scan path) were recorded by the eye tracker. In the next section analyses of the data will be given.

4.3 Results

In this chapter, analysis of the login success will be presented. Afterwards, analysis of the elapsed time to create and confirm password and time taken to login will be given. Then, the results of the analysis of eye tracking data (number of fixations, total fixation durations and Levenshtein String Distances) will be presented. Then in the next section participant's passwords were investigated in order to model similar participant choices. Lastly, the analysis of the collected data by means of the questionnaire will be presented.

4.3.1 Analysis of the login success

The first two analyses are about the login success of participants in their first and overall attempts on the black-grid-screen. Participants in the jumbled-image and coherent-image groups divided in to two other groups depends on the images they saw if they fail to login in the black-grid-screen .Therefore, black-grid-screen analyses were made just between the coherent-image and jumbled-image groups.

In the first login attempt

The three way Loglinear analysis produced a final model that retained login success and group type effects. The likelihood ratio of this model was $\chi2$ (4) = 0, p=1. This indicated that the interaction between login success and group type was significant, independently from the session. χ^2 (1) = 5.20, p<.05. Success ratios of the groups are shown in the Figure 4-9.

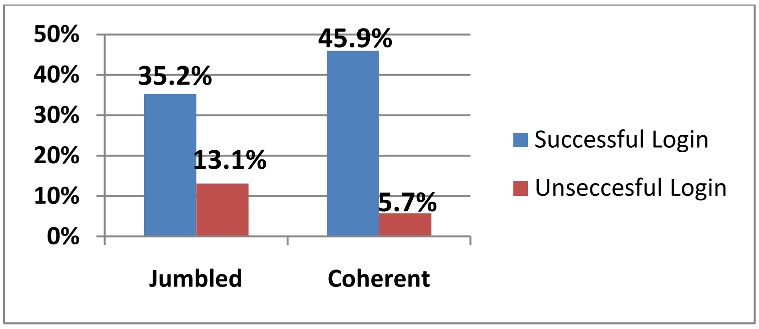


Figure 4-9 Success ratio of the groups in the first login attempts on the black-grid-screen in both sessions. In the jumbled-image group, 43 of 59 first login attempts and in the coherent-image group, 56 of 63 first login attempts on the black-grid-screen were successful.[30] Thus, 43 out of 122 (35.2%) and 56 out of 122 (45.9%) represents respectively success ratios of jumbled-image group and coherent-image group in the first login attempts on the black-grid-screen in both sessions.

Odd ratios indicated that the odds of successful login were 2.98 times higher in the coherent-image group than the jumbled-image one.

In the overall login attempt on the black-grid-screen

The participants, who selected the wrong password three times consecutively in the black-grid-screen, were provided with the image according to their groups.

The three way Loglinear analysis produced similar to the final model of the first attempt in black-grid-screen that retained login success and group type effects. The likelihood ratio of this model was $\chi 2$ (4) = 0, p=1. This indicated that the, interaction between login success and group type was significant, independently from the session. χ^2 (1) = 4.96, p<.05. Success ratios of the groups are given in the Figure 4-10.

[30] Sum of the first login attempts in both sessions was 59 for the jumbled-picture group and 63 for the coherent-picture group.

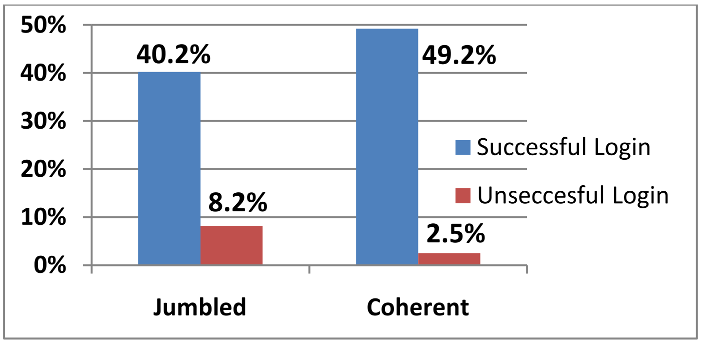


Figure 4-10 Success ratio of the groups, in the all login attempts (up to three) on the black-grid-screen, in both sessions. In the jumbled-image group, 49 of 59 and in the coherent-image group, 60 of 63 login attempts on the black-grid-screen (up to three) were successful. Thus, 49 out of 122 (40.2%) and 60 out of 122 (49.2%) represents respectively success ratios of jumbled-image group and coherent-image group in the all login attempts on the black-grid-screen in both sessions.

Odd ratios indicated that the odds of successful login were 4.08 higher in coherent-image group than the jumbled-image one.

Overall Success

In the overall success, participant's login success in and after the black-grid-screen were both taken into account. As shown in the Figure 4-11, the vast majority of participants were successful. In some groups there is not any failed login. So, it does not have enough data points to make statistical analysis meaningful.

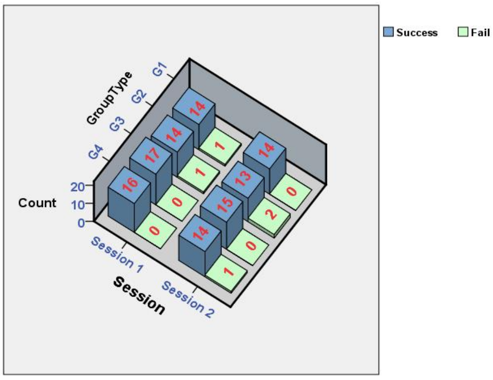


Figure 4-11 Numbers of the successful logins taken without discrimination as to section of the login (in or after the black-grid-screen) for each group and session.

Gender difference in login success

Chi-square test was performed to see whether there is a relationship between participants' gender and their login success. There was not any significant difference between gender and their login success in the black-grid-screen, independently from the session, χ^2 (1) = 1.96, p>.05. Success of each gender in each session is given in the Appendix C.1 in the Figure C-1.

Another categorical analysis was performed in order to see whether there is a relationship between participants' gender and their overall, in and after the black-grid-screen, login success. Since expected frequencies of participants smaller than 5, Fisher's exact test was performed. There was not any significant difference between gender and their login success, independently from the session (P=.664, FET). Success of each gender in each session is given in the Appendix C.1 in the Figure C-2.

In summary, interaction between the login success and the group type was significant, independently from the session both for the first login attempt (χ^2 (1) = 5.20, p<.05) and for the overall attempts on the black-grid-screen (χ^2 (1) = 4.96, p<.05).

4.3.2 Analysis of the elapsed time to create and confirm password and time taken to login

Time taken to create password

In order to meet normality and the homogeneity of variances assumptions logarithmic transformation was made. After that, to remove and reduce the impact of the outlier, that value was changed with the next highest score. A t-test was performed to see if there is difference in the time taken to create password between groups. On average, there was not any significant difference between the jumbled-image group (M=25177, SE=2875) and the coherent-image group (M=22368, SE=2026) in the time spent on password creation, t(61)=0.81, p>.05, r =.10.

Time taken to confirm password

In order to remove and reduce the impact of the outlier, logarithmic transformation was made. Then remaining few outliers was changed with the next highest score. Since data did not show normal distribution, Mann-Whitney test was used to see if there is difference in the time taken to confirm password between groups. Time taken to confirm password in the jumbled-image group (Mdn = 5748) did not differ significantly from the coherent-image group (Mdn = 5583), U= 470, z=-.344, ns, r=-.04

Time taken to login

59 of 63 participants logged in the both sessions. Two-way mixed ANOVA in which session factor is used as repeated-measure variable and group type is used as between group variable was performed to investigate the effect of group type in the time taken to login in each session. In order to remove and reduce the impact of the outlier and meet normality and the homogeneity of variances assumptions logarithmic transformation was made.

There was not a significant interaction effect between group type and the session, F (1, 57) =1.28, p>.05. To break down this interaction, t-tests were performed comparing time taken to login in each session across jumbled-image and coherent-image group. On average, in the first session, time taken to login in the jumbled-image group (M=9757, SE=1174) did not differ significantly from the coherent-image group (M=7846, SE=698), t (57) =1.06, p>.05. On the other hand, in the second session, participants in the jumbled-image group (M=10106, SE=1336) spend more time during login than the participants in the coherent-image group (M=7070, SE=582) (Figure 4-12). This difference was significant t(57)=2.02, p<.05 and it did represent small-sized effect r=.26.

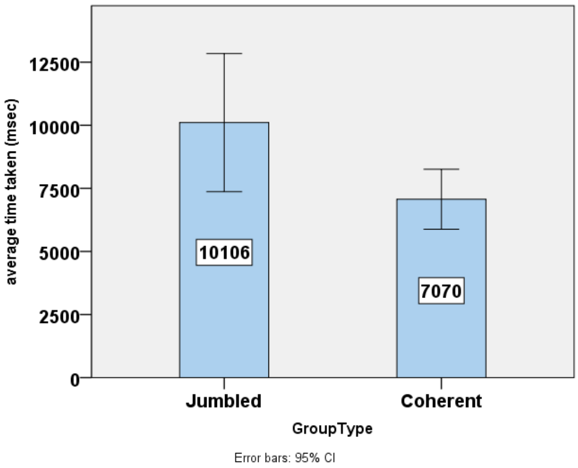


Figure 4-12 Mean login times of coherent-image and jumbled-image groups on the black-grid-screen in the second session. Time is in millisecond.

Participants spend approximately equal time during login between the first (M=8786, SE=683) and second session (M=8562, SE=741). And there was not a significant main effect of the session in the time taken to login, F (1, 57) =.374, p>.05.

Independently from the session, there was not a significant difference between the jumbled-image group (M=9932, SE=881) and the coherent-image group (M=7458, SE=453) in the time taken to login, F (1, 57) =2.85, p>.05.

In summary, there was no significant difference between jumbled-image and coherent-image groups in the time taken to create (t (61) =0.81, p>.05) and confirm password (U= 470, z=-.344, ns, r=-.04). On the other hand, in the second session, there is a significant difference between jumbled-image (M=10106, SE=1336) and coherent-image group (M=7070, SE=582) in the time taken to login t (57) =2.02, p<.05.

4.3.3 Eye movement analyses

In the login analyses, one participant had to be excluded due to calibration problems. In addition, three participants of the coherent-image group and one participant of the jumbled-image group did not attend to the second session.

Fixation Count

Fifty eight of 63 participants' fixation counts were calculated. In order to meet normality and the homogeneity of variances assumptions, reciprocal transformation was made. After that, to remove and reduce the impact of the outliers, those two values were changed with the mean plus two standard deviations. Two-way mixed ANOVA, in which session factor is repeated-measure variable and group type is between group variable, was performed to investigate the effect of group type in fixation count on the black grid for each session.

Fixation counts of participants in the first session (M=16.3, SE=1.42) was more than the fixation counts of participants in the second session (M=14.8, SE=1.72). And there was a significant main effect of the session in the fixation counts, F (1, 56) =10.2, p<.05.

There was not a significant interaction effect between group type and the session, F (1, 56) =2.96, p>.05. To break down this interaction, t-tests were performed comparing fixation counts in each session across jumbled-image and coherent-image group. On average, in the first session, there was not a significant difference between the coherent-image group (M=14.3, SE=1.64) and the jumbled-image group (M=18.3, SE=2.3) in fixation counts, t (56) =-.76, p>.05.

There was not a significant difference between the coherent-image group (M=12.7, SE=1.04) and the jumbled-image group (M=18.5, SE=1.90) in the fixation counts on the black grid F

(1, 56) =2.57, p>.05. On the other hand, in the second session, participants in the coherent-image group (M=11.1, SE=1.22) had less fixations on the black grid than the participants in the jumbled-image group (M=18.6, SE=3.09) (Figure 4-13). And this difference was significant t (56) =-2.00, p=.05 and it did represent a small-sized effect r=.26.

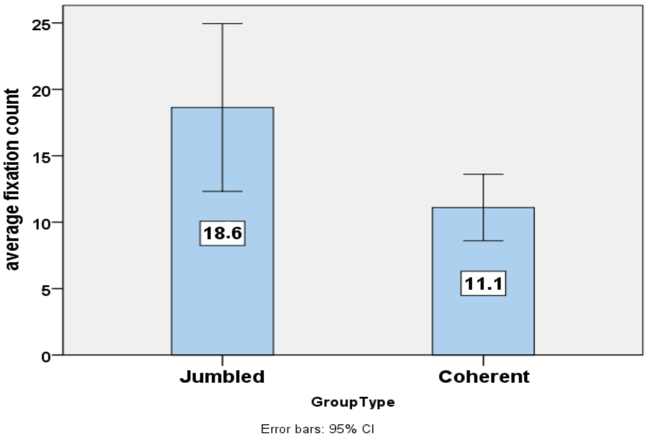


Figure 4-13 Mean average fixation counts of jumbled-image and coherent-image groups on the black grid in the second session.

Fixation count on pass-items and non pass-items, on the black grid

In the first session

58 of 63 participants' fixation counts were calculated. In order to meet normality and the homogeneity of variances assumptions, reciprocal transformation was made. Since there were some zero values, 1 was added to each value before calculating the inversion value. Two-way mixed ANOVA in which whether it is part of the password factor is used as repeated-measure variable and group type is used as between group variable was performed to investigate the effect of group type in fixation count on the black grid between the area of interests which were parts of their password and not parts of it in the first session.

Fixation counts of participants in the area of interests which were not parts of their password (M=2.89, SE=.24) was less than the fixation counts of participants in the area of interests which were parts of their password (M=1.13, SE=.15) (Figure 4-14). And there was a significant main effect of the whether AOI was part of the password in the fixation counts, F (1, 61) =79.9, p<.05.

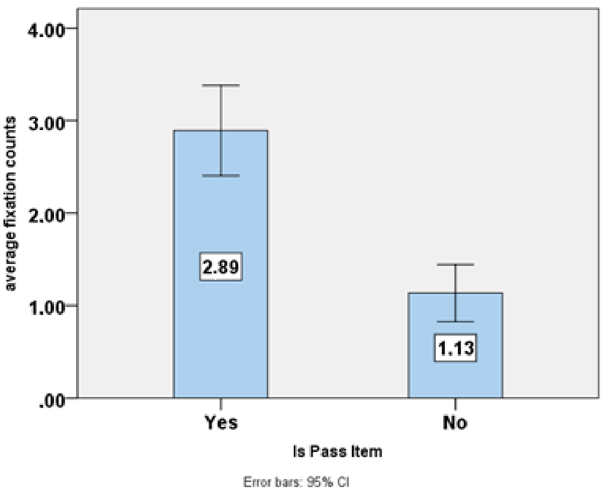


Figure 4-14 Mean average fixation counts on pass-items vs. non pass-items on the black grid in the first session.

In the second session

58 of 63 participants' fixation counts were calculated. In order to meet normality and the homogeneity of variances assumptions, reciprocal transformation was made. Since there were some zero values, 1 was added to each value before calculating the inversion value. Two-way mixed ANOVA in which whether it is part of the password factor is used as a repeated-measure variable and group type is used as between group variable was performed to investigate the effect of group type in fixation count on the black grid between the area of interests which were parts of their password and not parts of it in the second session.

Fixation counts of participants in the area of interests which were not parts of their password (M=1.05, SE=.21) was less than the fixation counts of participants in the area of interests which were parts of their password (M=2.55, SE=.23) (Figure 4-15). And there was a significant main effect of the whether AOI was part of the password in the fixation counts, F (1, 56) =111.3, p<.001.

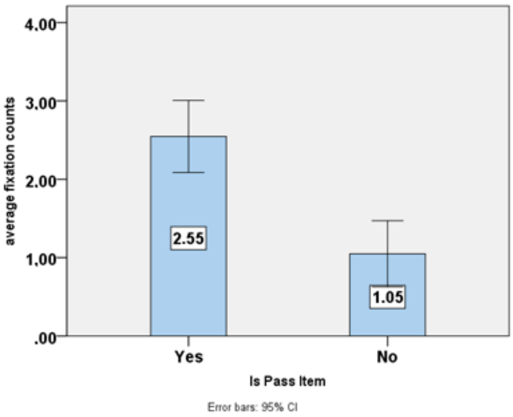
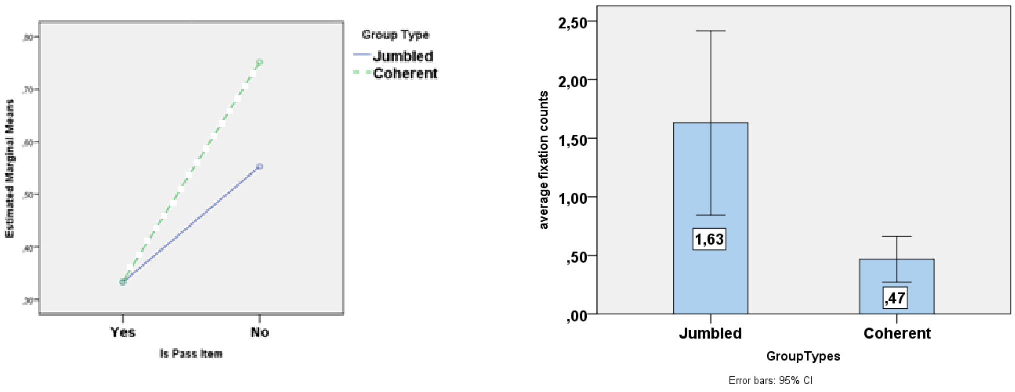


Figure 4-15 Mean average fixation counts on pass-items vs. non pass-items on the black grid in the second session.

There was a significant interaction effect between group type and whether AOI was part of the password on the black grid in the second session, F (1, 56) =10.7, p<.05. Estimated marginal means and t-tests were used to determine the nature of this interaction. The graph and the means clearly shows that fixation counts on the AOI which are pass-items are very similar for both the coherent-image (M=2.40, SE=.26) and jumbled-image group (M=2.70, SE=.39) (Figure 4-16(a)). And this difference was not significant, t (56) =-.40, p>.05. On the other hand, that fixation counts of the coherent-image group (M=.47, SE=.09) on the AOI which are not pass-items are less than the jumbled-image group (M=1.63, SE=.38) (Figure 4-16 (b)). And this difference was significant, t (56) =-3.17, p<.05.



(a)Fixation counts on the AOI which are pass-items are very similar for both coherent and jumbled image groups; whereas fixation counts on the AOI which are not pass-items are differ.

(b)Mean average fixation counts on the AOI which are non pass-items, in the second session.

Figure 4-16 Average fixation counts between group type and whether AOI was part of the password on the black grid in the second session

In summary, there was a significant main effect of the session in the fixation counts, F (1, 56) =10.16, p<.05. Fixation counts of participants in the first session (M=16.33, SE=1.42) was more than the fixation counts of participants in the second session (M=14.87, SE=1.72). In the second session, participants in the coherent-image group (M=11.10, SE=1.22) had less fixations on the black grid than the participants in the jumbled-image group (M=18.63, SE=3.09). And this difference was significant t (56) =-2.00, p=.05

In the first session, there was a significant main effect of the whether AOI was part of the password in the fixation counts, F (1, 61) =79.92, p<.05. Also, in the second session, there was a significant main effect of the whether AOI was part of the password in the fixation counts, F (1, 56) =111.3, p<.001. In the second session, fixation counts of the coherent-image group (M=.47, SE=.09) on the AOI which are not pass-items are less than the jumbled-image group (M=1.63, SE=.38). And this difference was significant, t (56) =-3.17, p<.05.

Average Total Fixation Duration

Fifty eight of 63 participants' total fixation durations were calculated.[31] In order to meet normality and the homogeneity of variances assumptions, reciprocal transformation was made. After that, to remove and reduce the impact of the outlier, that value was changed with the next highest values. Two-way mixed ANOVA in which session factor is used as repeated-measure variable and group type is used as between group variable was performed to investigate the effect of group type in fixation duration on the black grid for each session.

There was not a significant interaction effect between group type and the session, F (1, 56) =.46, p>.05. To break down this interaction, t-tests were performed comparing fixation durations in each session across jumbled-image and coherent-image group. On average, in the first session, there was not a significant difference between the coherent-image group (M=5787, SE=.54) and the jumbled-image group (M=7543, SE=1.02) in the fixation duration on the black grid, t (56) =1.51, p>.05.

On the other hand, in the second session, participants in the coherent-image group (M=5172, SE=.46) again had less fixation duration on the black grid than the participants in the

[31] Total fixation duration is the total length of the fixations within an AOI (Tobii, 2010) In the present study, even user did not fixated on the specific AOI, total fixation duration on that AOI was computed (total fixation duration includes zero).Details about the methodological issues of analyses are accessible under the "Data Collection and Methodological Assumptions for the Analyses" section.

jumbled-image (M=7725, SE=1.18). And this difference was significant t (56) =2.01, p<.05 and it did represent small-sized effect r=.26.

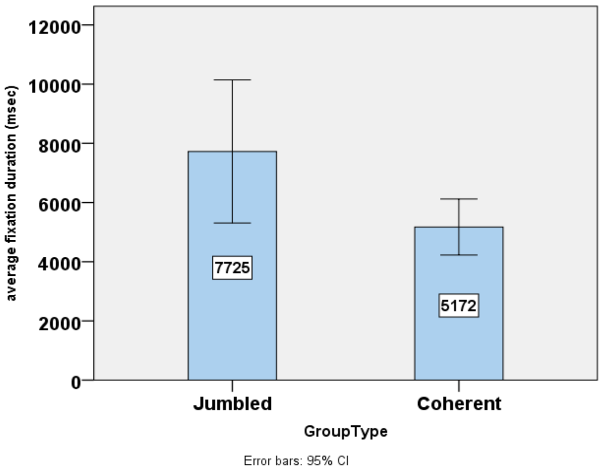


Figure 4-17 Mean average total fixation durations of jumbled-image and coherent-image groups on the black grid in the second session.

Independently from the session, participants in the coherent-image group (M=5480, SE=.35) had less total fixation duration on the black grid than the participants in the jumbled-image group (M=7634, SE=.77). And this difference was significant, F(1,56)=4.10, p<.05.

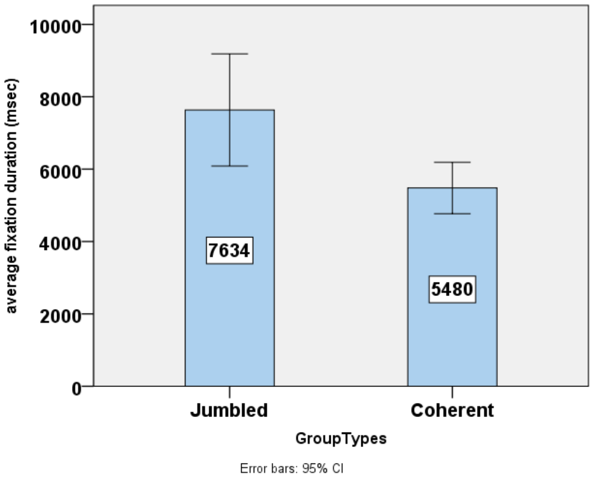


Figure 4-18 Mean average total fixation durations of jumbled-image and coherent-image groups on the black grid, independently from the session.

Total Fixation Duration on pass-items and non pass-items, on the black grid

In the first session

Fifty eight of 63 participants' total fixation duration was calculated. In order to meet normality and the homogeneity of variances assumptions, reciprocal transformation was made. Since there were some zero values, 1 was added to each value before calculating the inversion value. Two-way mixed ANOVA in which whether it is part of the password factor is used as repeated-measure variable and group type is used as between group variable was performed to investigate the effect of group type in total fixation duration on the black grid between the area of interests which were parts of their password and not parts of it in the first session.

Total fixation duration of participants in the area of interests which were not parts of their password (M=421, SE=.07) was less than the total fixation duration of participants in the area of interests which were parts of their password (M=1256, SE=.101). And there was a significant main effect of whether related AOI was part of the password in the total fixation duration, F (1, 61) =116.5, p<.05.

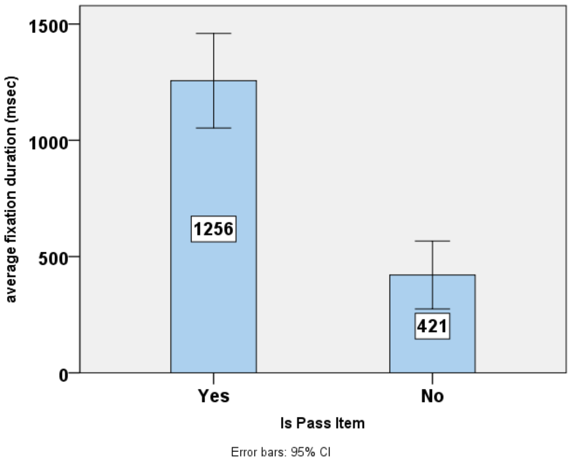


Figure 4-19 Mean average total fixation durations on pass-items vs. non pass-items on the black grid in the first session.

In the second session

58 of 63 participants' total fixation duration were calculated. In order to meet normality and the homogeneity of variances assumptions, reciprocal transformation was made. Hence, there were some zero values, 1 was added to each value before calculating the inversion value. Since ANOVA is generally robust to violations as long as sample sizes in each group (29 per each group in our study) are equal and not unreasonable small (<5), even though AOI which are not the parts of password violated normality assumption, two-way mixed ANOVA in which whether it is part of the password factor is used as repeated-measure variable and group type is used as between group variable was performed to investigate the effect of group type in total fixation duration on the black grid between the area of interests which were parts of their password and not parts of it in the second session.

Total fixation duration of participants in the area of interests which were not parts of their password (M=375, SE=.07) was less than the total fixation duration of participants in the area of interests which were parts of their password (M=1292, SE=.13). And there was a significant main effect of the whether AOI was part of the password in the total fixation duration, F (1, 56) =142.3, p<.001.

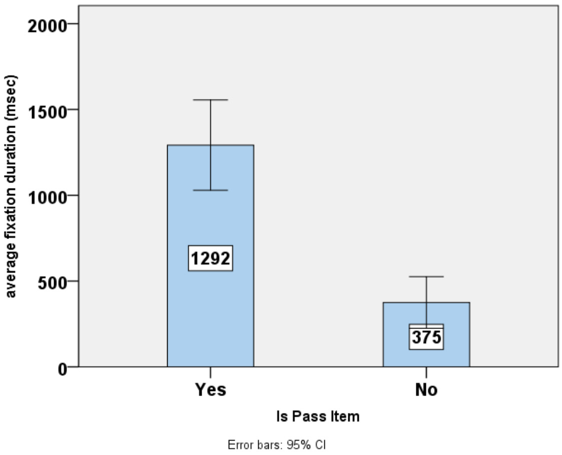
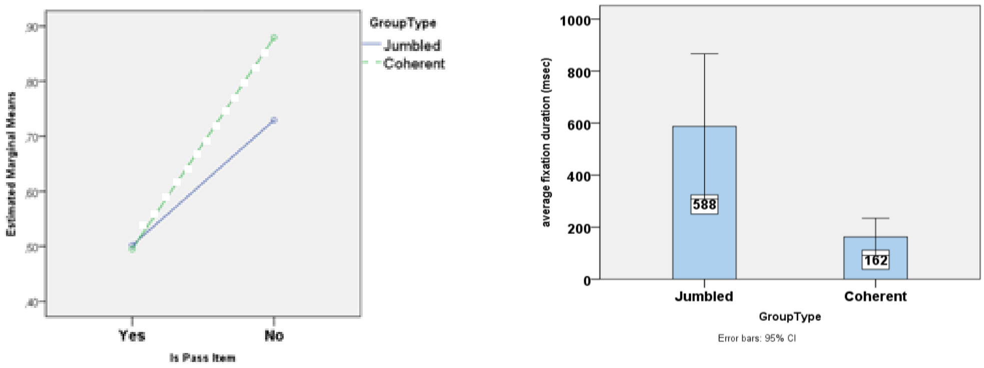


Figure 4-20 Mean average total fixation durations on pass-items vs. non pass-items on the black grid in the second session.

There was a significant interaction effect between group type and whether AOI was part of the password on the black grid in the second session, F (1, 56) =9.38, p<.05. Estimated marginal means and t-tests were used to determine the nature of this interaction. The graph and the means clearly shows that total fixation duration on the AOI which are the parts of password are very similar for both the coherent-image (M=1354, SE=.20) and jumbled-image group (M=1230, SE=.17). And this difference was not significant, t (56) =-.214, p>.05. On the other hand, total fixation duration of the coherent-image group (M=162, SE=.034) on the AOI which are not pass-items are less than the jumbled-image group (M=588, SE=.14). And this difference was significant, t (56) =-3.06, p<.05.



(a) Total fixation durations on the AOI that are pass-items are very similar for both coherent and jumbled image groups, whereas AOI which are not password items are differ.

b) Mean average total fixation durations on the AOI which are not pass-items, in the second session.

Figure 4-21 Average total fixation durations between group type and whether AOI was part of the password on the black grid in the second session

Participants in the coherent-image group had less total fixation duration (M=758, SE=.13) than the jumbled-image group (M=909, SE=.12). And this difference was significant, F (1, 56) =4.70, p<.05.

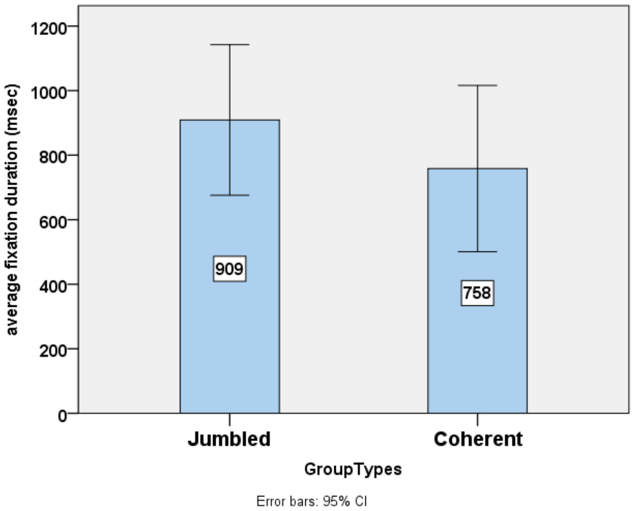


Figure 4-22 Mean average total fixation durations of jumbled-image and coherent-image groups.

In summary, in the second session, participants in the coherent-image group (M=5172, SE=.46) had less total fixation duration on the black grid than the participants in the jumbled-image group (M=7725, SE=1.18). And this difference was significant t (56) =2.01, p<.05. Independently from the session, participants in the coherent-image group (M=5480, SE=.35) had less total fixation duration on the black grid than the participants in the jumbled-image group (M=7634, SE=.77). And this difference was significant F (1, 56) =4.10, p<.05.

In the first session, there was a significant main effect of whether related AOI was part of the password in the total fixation duration, F (1, 61) =116.5, p<.05. Also, in the second session, there was a significant main effect of whether related AOI was part of the password in the total fixation duration, F (1, 56) =9.38, p<.05. In the second session, total fixation duration of the coherent-image group (M=162, SE=.03) on the AOI which are not password items are less than the jumbled-image group (M=588, SE=.14). And this difference was significant, t (56) =-3.06, p<.05.

Levenshtein String Distance

Login attempts on the black grid

Fifty eight of 63 participants' Levenshtein String Distance (LSD) concerning login attempts on the black grid was calculated. In order to meet normality and the homogeneity of variances assumptions, logarithmic transformation was made. After that, to remove and reduce the impact of the outliers, those two values were changed with the mean plus two standard deviations. Two-way mixed ANOVA in which session factor is used as repeated-measure variable and group type is used as between group variable was performed to investigate the effect of group type in LSD on the black grid for each session.

LSD of participants in the first session (M=8.22, SE=.94) was more than the LSD participants in the second session (M=6.97, SE=1.08). And there was a significant main effect of the session in the LSD, F (1, 56) =9.69, p<.05.

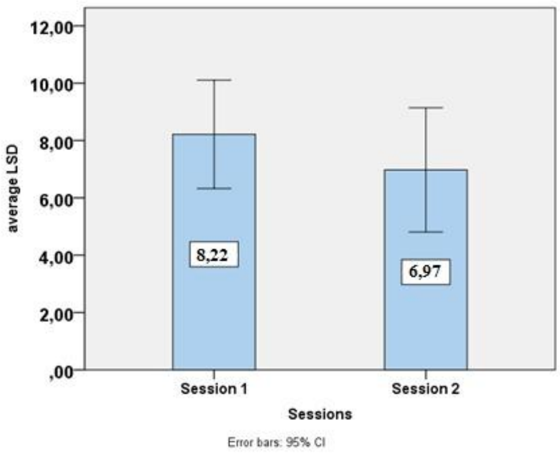
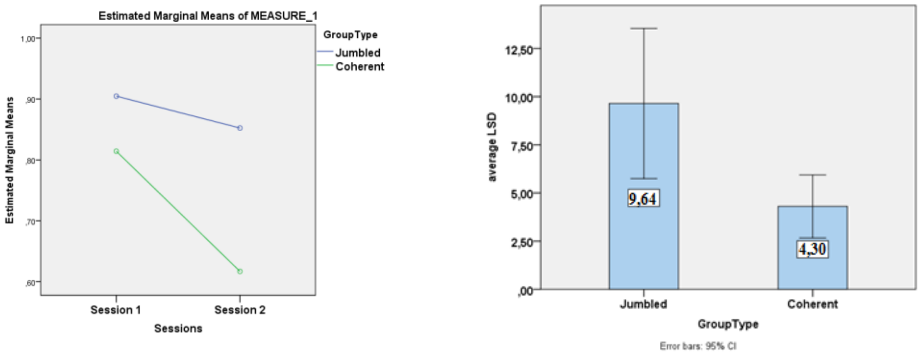


Figure 4-23 Average LSD of participants in the first and second session.

There was not a significant interaction effect between group type and the session, F (1, 56) =3.27, p>.05. To break down this interaction, t-tests were performed comparing fixation counts in each session across jumbled-image and coherent-image group. On average, in the first session, there was not a significant difference between the coherent-image group (M=7.08, SE=1.16) and the jumbled-image group (M=9.35, SE=1.47) in the LSD on the black grid t (56) =1.16, p>.05.

On the other hand, in the second session, participants in the coherent-image group (M=4.30, SE=.80) again had less LSD on the black grid than the participants in the jumbled-image group (M=9.64, SE=1.90). And this difference was significant t (56) =2.57, p<.05 and it did represent medium-sized effect r=.32.



(a) LSD of participants on the black grid is similar for both coherent and jumbled image groups in the first session, whereas it differs in the second session.

b) Mean average LSD of jumbled-image and coherent-image groups on the black grid, in the second session.

Figure 4-24 Average LSD of participants on the black grid.

Participants in the coherent-image group (M=5.69, SE=.72) had less LSD on the black grid than the participants in the jumbled-image group (M=9.49, SE=1.19). And this difference was significant F (1, 56) =4.72, p<.05.

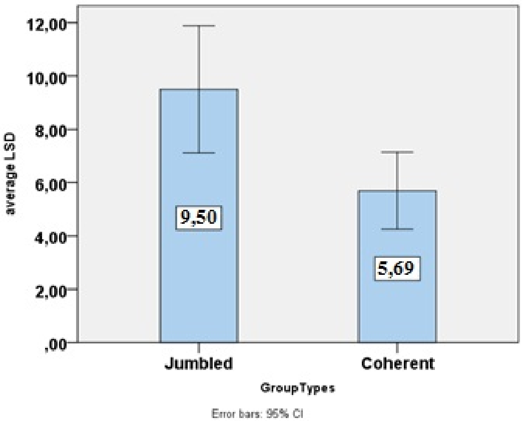


Figure 4-25 Average LSD of jumbled-image and coherent image group.

Password Creation

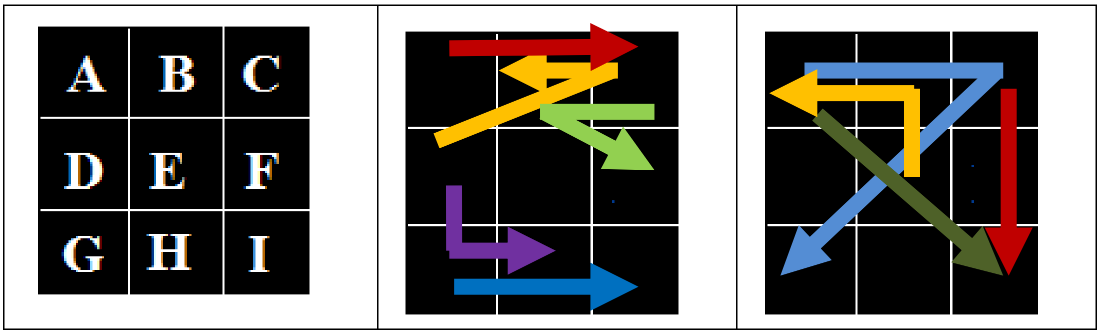
All 63 participants' Levenshtein String Distance (LSD) concerning password creation scenes were calculated. In order to meet normality and the homogeneity of variances assumptions, square root transformation was made. After that, to remove and reduce the impact of the outliers, those two values were changed with the next highest value. T-test was performed to investigate the effect of group type in LSD that generated during password creation. There was not a significant difference between the coherent-image group (M=33.55, SE=3.84) and the jumbled-image group (M=39.57, SE=4.31) in the LSD on the password creation screens, t (61) =1.22, p>.05.

In summary, in the login task, LSD of participants in the first session (M=8.22, SE=.94) was more than the LSD of participants in the second session (M=6.97, SE=1.08). And there was a significant main effect of the session in the LSD, F (1, 56) =9.69. Participants in the coherent-image group (M=5.69, SE=.72) had less LSD on the black grid than the participants in the jumbled-image group (M=9.49, SE=1.19). And this difference was significant F (1, 56) =4.72, p<.05.

However, there was not a difference between the coherent-image group and the jumbled-image group in the LSD on the password creation screen, t (61) =1.22, p>.05

4.3.4 Discovering patterns

By using eyePatterns application, participant's similar password choices were discovered with the parameters: minimum number of sequences is 2 and length of pattern 3 or 4. There was not any pattern with the length 4, which means each participant chose passwords that sorted in different locations. Letter code of locations and discovered patterns in groups is shown in Figure 4-26 and just below it; ratio of seen sequences is given in the Table 4-5.



(a) Letter code of locations

(b)Discovered patterns in the Jumbled-image group

(c) Discovered patterns in the Coherent-image group

Figure 4-26 Letter code of locations and discovered patterns in groups

Table 4-5 Ratio of password sequences selected at least by two participants. In the jumbled image group 5 patterns and in the coherent-image group 4 patterns are detected. In the table, overall ratios of each pattern is also given and ratios of each pattern for each groups are calculated by dividing observed count by the total numbers of participants

		Jumbled-image	**Coherent-image**	**Overall**
Patterns	**GHI**	3 (%4.76)		4 (%6.35)
	ACG		3 (%4.76)	3 (%4.76)
	DGH	2 (%3.17)		3 (%4.76)
	EBA		2 (%3.17)	2 (%3.17)
	CFI		2 (%3.17)	2 (%3.17)
	AEI		2 (%3.17)	2 (%3.17)
	DBF			2 (%3.17)
	ABC	2 (%3.17)		2 (%3.17)
	CBF	2 (%3.17)		2 (%3.17)
	DCB	2 (%3.17)		2 (%3.17)

Hierarchical clustering graph of passwords is given in Figure 4-27.

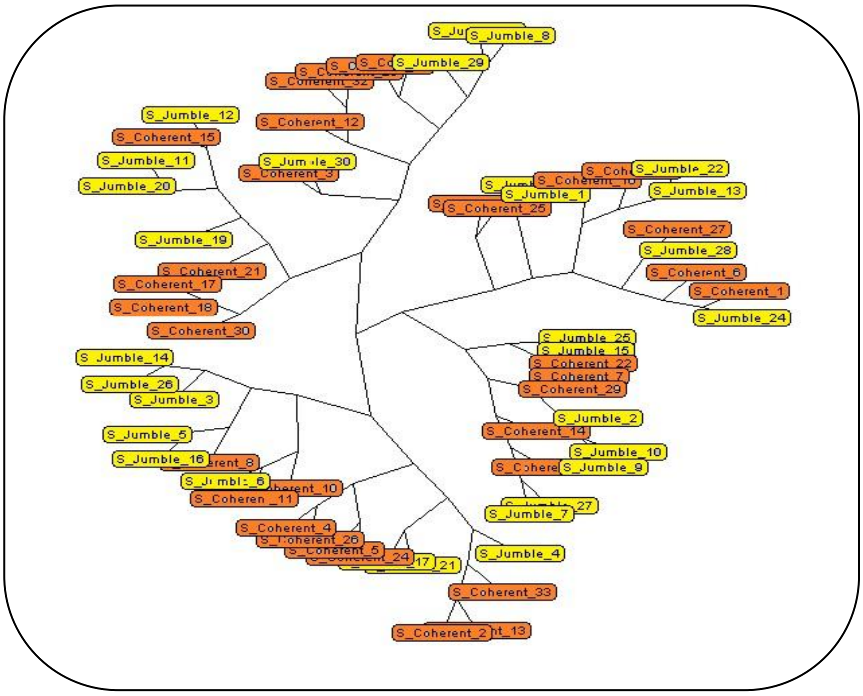
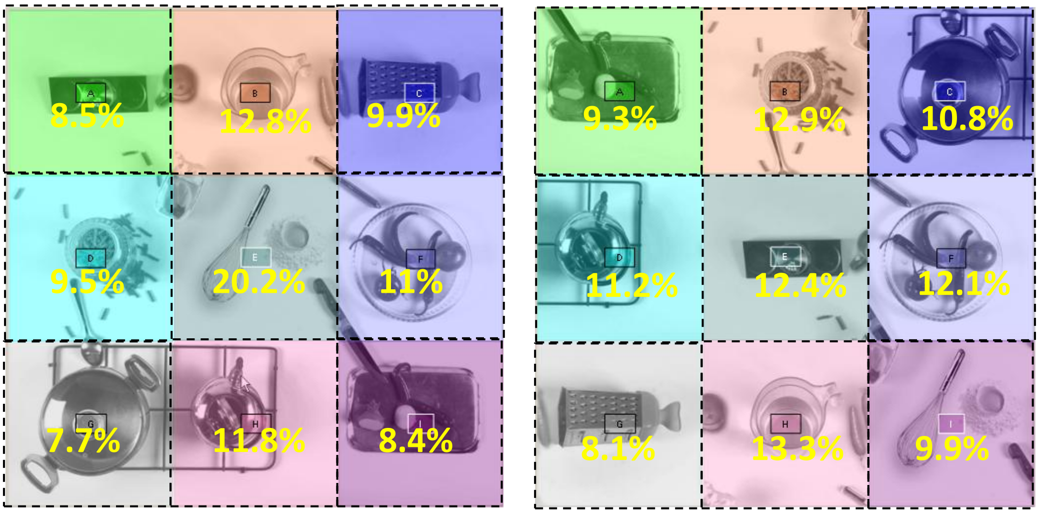


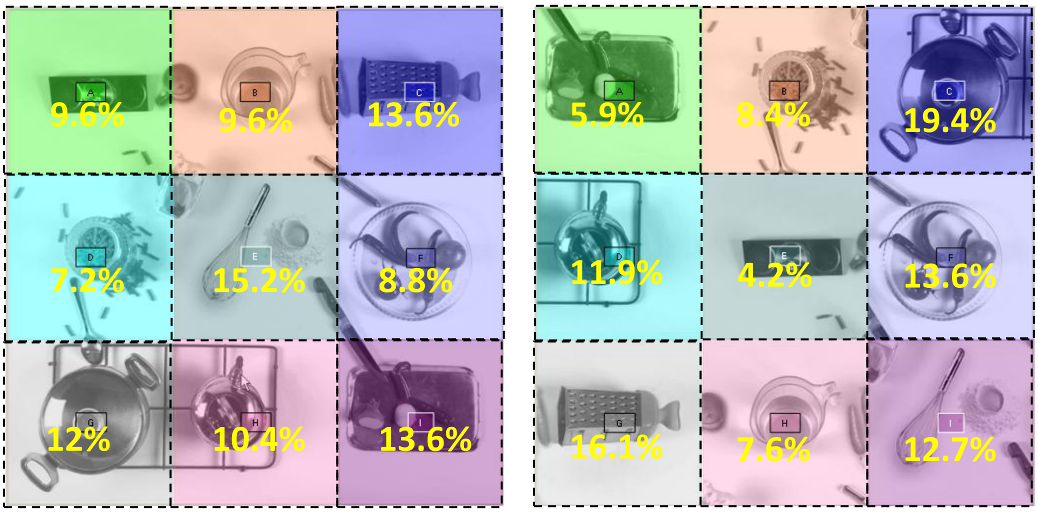
Figure 4-27 Clustering graph of passwords. Yellow nodes represent the passwords generated by participants in the jumbled-image group and the orange ones represent the passwords generated by participants in the coherent-image group. If there are fewer branches between 2 sequences, it can be concluded that those sequences are more similar.

In addition to that, each 9 item was examined in terms of how many times and in which order it was used as parts of the participant's password. The ratios of how many times they were picked as the pass-items are given in the Figure 4-28 and Figure 4-29.



(a) Coherent-image group (b) Jumbled-image group

Figure 4-28 Ratios of how many times each pieces were looked at.



(c) Coherent-image group (d) Jumbled-image group

Figure 4-29 Ratios of how many times each pieces were picked as the password item.

The number that each AOI is used as pass-item and in which order it was used as parts of the participant’s password is given in the Figure 4-30.

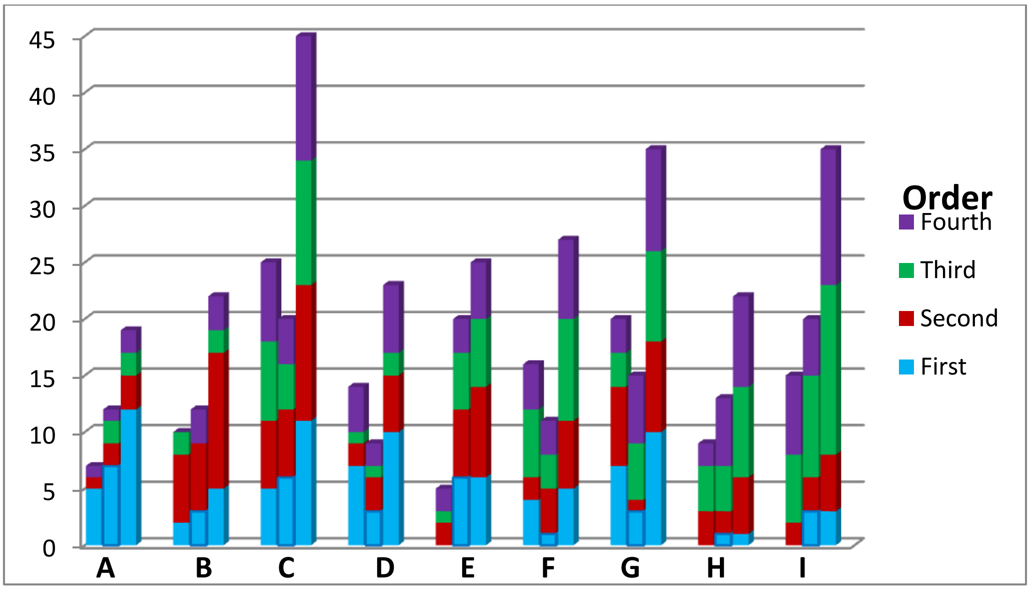


Figure 4-30 The number that each AOI is used as pass-item
In each AOI (A to I), bars from left to right represent usage of the related AOI in the passwords chosen by the participants in the either jumbled-image group, coherent-image group or both groups. In addition, colors in the bar represent the order in which the related AOI used as pass-item.

4.3.5 Analysis of the collected data by means of the questionnaire

Ratio of password creation strategies used by participants is shown in the Figure 4-31 regardless of the group type.

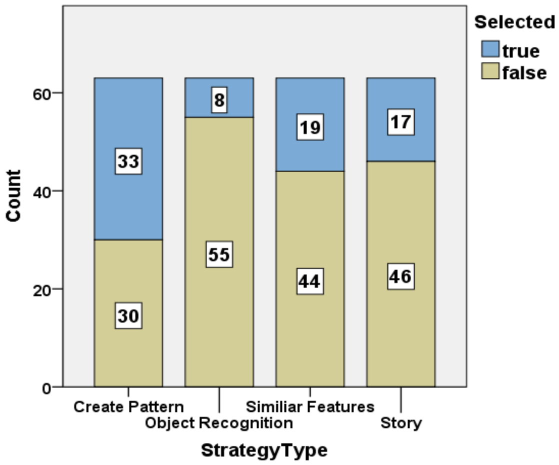


Figure 4-31 The number that each password-creation strategy is whether selected or not.

In the first session, for each group (coherent-image or jumbled-image), number of the chosen password strategies and total number of successful login attempts of the participants used related strategy is given in the Figure 4-32.

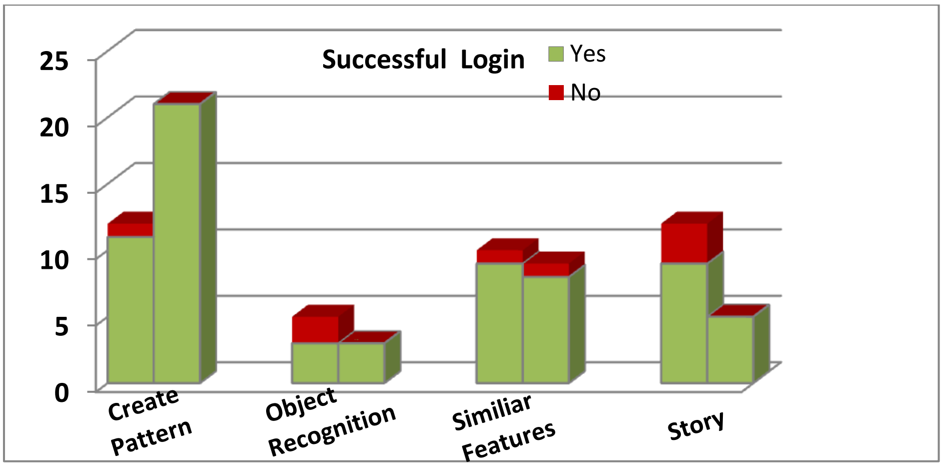


Figure 4-32 Number of successful login attempts for the specific strategy, in the first session In each password-creation strategy, bars from left to right represent number of chosen strategy in the either jumbled-image group or coherent-image group. In addition, colors in the bar represent login success of participants who chose the related strategy, in the first session.

In the second session, for each group (coherent-image or jumbled-image), number of the chosen password strategies and total number of successful login attempts of the participants used related strategy is given in the Figure 4-33.

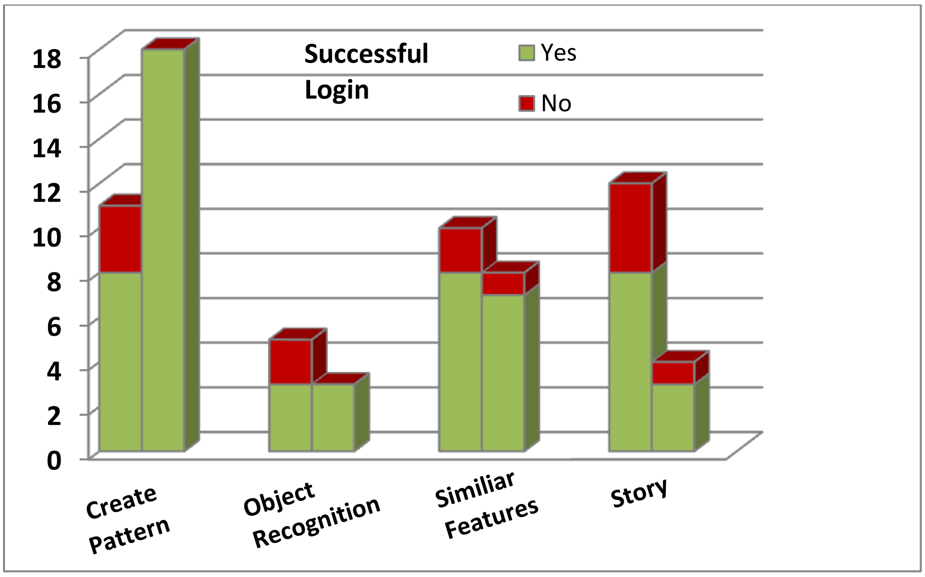


Figure 4-33 Number of successful login attempts for the specific strategy, in the second session In each password-creation strategy, bars from left to right represent number of chosen strategy in the either jumbled-image group or coherent-image group. In addition, colors in the bar represent login success of participants who chose the related strategy, in the second session.

In the below tables and figure, participant's perception in the usability of the given graphical passwords schemes is given.

Table 4-6 Participants' idea on ease of creating graphical password.

		Percentage
Create Easily	Completely Disagree	3.17%
	Disagree	4.76%
	Maybe	9.52%
	Agree	47.6%
	Completely Agree	34.9%

Table 4-7. Participants' idea on ease of remembering graphical password.

		Percentage
Remember Easily	Completely Disagree	1.59%
	Disagree	4.76%
	Maybe	23.8%
	Agree	39.7%
	Completely Agree	30.2%

Table 4-8 Participants' idea on usability of graphical password.

		Percentage
More Usable	Completely Disagree	1.59%
	Disagree	9.52%
	Maybe	49.2%
	Agree	31.8%
	Completely Agree	7.94%

In the below clustered and stacked column charts, participant's perception in the usability of the given graphical passwords schemes is shown on the basis of groups.

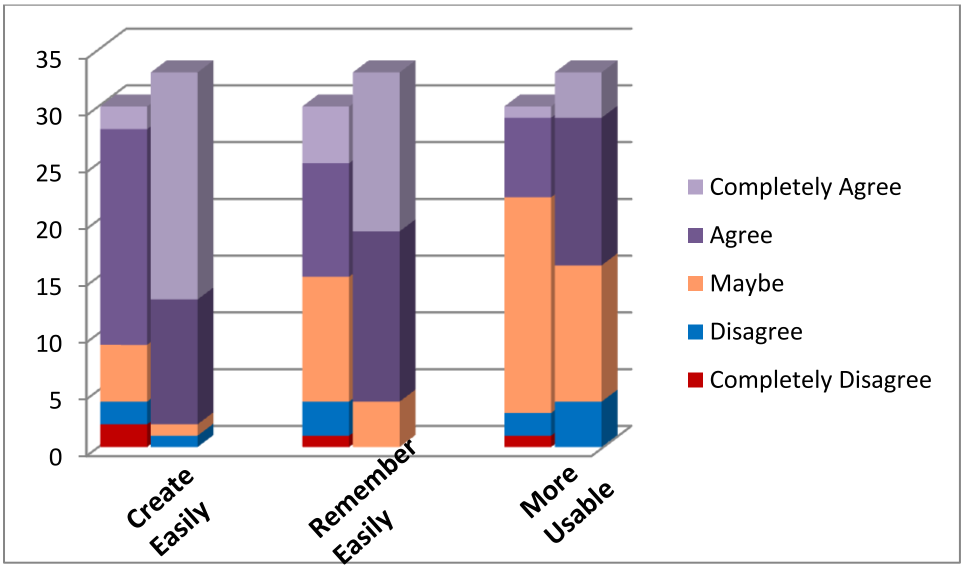


Figure 4-34 Participants' perception in the usability
Bars from left to right represents choices of participants who is in the either jumbled-image group or coherent-image group.

CHAPTER 5

DISCUSSION AND CONCLUSION

In this chapter, the study and its findings are summarized. Then, different aspects of the results are discussed in more detail. Lastly, conclusions are drawn on the basis of the results while stating the limitations of the study and relating to potential future research.

5.1 Summary of the study

Maintaining information security, most of the time, is not considered and is not part of the major goal. There is a bidirectional relation between usability and security. The design of the system influence users' behaviours and behaviours of users can potentially affect the security of the system (Adams et al., 1999). Providing both usability and security is a difficult task. It was observed that text-based authentication systems are vulnerable to attacks mainly because of the limited capacity of humans. Graphical password systems, in which passwords are created on images instead of alphanumeric characters, have been proposed as an alternative authentication method to provide users with usable yet secure systems.

The usability and security of the click-based graphical passwords, a kind of cued-recall system in which a pixel based image acts as a cue to help trigger the user's memory, were investigated (e.g., Dirik et al. 2007; Salehi-Abari et al. 2008; Chiasson et al., 2009; van Oorschot, & Thorpe, 2011). However, the studies that systematically investigate the properties (in particular, coherence) of the provided image in the proposed graphical password scheme in terms of usability and security have been lacking to date.

The studies that investigate the scene memory are generally conducted on the functions, such as change blindness and visual search that take place in interaction with scene memory (Hollingworth, 2009; Desimone, & Duncan, 1995; Hollingworth et al., 2008; Rensink, 2000; Simons, 1997). However, password notion is different than both visual search and change blindness.

The studies (i.e., Biederman, 1972; Biederman et al., 1973) reported that the objects in a *coherent image* were identified and recognized more quickly and efficiently than the objects in a *jumbled image*. In addition, effect of inter-object relations which is a spatially coherent relationship between the arrangements of objects in the scene, was studied in the recognition domain (Hock et al., 1978). It was reported that image with the familiar inter-object relations facilitate object recognition rather than the image with the novel inter-object relations. Moreover, Brockmole et al. (2006) reported that changes in the global context have support in favor of *better retrieval cue* rather than the changes in the local context that is a subset of spatially and semantically related objects. In the light of these findings, the study was designed to examine the differences in usability and security of the graphical password scheme by taking into account the coherence of the displayed image.

A desktop software application, written in Microsoft C# programming language and build with .Net Framework 3.5, was developed. Two pilot studies, one follow up study and the main experiment were conducted on different versions of the developed application to investigate the research questions.

The eye-tracking technology was also used to measure the fixation count, total fixation duration and to calculate the Levenshtein String Distance that represents the similarity between two strings (i.e., password string vs. string created during the first section of the login task and password string vs. string created during password creation).

In the first pilot study, 20 participants (10 females and 10 males) conducted the study on their own computers on which the developed application was installed. The participants were presented either a bedroom image or a set of predefined, separate items in the bedroom image. In the coherent image group, the participants picked their passwords by clicking four pieces on the bedroom image and in the separated image group the participants picked their passwords by clicking on four separate items. The study consisted of two sessions. The first session was divided into three phases: *practice, password generation* and *retention.* The second session was conducted after one week and in that session, the participants tried to log in to the system with their user names and passwords. They were allowed to try three times until they recalled their password. Otherwise, they were evaluated as failed to access to the system. By means of the application and questionnaires, demographic and task specific information were collected from the participants. The results of this pilot study revealed that coherence in images may not contribute to memorability of graphical passwords, which was not predicted by the hypothesis. On the other hand, a set of technical and methodological problems were identified in the first pilot study, the software interface and the experimental design was revised accordingly and a second pilot study was conducted.

The goal of the second pilot study was to explore design issues that were relevant to data collection in different modalities (e.g., eye tracking) and to resolve a set of technical issues that were observed in the first pilot study. In this pilot study, in addition to the obtained data from application and questionnaires, eye movement data (fixation count, total fixation duration) were obtained by eye-tracker. In this pilot study, Biederman's (1972) method was used to disrupt scene; and his proposed terms (i.e., jumbled and coherent) were adopted to describe the image types. A kitchen image was presented to the coherent-image group and the jumbled version of the image was presented to the jumbled-image group. Initially, four participants attended the second pilot study. A further follow up study was conducted by six participants in order to test the developed applications, provided images, questionnaires and to measure the average completion time of the tasks before starting to conduct the main experiment. Based on the lessons learned in the pilot studies, the main experiment was designed. For the design of the main experiment, professional support was received in order to have the photos taken under sufficient light conditions and ensure that there is at least one object which can be named in each cell. Then, it was converted in to gray scale to reduce the effect of visual saliency may cause different effects between groups by color contrast.

Sixty-three participants (29 females, 34 males) participated in the main experiment. Participants were familiar to Windows 7 logon screen. In order to simulate the Windows 7 logon screen interface, the application was modified by using Windows 7 logon screen background images. The experiment consisted of two sessions. The second session was performed 3 days after the first session. Fifty-nine of 63 participants participated in the second session. In the first session, each participant created a graphical password and in the second session s/he tried to remember it in order to authenticate to the system. The login phase consisted of two sections, firstly the participant was asked to re-select his/her password on the black grid by supposing that the image containing the cells s/he picked as his/her pass-items was covered with a black blanket. After three unsuccessful login attempts in the black-grid-screen, in the second section of the login process, the participant was provided with an image depending on his/her group type and wanted to pick his/her password. The participants were divided into groups according to the type of image they were presented in the password creation (jumbled vs. coherent image) and in the second section of the login tasks (jumbled, shuffled jumbled, coherent, shuffled coherent image).

5.2 Discussion

In the main experiment, login process consisted of two sections. In the first section, participants were asked to pick their password on black grid consist of 9 cells by supposing that the image containing the pieces they picked as their pass-items was covered with a black blanket. After three unsuccessful login attempts in the black-grid-screen, in the second section of the login process, the participant was provided with an image depending on his/her group type and wanted to pick his/her password. In the first session, 57 of 63 participants successfully logged in to the system by picking their password on the black grid without seeing the image they chose their password. Similarly, in the second session 52 of 59 participants (4 participants did not attend to the second session) did not need to see the image to remember their password. Therefore, even participants were divided into groups according to the type of image they were presented in the password creation (jumbled vs. coherent image) and in the second section of the login tasks (jumbled, shuffled jumbled, coherent, shuffled coherent image), analyses were generally performed only on two groups; either jumbled or coherent instead of the four groups depending on the presented image in the second section of the login task. In this section results that are obtained from the analysis are summarized and discussed.

Firstly, participants' login success on the black-grid screen was calculated. The results demonstrate that participants in the coherent-image group are more successful than the participants in the jumbled-image group in both of the first login attempts on the black-grid-screen and the overall login attempts on the black-grid-screen. This finding is consistent with the previous studies of Biederman (1972) and Biederman et al. (1973) in which it was reported that objects in the coherent image was recognized and identified more efficiently and quickly than the jumbled one.

Secondly, the analyses of the elapsed time to create and confirm password and the time to login were conducted. The results indicate that there was not a significant difference between groups in the time spent for password creation and password confirmation. On the other hand, in the second session, there was a significant difference between jumbled-image and coherent-image group in the time taken to login. Participants in the jumbled-image group spent more time during login than the participants in the coherent-image group. In the analysis of the login success, it was observed that participants in the jumbled-image group had some password retention problems which affect their success ratio on the black-grid-screen. Findings of the analysis of the time taken to login supported the retention difficulties of the participants in the jumbled-image group. This finding is consistent with the finding of

Foulsham et al. (2011) study related with visual search and stated that it took more time to focus on the target in the fine scrambled (divided into 8*16 grid and jumbled) than in the coarse-normal (divided into 4*4 grid), coarse scrambled (divided into 4*4 grid and jumbled) and fine normal (divided into 8*16 grid). Moreover, there was no difference between groups in the password creation and confirmation. Only difference between groups was observed in the time taken to login. Thus, there was not any difference in the ease of use of the application between groups.

Thirdly, to investigate the source of the password retention difficulties of the participants in the jumbled-image group and to elaborate the differences between the jumbled image group and the coherent image group in more detail, eye movement analyses were conducted. The results revealed that fixation counts of the participants in the first session were higher than the fixation counts of participants in the second session. Moreover, there was a significant main effect of the session in the fixation counts. In the second session, the participants in the coherent-image group had fewer numbers of fixations on the black grid than the participants in the jumbled-image group; this difference was significant. In both sessions, there was a significant main effect of whether the AOI was part of the password (pass-item) in the fixation counts. Finally, in the second session, while fixation counts on the AOI, which were pass-items, were very similar for both the coherent-image and the jumbled-image group, fixation counts in the coherent-image group on the AOI, which were not pass-items, were fewer than the jumbled-image group; and this difference was significant. These results demonstrate that both of the participants in the jumbled-image group and in the coherent-image group had approximately equal numbers of fixations on the parts of the pass item. However, the participants in the jumbled-image group had more fixations on the non-pass items than the participants in the coherent-image group. These findings support the results of the second analysis which demonstrate that in the second session, the participants in the jumbled-image group spend more time during login than the participants in the coherent-image group. Following the analysis of the fixation counts, average total fixation durations of the participants were calculated. Independently from the session, participants in the coherent-image group had less average total fixation duration on the black grid than the participants in the jumbled-image group. A significant difference was obtained, which stems from the difference between the groups in the second session. In addition, in both sessions, there was a significant main effect of whether AOI was part of the password (pass-item) in the average total fixation duration. In the second session, while average total fixation duration on the AOI, which were pass-items, were similar for both the coherent-image and the jumbled-image group, average total fixation duration of the coherent-image group on the

AOI, which were not pass-items, were fewer than the jumbled-image group; and this difference was significant. These findings are similar to the results obtained for fixation counts. In the last phase of the eye movement analyses, Levenshtein String Distance (LSD) was calculated to measure the distance between the password string and the string created on either in the first section of the login or in the password creation. In the password creation task, there was not a significant difference between the coherent-image group and the jumbled-image group in the LSD. However, in the login task, the LSD of participants in the first session was more than the LSD of participants in the second session with a significant main effect of the session in the LSD. The participants in the coherent-image had fewer LSD on the black grid than the participants in the jumbled-image group, which was also a significant difference. These findings were consistent with the results of the fixation counts and average total fixation duration analyses.

Fourthly, the patterns of user choices in graphical passwords were investigated to model similar participant choices. There was not any pattern with the length 4, which means each participant chose passwords that sorted in different locations. When minimum number of sequences was selected as 2 and length of pattern was selected as 3, five patterns for the jumbled-image group and four patterns for the coherent image group were observed. In the hierarchical clustering graph of passwords, there were more branches between sequences of the passwords of the participants in the jumbled-image group which means that sequences are less similar. Also, the ratios of how many times of each 9 items were picked as the password item were calculated. In the coherent-image group, the ratios of the items were in a range between 7.2% and 15.2% whereas in the jumbled-image group, it was in a range between 4.2% and 19.4%.

Lastly, the preferred strategies for choosing passwords and the perception of users in the usability of the given graphical password scheme were analyzed by means of the data collected by the questionnaire. The most frequently used strategy was pattern creation (33 of the 63) and the least frequently used one was object recognition (8 of the 63). In the jumbled-image group, the participants mostly used story and pattern creation strategies. In the first session, 11 of 12 participants, who reported that s/he used the pattern creation strategy, and 9 of 12 participants, who reported that s/he used the story strategy, were successfully logged into the system. In the second session, 8 of 11 participants (11 of 12 participants attended to the second session), who reported that s/he used the pattern creation strategy, and 7 of 11 participants (11 of 12 participants attended to the second session), who reported that s/he used the story strategy, successfully logged into the system. In the coherent-image group, the participants mostly used pattern creation strategies. In the first session, 20 of 20 participants reported that s/he used the pattern creation strategy, successfully logged into the system. In the second session, 17 of 17 participants (17 of 20 participants attended to the second session), reported that s/he used the pattern creation strategy, successfully logged in to the system. More participants in the coherent-image group agreed or completely agreed on the following issues; created password easily, would remember password easily and the suggested password scheme was more usable. However majority of participants in the jumbled-image group were not sure on those issues and selected the *maybe* choice.

5.3 Conclusion

The results of the main experiment revealed that there was an effect of the coherence of the image used in the graphical password system on usability and security issues in several aspects. In particular, the analyses of the login success revealed that the coherent-image increases the memorability of the password on the black-grid-screen which is a usability factor. It also revealed that, the vast majority of the participants successfully logged in to the system by remembering the locations of each pass-item and picking them on the black-grid-screen without seeing the image they chose their password on it. There are some studies showing that obtaining the image, presented to create password on it, facilitates to make automated attacks (e.g. Dirik et al., 2007; Thorpe, & Oorschot, 2007). If in graphical passwords, users generally memorize the location of the items instead of the contents of them, it would not be necessary to display the image during login process. Not displaying image might decrease the success of the automated attacks and makes the system more secure against the guessing, brute force and dictionary attacks. However, with the increasing prevalence of information technology in every aspect of our lives, most people have to deal with a lot of knowledge based authentication items in order to access secure systems (Sasse, & Flechai, 2005). Therefore, not displaying image might also reduce the number of successful login attempts which directly affect the memorability, a usability issue.

The results of the analyses of the elapsed time to create and confirm password revealed that there was not a difference between jumbled-image and coherent-image group in understanding the usage of application and the password creation and confirmation tasks. It indicates that, the confounding factor on those tasks did not occur, which is a desirable situation. According to usability-metrics-table (presented in the *usability aspect of the graphical password* heading), value of the *easy to use* and *easy to create* attributes of the *satisfaction*, one of three usability features, are the same for both of the jumbled-image and coherent-image group. In the second session, participants in the jumbled-image group spent more time during login than the participants in the coherent-image group. This finding provided further evidence of the password retention problems of the participants in the jumbled-image group. According to the usability-metrics-table, increasing in the required time to log in to the system limits utilization of the usage of jumbled-images in real-world practice. Utilization in the real-world-practice is an attribute of *efficiency*, which is one of the three usability features.

As stated in the Just and Carpenter (1980) study "there is no appreciable lag between what is fixated and what is processed" (i.e., eye-mind hypothesis). In addition, according to Renshaw et al. (2004), users spend more time when they have difficulty in processing the display. If conclusions of the eye movement analyses are stated in the context of these assumptions, it can be concluded that participants in the jumbled-image group had more difficulty in remembering the passwords and made more ineffective search than the participants in the coherent-image group (Goldberg, & Kotval, 1998; Kotval, & Goldberg, 1998). In addition, since participants in the coherent-image had fewer LSD on the black grid than the participants in the jumbled-image group, eye movements of the participants in the coherent-image group on the black grid was more similar to their passwords. Therefore, it can be concluded that passwords of the users' in the coherent-image group are more vulnerable to the guessing and automated attacks provided that the eye movements of the user is obtained.

Analyses of the patterns of user choices in graphical passwords produced no clear results. One possible reason is the small password space (9^4) of the given graphical password scheme. In addition, in most of the analysis, significant differences between groups were observed only in the second session. It may be due to the effect of the retention interval on the visual memory.

The first hypothesis of the study, which stated that the coherence of the image used in the graphical password system affects the memorability, is partially accepted since difference between two groups was observed only in the login attempts on the black-grid-screen. On the other hand, since most of the participants successfully logged in to the system on the black-grid-screen without seeing the image they chose their password. The second hypothesis of the study, which stated that the coherence of the image used in the graphical password system affects the used strategy to create password, is accepted since the most common strategies used by participants differ for each group. For instance, the participants in the coherent-image group generally used the pattern creation strategy whereas the participants in the jumbled-image group generally made a story in order to choose and remember their password. Since participants in the coherent-image group remembered passwords better than the participants in the jumbled-image group and generally used the pattern creation strategy, it can be concluded that the pattern creation strategy is in direct proportion to memorability. Finally, the last hypothesis of the study stated that the coherence of the image affects the time taken to login is also partially accepted since difference between groups observed only in the second session. In the second session, the participants in the jumbled-image group spent more time during login than the participants in the coherent-image group.

In the present study, coherence effect is examined under different topic, graphical passwords, and it reveals the positive effect of coherence in usability and security of graphical passwords. Indeed coherence is a topic of interdisciplinary research. For instance, in linguistic, it was discussed under the discourse coherence topic.

“The first flight to Frankfurt this morning was delayed. The second one was on time.” (Wolf, 2005). Even there are no explicit markers between two sentences; anyone can understand by adding the missing links automatically that there is a contrastive meaning between sentences. Discourse coherence provides continuity in context and meaning between texts in written or spoken communication. That is why discourse coherence is crucial for communication. It has been studied in various disciplines such as linguistics, psychology and computer science. Linguistics investigate markers providing coherence in language, computer scientists develop software in order to produce and detect discourse coherence automatically and psychologists explore the effect of coherence in cognitive processing (Louwerse, & Graesser, 2005; Wolf, 2005).

In addition to studies on the effect of coherence in language, there are also studies on selective visual attention, change blindness and visual search showing the importance of coherence (e.g., Edelman, 2008; Henderson et al., 1999; Neider, & Zelinsky, 2006; Rensink, 2001). These findings suggest that coherence is not a topic just for language or visual memory; it interacts with many aspects of human cognition

5.4 Implications of the Study

The study has implications for cognitive scientists and graphical password designers. There have been studies shown the effect of coherence in language, visual search, visual attention, and change blindness (e.g., Edelman, 2008; Henderson et al., 1999; Neider, & Zelinsky, 2006; Louwerse, & Graesser, 2005; Rensink, 2001; Wolf, 2005). From the cognitive science perspective, the present study shows that coherence has also impacts on graphical passwords.

Dirik et al.'s study (2007) has shown that in a graphical password system, the presented image has an impact on both security and usability. The results of the study would be helpful to determine the image type for the new proposed graphical password schemes. Using coherent images instead of jumbled one has more advantages in both usability and security of graphical password scheme. In addition, in the experiment, during the first phase of login, all participants, regardless of their group membership, were asked to enter their password on the black-grid-screen by supposing that the image containing the cells they picked their password was covered with a black blanket. Black-grid-screen is designed to understand participants' strategy of creating password that is whether they choose their passwords by memorizing the spatial location of the sections. Majority of participants were successfully authenticated by choosing their password on the black grid without seeing the image related to their groups. In other words, spatial index of the objects in the given graphical password scheme is enough for most of the participants to successfully authenticate. These two findings can be taken into consideration when designing new graphical password scheme.

5.5 Limitations and Future Studies

The main limitations of the study are the following. Firstly, in this study, theoretical password space of the proposed graphical password scheme is small. Small password space may limit to observe participants' password creation strategy and common patterns chosen by them. Secondly, only one image with two versions of it either coherent or jumbled was used to create password on it. However, in order to generalize findings of the present study, it is necessary to obtain similar results on different images containing different objects in it. Experiment can be repeated with different images divided more than 9 sections that each includes nameable objects. Thirdly, according to the design of the application, only the participants, who did not remember password on the black-grid-screen, were provided with the image to pick password. Since most of the participants remembered password on the black-grid-screen, collected data was not enough for the situation in which participants did not remember password on the black-grid-screen. This design issue prevented to see and compare results for both situations; logging to the system by picking the pass-items either on black-grid-screen or on the image displayed depends on the groups of the user. Furthermore, the number of participants and their demographic diversity should be as large as possible for the aim of generalizing the result of the present study. Although the number of participants was large enough to observe the difference between groups and decide whether the difference was significant, more participants would allow us to obtain different trends and occasions.

In a further study, different images with different theoretical password space, semantic context and visual properties can be used. For instance, the provided image for creating password could be colorful instead of being grayscale. Moreover, time interval can be increased and login task can be repeated in different time intervals. In addition, the effect of multiple graphical passwords can be examined. Lastly, usability aspects of the proposed graphical password scheme were emphasized more than security aspect. Detail analyses of security aspect would be needed as well.

REFERENCES

Adams, A., & Sasse, M. A. (1999). Users are not the enemy: Why users compromise security mechanisms and how to take remedial measures. *Communications of the ACM, 42*(12), (pp. 40–46).

Aglioti, S., Smania, N., Barbieri, C., & Corbetta, M. (1997). Influence of stimulus salience and attentional demands on visual search patterns in hemispatial neglect. *Brain Cogn*, *34*(3), (pp. 388-403).

Akula, S., & Devisetty, V. (2004). Image based registration and authentication system. *Proceedings of Midwest Instruction and Computing Symposium* (Vol. 4).

Aloul, F., Zahidi, S., & El-Hajj, W. (2009). Two factor authentication using mobile phones. *IEEEACS International Conference on Computer Systems and Applications*, *4*(2), (pp. 641-644). Ieee.

Alsulaiman F., & Saddik. A. E. (2006). A novel 3D graphical password schema: *In IEEE International Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems*

Angeli, A. D., Coventry, L., Johnson, G., & Coutts, M. (2003). Usability and user authentication: Pictorial passwords vs. PIN. McCabe, P.T. (Ed.), *Contemporary Ergonomics 2003*, 253-258. Taylor & Francis, London (pp. 253–258).

Antes, J. R., Penland, J. G., & Metzger, R. L. (1981). Processing global information in briefly presented pictures. *Psychological Research*, *43*(3), (pp. 277-292).

Aviv, A. J., Gibson, K., Mossop, E., Blaze, M., & Smith, J. M. (2010). Smudge Attacks on Smartphone Touch Screens. *In Usenix 4th Workshop on Offensive Technologies*. Usenix Association.

Bar, M., & Ullman, S. (1996). Spatial context in recognition. *Perception*, *25*(3), (pp. 343-352).

Barral, Claude. (2010). *Biometrics & Security: Combining Fingerprints, Smart Cards and Cryptography*. (Doctoral thesis, Ecole polytechnique fédérale de Lausanne, Lausanne, Switzerland). Retrieved from http://biblion.epfl.ch/EPFL/theses/2010/4748/EPFL_TH4748.pdf

Ben-Asher, N., Meyer, J., Möller, S., & Englert, R. (2009). An Experimental System for Studying the Tradeoff between Usability and Security 2009 *International Conference on Availability Reliability and Security*, (pp. 882-887). Ieee

Bhattacharyya, D., Ranjan, R., A, F. A., & Choi, M. (2009). Biometric authentication: A review. *Biometric Technology Today, 2(3),* (pp. 13-28). Retrieved from http://www.sersc.org/journals/IJUNESST/vol2_no3/2.pdf

Bicakci, K., Atalay, N. B., Yuceel, M., Gurbaslar, H., & Erdeniz, B. (2009). Towards Usable Solutions to Graphical Password Hotspot Problem. *33rd Annual IEEE International Computer Software and Applications Conference*, (pp. 318-323). Ieee.

Bicakci, K., Yuceel, M., Erdeniz, B., & Gurbaslar, H. (2009). Graphical Passwords as Browser Extension : Implementation and Usability Study. *Ifip International Federation For Information Processing*, (107), (pp. 15-29).

Biddle R., Chiasson S., & Oorschot, P.C. (2010). Graphical Passwords: Learning from the first twelve years. *ACM Computing Surveys* 44(4).

Biederman, I. (1972). Perceiving real-world scenes. *Science*, 177, (pp. 77–80).

Biederman, I. (1981). On the semantics of a glance at a scene. In: Kubovy M, Pomerantz JR, editors. *Perceptual organization. Hillsdale (NJ): Lawrence Erlbaum* (pp. 213—253).

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, 14, (pp. 143– 177).

Biederman, I., Glass, A. L., & Stacy, E. W. (1973). Searching for objects in real-world scenes. *Journal of Experimental Psychology*, *97*(1), (pp. 22-27).

Biederman, I., Rabinowitz, J. C., Glass, A. L., & Stacy, E. W. (1974). On the information extracted from a glance at a scene. *Journal of Experimental Psychology*, *103*(3), (pp. 597-600). Elsevier.

Birge, C. (2009). Enhancing research into usable privacy and security. *Proceedings of the*

27th ACM international conference on Design of communication, Bloomingto, (pp. 221-226). ACM.

Blonder, G., (1996). United states patent, *United States Patent* 5559961.

Boyce, S. J. Pollatsek, A., & Rayner, K. (1989). Effect of background information on object Identification. *Journal of Experimental Psychology: Human Perception and Performance*, 15, (pp. 556-566). PubMed

Brockmole, J. R., Castelhano, M. S., & Henderson, J. M. (2006). Contextual cueing in naturalistic scenes: Global and local contexts. *Journal of experimental psychology Learning memory and cognition*, *32*(4), (pp. 699-706). Washington, DC: American Psychological Association.

Brockmole, J. R., & Henderson, J. M. (2006). Using real-world scenes as contextual cues for search. *Visual Cognition*, 13, (pp. 99–108).

Brostoff, S., & Sasse, M. (2000). Are passfaces more usable than passwords? A field trial investigation. Y. Wærn, S. McDonald, & G. Cockton, (Eds.), *Computer*, (pp. 1-20). Springer-Verlag London ltd.

Castelhano, M. S., & Henderson, J. M. (2007). Initial scene representations facilitate eye movement guidance in visual search. *Journal of Experimental Psychology: Human Perception and Performance,* 33(4), (pp. 753-763).

Chiasson, S. (2008). *Usable autentication and click-based graphical passwords*. (Doctoral thesis, Carleton University, Ottowa, Canada). Retrieved from http://hotsoft.carleton.ca/~sonia/content/Chiasson_PhDThesis2008_UsableAuthentication.pdf

Chiasson, S., Biddle, R., & Van Oorschot, P. (2007). A second look at the usability of click-based graphical passwords: *In ACM Symposium on Usable Privacy and Security* (SOUPS).

Chiasson, S., Forget, A., Biddle, R., & Van Oorschot, P. C. (2008). Influencing users towards better passwords: persuasive cued click-points. *Technology, (Hci),* (pp. 121-130). British Computer Society.

Chiasson, S., Forget, A., Biddle, R., & van Oorschot, P. C. (2009). User interface design affects security: Patterns in click-based graphical passwords. *International Journal of Information Security, Springer* 8, 6, (pp. 387-398).

Chiasson, S., Van Oorschot, P. C., & Biddle, R. (2006). A usability study and critique of two password managers. *Symposium A Quarterly Journal In Modern Foreign Literatures* (pp. 1-16). USENIX Association.

Chiasson, S., Van Oorschot, P., & Biddle, R. (2007). Graphical Password Authentication Using Cued Click Points. J. Biskup & J. López, (Eds.), *In European Symposium On Research In Computer Security (ESORICS),LNCS 4734*, (pp. 359-374). Springer Berlin Heidelberg.

Chun, M. M. (2003). Scene perception and memory. In D. Irwin and B. Ross (Eds.) *Psychology of Learning and Motivation: Advances in Research and Theory: Cognitive Vision, Vol.* 42 (pp. 79-108). Academic Press, San Diego, CA.

Coventry, L. (2005). Usable Biometrics. In L. Cranor and S. Garfinkel (Eds.) *Security and Usability :Designing Secure Systems that People Can Use*. (pp. 175-198), O'Reilly.

Craik, F., & McDowd, J. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13, 3, (pp. 474–479).

Cranor, L. F., & Garfinkel, S. (2005). *Security and Usability: Designing Secure Systems That People Can Use*. L. Cranor & S. Garfinkel, (Eds.), O'Reilly.

Davis, D., Monrose, F., & Reiter, M. K. (2004). On user choice in graphical password schemes. *In Proceedings of the 6th USENIX Security Symposium ,Vol.* 21,(p. 11). Usenix Association.

De Angeli, A., Coventry, L., Johnson, G., & Renaud, K. (2005). Is a picture really worth a thousand words? Exploring the feasibility of graphical authentication systems. *International Journal of Human-Computer Studies*, *63*(1-2), (pp. 128-152). Elsevier.

De Graef, P., Christiaens, D., & d'Ydewalle, G. (1990). Perceptual effects of scene context on object identification. *Psychological Research*, *52*(4), (pp. 317-329). Springer.

Derawi, M. O., Nickel, C., Bours, P., & Busch, C. (2010). Unobtrusive User-Authentication on Mobile Phones using Biometric Gait Recognition. *Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing,* (pp. *306–311). IEEE.*

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, (pp. 193-222).

Dhamija, R., & Perrig, A. (2000). Déjà Vu: a user study using images for authentication.

SSYM00 Proceedings of the 9th conference on USENIX Security Symposium, 4. Usenix Association.

Dirik, A., Menon, N., & Birget, J. (2007). Modeling user choice in the Passpoints graphical password scheme. *In 3rd ACM Symposium on Usable Privacy and Security (SOUPS).*

Dunphy, P., Nicholson, J., & Olivier, P. (2008). Securing passfaces for description. *Proceedings of the 4th symposium on Usable privacy and security SOUPS 08*, 24. ACM Press.

Dunphy, P., & Yan, J. (2007). Do background images improve "draw a secret" graphical passwords? *Proceedings of the 14th ACM conference on Computer and communications security CCS 07* (p. 36). ACM Press

Edelman, S. (2008). *Computing the mind: how the mind really works*. New York: Oxford University Press.

Epstein, R., Harris, A., Stanley, D., & Kanwisher, N. (1999). The parahippocampal place area: Recognition, navigation, or encoding. *Neuron,* 23, (pp. 115–125).

Florencio, D., & Herley, C. (2007). A large-scale study of web password habits. P. F. Patel Schneider & P. Shenoy, (Eds.). *Proceedings of the 16th international conference on World Wide Web (WWW) 07*,20(3), (pp. 657). ACM Press

Foulsham, T., Alan, R. & Kingstone, A. (2011). Scrambled eyes? Disrupting scene structure impedes focal processing and increases bottom-up guidance. *Attention, Perception and Psychophysics,* 73 (7), (pp. 2008-2025).

Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, 108, (pp. 316-355).

Gao, H., Guo, X., Chen, X., Wang, L., & Liu, X. (2008). YAGP: Yet Another Graphical Password Strategy. *Annual Computer Security Applications Conference ACSAC*, (pp. 121-129). Ieee.

Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: Exploring mechanisms for face recognition. *Vision Research,* 37, (pp. 1673–1682).

Gaw, S., & Felten, E. W. (2006). Password management strategies for online accounts. *Proceedings of the second symposium on Usable privacy and security SOUPS 06*, (pp. 44). ACM Press

Goldberg, J. H., & Kotval, X. P. (1998). Eye movement-based evaluation of the computer interface. In: S. K. Kumar (ed.), *Advances in Occupational Ergonomics and Safety* (pp. 529-532). Amsterdam: ISO Press.

Goldberg, J., Hagman, J., & Sazawal, V. (2002). Doodling our way to better authentication. *CHI 02 extended abstracts on Human factors in computing systems CHI 02,* (pp. 868). ACM Press.

Green, C., & Hummel, J. (2004). Functional Interactions Affect Object Detection in Non Scene Displays. *Proceedings of the 26th Annual Meeting of the Cognitive Science Society Vol.* 1, (pp. 488-493).

Green, C., & Hummel, J. E. (2006). Familiar interacting object pairs are perceptually grouped. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(5), (pp. 1107-1119). Elsevier.

Grudin, J. (1992). Utility and usability: research issues and development contexts. *Interacting with Computers*, *4*(2), (pp. 209-217). Elsevier.

Hafız, M. D., Abdullah, A. H., Ithnin, N., & Mammi, H. K. (2008). Towards Identifying Usability and Security Features of Graphical Password in Knowledge Based Authentication Technique. (2008) *Second Asia International Conference on Modelling Simulation (AMS)*, (pp. 396-403). Ieee

Hayashi, E., Dhamija, R., Christin, N., & Perrig, A. (2008). Use Your Illusion: Secure Authentication Usable Anywhere. *Proceedings of the 4th symposium on Usable privacy and security*, (pp. 35-45). ACM.

Heckle, R. R., Patrick, A. S., & Ozok, A. (2007). Perception and acceptance of fingerprint biometric technology. *Informing Science*, (July), (pp. 153-154). ACM Press

Henderson, J. M. (1992). Object identification in context: the visual processing of natural scenes. *Canadian Journal Of Psychology*, *46*(3), (pp. 319-341).

Henderson, J. M., & Ferreira, F. (2004). Scene perception for psycholinguists. In J. M. Henderson and F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 1-58). New York: Psychology Press.

Henderson, J. M., & Hollingworth, A. (1999). High level scene perception. *Annual Review of Psychology*, 50, (pp. 243-271).

Henderson, J. M., Weeks, P. A. J., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology Human Perception and Performance,25*(1), (pp. 210-228). American Psychological Association.

Herley, C., Van Oorschot, P., & Patrick, A. (2009). Passwords: If We're So Smart, Why Are We Still Using Them? In R. Dingledine & P. Golle, (Eds.)*Financial Cryptography and Data Security, 5628*, (pp. 230-237). Springer.

Hintzbergen, J., Hintzbergen, K., Smulders, A., & Baars, H. (2010). *Foundations of information security: based on ISO27001 and ISO27002,* Van Haren Publishing, (pp.9-27)

Hock, H. S., Romanski, I. L., Galie, A., & Williams, G. S. (1978). Real-world schemata and scene recognition in adults and children. *Memory and Cognition*, *6*(4), (pp. 423-431). Psychonomic Society.

Hollingworth, A. (2005). Memory for object position in natural scenes. *Visual Cognition*, *12*(6), (pp. 1003-1016). Psychology Press.

Hollingworth, A. (2006). Scene and position specificity in visual memory for objects. *Journal of experimental psychology Learning memory and cognition*,*32*(1), (pp. 58-69). American Psychological Association: US.

Hollingworth, A. (2009). Two forms of scene memory guide visual search: Memory for scene context and memory for the binding of target object to scene location. Visual Cognition, 17(1), (pp. 273-291). Psychology Press.

Hollingworth, A., Richard, A. M., & Luck, S. J. (2008). Understanding the function of visual short-term memory: Transsaccadic memory, object correspondence, and gaze correction. *Journal of Experimental Psychology: General*, 137, (pp. 163-181).

Hong, D., Man, S., Hawes, B., & Mathews, M. (2004). A password scheme strongly resistant to spyware*: In Proceedings of International conference on security and management.* Las Vegas, NV

GrIDSure. 2009. *GrIDsure* corporate website. Retrieved on July 16, 2012 http://www.gridsure.com

Inglesant, P., & Sasse, M. (2010). The true cost of unusable password policies: password use in the wild. *Security, 1*, (pp. 383-392). ACM.

International Organization for Standardization (ISO) 9241-11. (1998). Ergonomic requirements for office work with visual display terminals: Guidance on Usability (Part 11), ISO, Geneva.

Jansen, W. (2003). Authenticating Users on Handheld Devices. *Proceedings of Canadian Information Technology Security Symposium*.

Jermyn, I., Mayer, A., Monrose, F., Reiter, M. K., & Rubin, A. D. (1999). The Design and Analysis of Graphical Passwords. *Proceedings of the 8th USENIX Security Symposium*, Vol. 8, (pp. 1-14). Usenix Association.

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 26, (pp. 683–702).

John, M. S. (2010). Enhancing security of Pass Points system using variable tolerance. *Knowledge Creation Diffusion Utilization*, *274*, (pp. 270-274).

Johnston, J., Eloff, J.HP., & Labuschagne, L. (2003). Security and human computer interfaces. Computers & Security 22 (8), (pp. 675-684).

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), (pp. 329-354).

Kainda, R., Fléchais, I., & Roscoe, A. W. (2010). Security and Usability: Analysis and Evaluation. *International Conference on Availability Reliability and Security*,(pp. 275-282). Ieee.

Kim, J. G., & Biederman, I. (2011). Where do objects become scenes? *Cerebral Cortex*, *21*(8), (pp. 1738-1746).

Kintsch, W. (1970). Models for free recall and recognition. In D. A. Norman (Ed.), *Models of Human Memory* (pp. 333-372). Academic Press.

Kirkpatrick, E. A. (2002). An experimental study of memory. *Psychological Review, 1*(6), (pp. 602-609).

Kotval, X. P., & Goldberg, J. H. (1998). Eye movements and interface components grouping: An evaluation method. *In: Proceedings of the 42nd Annual Meeting of the Human Factors and Ergonomics Society* (pp. 486–490). Santa Monica: Human Factors and Ergonomics Society.

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? M. Tistarelli & M. Nixon, (Eds.) *Vision Research*, *41*(25-26), (pp. 3559-3565). Elsevier.

Lampson, B. W. (2004). Computer security in the real world. *Computer*, *37*(6), (pp. 37-46). IEEE Computer Society.

Lashkari, A. H., Zakaria, O., Salleh, R., & Farmand, S. (2009). A wide-range survey on Recall-Based Graphical User Authentications algorithms based on ISO and Attack Patterns. *International Journal of Computer Science and Information Security (IJCSIS), Vol*. 6, No. 3

Lin, D., Dunphy, P., Olivier, P., & Yan, J. (2007). Graphical passwords & qualitative spatial relations. *Proceedings of the 3rd symposium on Usable privacy and security SOUPS 07,* (pp. 161). ACM Press.

Louwerse, M.M., & Graesser, A.C. (2005). Coherence in discourse. In K. Brown (Ed.), *Encyclopedia of language and linguistics,(2nd ed).* Oxford: Elsevier.

Luca, A. D. ,Denzel, M., & Hussmann, H. (2009). Look into my eyes!: can you guess my password?. *In Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS '09)*. ACM, New York, NY, USA,

Madigan, S. (1983). Picture memory. In J. C. Yuille (Ed.), *Imagery, memory, and cognition*: Essays in honor of Allan Paivio (pp. 65-89). Hillsdale, NJ: Lawrence Erlbaum Associates.

Man, S., Hong, D., & Mathews, M. (2003). A shoulder- surfing resistant graphical password scheme: *In Proceedings of international conference on security and management*. Las Vegas, NV

Matyáˇs, V., & Ríha, Z. (2002). Biometric Authentication Security and Usability. *In Proc. 6th IFIP TC6/TC11 Conf. Commun. Multimedia Security*, (pp. 227–239).

Miller, R. B. (1971). Human ease of use criteria and their tradeoffs. *IBM Technical Report* TR 00.2185. Poughkeepsie, NY: IBM Corporation.

Minsky, M. (1974). A framework for representing knowledge. (P. Winston, Ed.)*The Psychology of Computer Vision*. McGraw-Hill.

Nali, D., & Thorpe, J. (2004). Analyzing user choice in graphical passwords.*School of Computer Science Carleton University Tech Rep TR0401*, (pp. 1-6). Citeseer.

Neider, M. B., & Zelinsky, G. J. (2006). Scene context guides eye movements during visual search. Vision Research, 46(5), (pp. 614-621).

Nelson, K. E., & Kosslyn, S. M. (1976). Recognition of previously labeled or unlabeled pictures by 5-year-olds and adults. *Journal of Experimental Child Psychology*, *21*(1), (pp. 40-45).

Nicoletti, R., & Umilfft, C. (1989). Splitting visual space with attention. *Journal of Experimental Psychology: Human Perception and Performance*, 15, (pp.164-169).

Nielsen, J. (1994a). *Usability Engineering*. San Diego, Calif. Academic Press.

Nielsen, J. (1994b). Heuristic evaluation. In J. Nielsen and R. L. Mack (Eds.), *Usability Inspection Methods* (pp. 25-64). New York: John Wiley and Sons, Inc.

Nielsen, J., & Levy, J. (2003). Measuring usability: Preference vs. performance. *Communications of the ACM*, 37(4), (pp. 66-75).

Norman, D. A. (2009). When Security Gets in the Way. *interactions*, *16*(6), (pp. 60-63). ACM.

O'Gorman, L. (2003). Comparing passwords, tokens, and biometrics for user authentication. *Proceedings of the IEEE*. IEEE.

Olson, I. R., & Chun, M. M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition*, 9, (pp. 273-302).

Orozco, M., Malek, B., Eid, M., & El Saddik, A. (2006). Haptic-Based Sensible Graphical Password. *Proceedings of Virtual Concept*, *56*(7), (pp. 1-4). Citeseer.

Pering, T., Sundar, M., Light, J., & Want, R. (2003). Photographic authentication through untrusted terminals. *Ieee Pervasive Computing*. IEEE Educational Activities Department.

Paivio, A., (1971). *Imagery and Verbal Processes*. Holt, Rinheart & Winston, New York.

Paivio, A. & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology*, 5(2), (pp.176-206).

Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, *9*(4), (pp. 441-474). Elsevier.

Parker, D. B. (1998). *Fighting Computer Crime: A New Framework for Protecting Information* (p. 512). Wiley.

Parker, D. B. (2002). Toward a New Framework for Information Security. In S. Bosworth & M. Kabay (Eds.), *Computer Security Handbook* (p. 1224). John Wiley & Sons, Inc.

Partala, T. (2009). The Combined Walkthrough : Measuring Behavioral , Affective , and Cognitive Information in Usability Testing. *Measurement*, *5*(1), (pp. 21-33).

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, 44, (pp. 369-378).

Por, L. Y., Lim, X. T., Su M. T., & Kianoush F. (2008). The Design and Implementation of Background Pass-Go Scheme Towards Security Threats. *Journal of WSEAS Transactions on Information Science and Applications, Issue 5, Vol.* 6, (pp. 943-952), *SCOPUS-Cited Publication*.

Raaijmakers, J. G., & Shiffrin, R. M. (1992). Models for recall and recognition. *Annual Review of Psychology*, *43*(1), 205-234. Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA.

Renaud, K. (2005). Evaluating Authentication Mechanisms. In L. F. Cranor & S. Garfinkel (Eds.), *Security and Usability :Designing Secure Systems that People Can Use.* (pp. 103-128). O'Reilly & Associates.

Renshaw, J., Finlay, J., Tyfa, D. & Ward, R. (2004). Understanding visual influence in graph design through temporal and spatial eye movement characteristics. *Interacting with Computers,* 16: (pp. 557-578).

Rensink, R. A. (2000). Visual search for change: A probe into the nature of attentional processing. *Visual Cognition*, 7, (pp. 345-376)

Rensink, R. A. (2001). Change blindness: Implications for the nature of visual attention. *Vision and Attention* (pp. 169-188). Springer.

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, (pp. 1157-1165).

Riley, S. (2006). Password security: what users know and what they actually do. *Usability News, 8*(1), (pp. 2833-2836). ACM.

Salehi-Abari, A., Thorpe, J., & van Oorschot, P. (2008). On purely automated attacks and click-based graphical passwords. *In Annual Computer Security Applications Conf. (ACSAC).*

Sasse, M., Brostoff, S., & Weirich, D. (2001). Transforming the Weakest Link - A Human Computer Interaction Approach to Usable and Effective Security. *BT Technology Journal*, *19*(3), (pp. 122-131). Springer

Sasse, M. A., & Flechai, I. (2005). Usable Security. In L. Cranor and S. Garfinkel (Eds.) *Security and Usability :Designing Secure Systems that People Can Use*. O'Reilly, (pp. 13-31)

Schneier, B. (2004). *Secrets and Lies: Digital Security in a Networked World*, Wiley.

Simon, W. L. (2002). Controlling the Human Element of Security. *Current Opinion in Neurobiology*, (pp. 368), Wiley Publishing.

Simons, D. J., & Levin, T. (1997). Change blindness. *Trends in Cognitive Sciences*, *1*(7), (pp. 261-267). Elsevier.

Spiro, R. J. (1977). Remembering information from text: The state of schema approach. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 137-165). Hillsdale, NJ : Lawrence Erlbaum Associates, Publishers.

Steering Committee on the Usability, Security, and Privacy of Computer Systems, National Research Council. (2010). Current Research at the Intersection of Usability, Security, and Privacy. *Toward Better Usability, Security, and Privacy of Information Technology: Report of a Workshop*. Washington, DC: The National Academies Press

Sternberg, R. J. (2003). *Cognitive theory* (3rd ed.). Belmont, CA: Thomson Wadsworth.

Stoffer, T. H. (1991). Attentional focussing and spatial stimulus-response compatibility. *Psychological Research*, 53, (pp. 127-135).

Stubblefield, A., & Simon, D. R. (2004). *Inkblot Authentication.Test*, (MSR-TR-2004-85), (pp.1-16). Microsoft Research. Retrieved from http://research.microsoft.com/pubs/70086/tr-2004-85.pdf

Suo, X., Zhu, Y., & Owen, G. S. (2005). Graphical Passwords: A Survey. *21st Annual Computer Security Applications Conference ACSAC05*, (pp. 463-472). Ieee.

Tafasa. 2010. *Patternlock*. Retrieved on July, 2012, http://www.tafasa.com/patternlock.html.

Takada, T., & Koike, H. (2003). Awase-E: Image-based authentication for mobile phones using user's favorite images. In L. Chittaro (Ed.),*HumanComputer Interaction with Mobile Devices and Services*, Vol. 2795, (pp. 347-351). Springer Berlin, Heidelberg.

Tao, H., & Adams, C. (2008). Pass-Go: A proposal to improve the usability of graphical passwords. *International Journal of Network Security*, *7*(2), (pp. 273-292). Citeseer.

Therriault, D. J., Yaxley, R. H., & Zwaan, R. A. (2009). The role of color diagnosticity in object recognition and representation. *Cognitive Processing, 10*(4), (pp. 335-342).

Theofanos, M. F., & Pfleeger, S. L. (2011). Shouldn't All Security Be Usable? *Test*, *9*, (pp. 12-17).

Therriault, D. J., Yaxley, R. H., & Zwaan, R. A. (2009). The role of color diagnosticity in object recognition and representation. *Cognitive Processing*, *10*(4), (pp. 335-342).

Thorpe, J., & Oorschot P. C., (2004). Graphical Dictionaries and the Memorable Space of Graphical Passwords: *In Proceedings of the 13th USENIX Security Symposium*. San Deigo, USA

Thorpe, J., & Oorschot, P. C. V. (2007). Human-seeded attacks and exploiting hot-spots in graphical passwords. *Sort*, (pp. 1-16). Usenix Association.

Tobii. (2010). Tobii Eye Tracking. *Technology*. Retrieved on July 16, 2012 from http://www.tobii.com/Global/Analysis/Training/WhitePapers/Tobii_EyeTracking_Introduction_WhitePaper.pdf?epslanguage=en

Tognazzini, B. (2005). Design for Usability. In L. Cranor and S. Garfinkel (Eds.) *Security and Usability :Designing Secure Systems that People Can Use*. O'Reilly, (pp. 31-46).

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal Of Verbal Learning And Verbal Behavior*, *5*(4), (pp. 381-391). Elsevier.

Valentine, T. (1999). An evaluation of the Passface personal authentication system. *Technical Report Goldsmith College University of London.*

Van Oorschot, P. & Thorpe, J. (2011). Exploiting predictability in click-based graphical passwords. *Journal of Computer Security,* 19(4) (pp. 669-702).

Varenhorst, C. (2004). Passdoodles: A lightweight authentication method. *Research Science Institute*.

Vuilleumier, P., Henson, R. N., Driver, J., & Dolan, R. J. (2002). Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nature Neuroscience*, *5*(5), Nature America Inc.

Yokosawa, K., & Mitsumatsu, H. (2003). Does disruption of a scene impair change detection? *Journal of Vision*, *3*(1), (pp. 41-48). Association for Research in Vision and Ophthalmology.

Weinshall, D. (2006). Cognitive authentication schemes safe against spyware. (2006). *IEEE Symposium on Security and Privacy SP06*, (pp. 300). Ieee.

Wharton, C., Rieman, J., Lewis, C., & Polson, P. (1994). The cognitive walkthrough method: A practitioner's guide. In J. Nielsen and R. L. Mack (Eds.) *Usability Inspection Methods* (pp. 105-140). New York: John Wiley & Sons.

Wiedenbeck, S., Waters, J., Birget, J.-C., Brodskiy, A., & Memon, N. (2005). PassPoints: Design and longitudinal evaluation of a graphical password system. *International Journal of Human-Computer Studies*, *63*(1-2), (pp. 102-127). Academic Press, Inc.

Wiedenbeck, S., Waters, J., & Brodskiy, A. (2005). Authentication Using Graphical Passwords : Effects of Tolerance and Image Choice. *Security,* (pp. 1-12).

Wiedenbeck, S., Waters, J., Sobrado, L., & Birget, J.-C. (2006). Design and evaluation of a shoulder-surfing resistant graphical password scheme. *Proceedings of the working conference on Advanced visual interfaces AVI 06,* (pp. 177-184). ACM Press.

Wolf, F. (2005). Coherence in natural language: Data structures and applications. (Doctoral thesis, Massachusetts Institute of Technology, Cambridge, USA). Retrieved from http://dspace.mit.edu/bitstream/handle/1721.1/28854/60405129.pdf?...1

Zurko, M. E., & Simon, R. T. (1996). User-centered security. *Proceedings of the 1996 workshop on New security paradigms NSPW 96,* Vol. 1, (pp. 27-33). ACM Press

APPENDICES

Appendix A: The First Pilot Study

This appendix is composed of two parts. In the first part A.1, questionnaires of the first pilot study are presented. The second part A.2 is an appendix for the participants of the first pilot study.

A.1 Appendix to the questionnaires of the first pilot study

A.1.1 Demographic questions of the first pilot study (Before experiment) – original one is in Turkish

1. Name Surname :

2. Gender : ( ) Male ( ) Female

3. Age :

4. Job :

5. Grade-level :

A.1.2 Questionnaires for the first pilot study (After experiment) – original one is in Turkish

	Agree	Neutral	Disagree
I create password easily.			
I think that I will remember my password next week.			
Usability of this kind of password is better than numerical passwords.			
Safety of this kind of password is better than numerical passwords.			

A.1.3 Questionnaires for the first pilot study(2) (After one week retention) – original one is in Turkish

1. If you remember your password correctly,

   1.1) How do you remember your password?

   1.2) Do you remember each item obviously?

2. If you do not remember your password correctly,

   2.1) What is the possible reason?

   2.2) Do you remember any item obviously?

A.2 Appendix to the participants of the first pilot study

Table A-1 Education Level of Participants

		Number of participants
Graduate degree	High School degree	2
	Bachelor degree	9
	Master or PhD degree	9

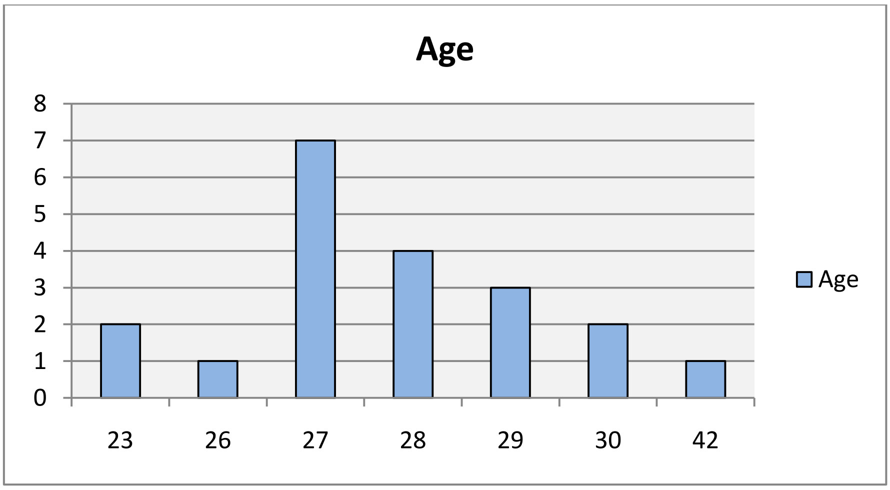


Figure A-1 Age Distribution of Participants

Appendix B: The Main Experiment

This appendix is composed of four parts. In the first part B.1, statistical information about the participants of the main experiment is given. In the second part B.2, interfaces of the application are listed. In the third part B.3, images used in the main experiment are presented and finally in the fourth part B.4, questionnaires displayed in the main application are given in English.

B.1 Appendix to the participants of the main experiment

Table B-1 Age Range of Participants

	Session 1		**Session 2**	
	Female	**Male**	**Female**	**Male**
G1	7	8	7	7
G2	7	8	7	8
G3	9	8	8	7
G4	6	10	6	9

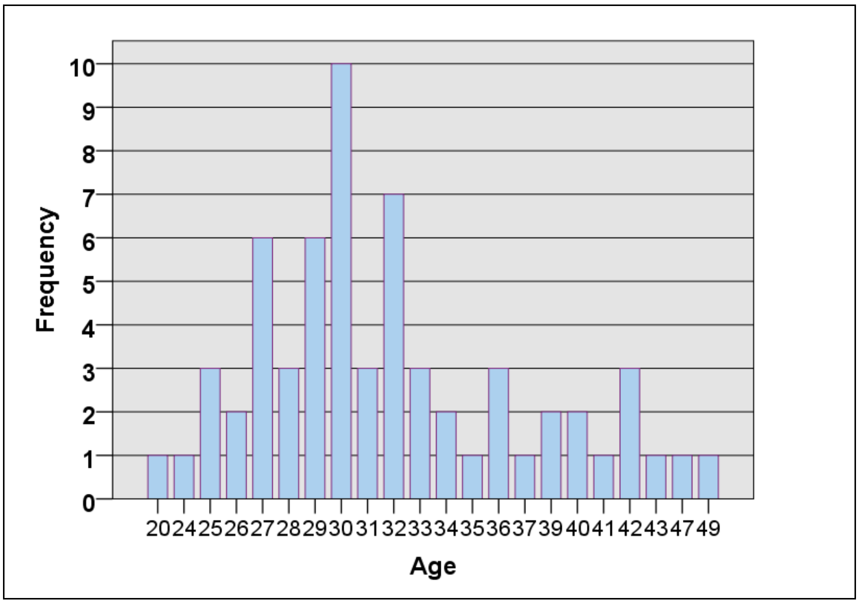


Figure B-1 Age Distribution of Participants

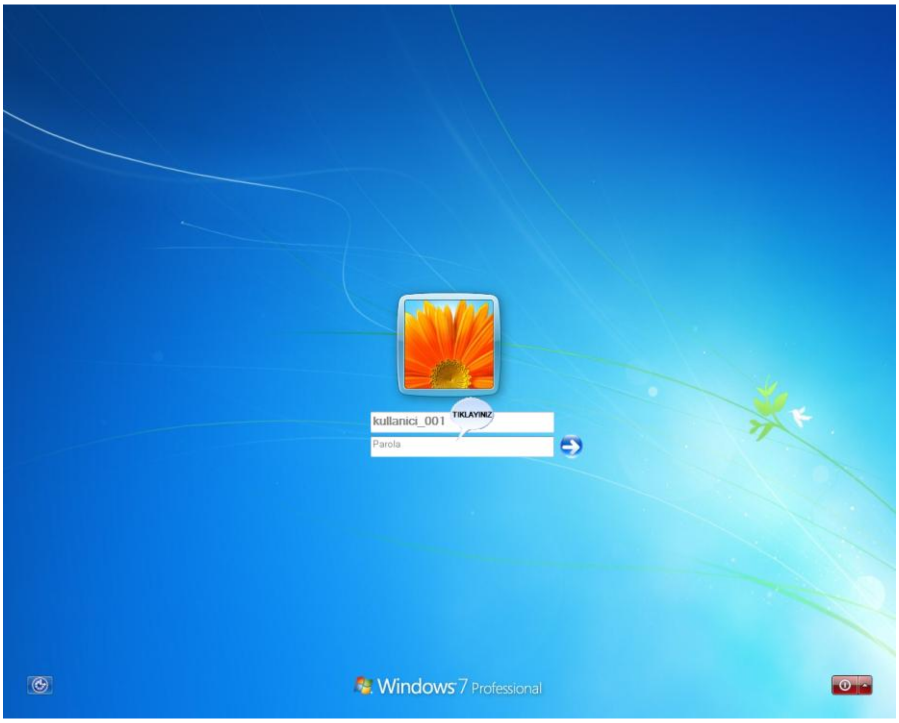
The age range of the 63 participant was 20-49 years and the mean age was 32.05 years.

Table B-2 Other informations about participants collected via a questionnaire

		G1	G2	G3	G4
Which handed	Right	14	15	15	15
	Left	1	0	2	1
Usage of eyeglasses during the study	Yes	5	6	3	4
	No	10	9	14	12

B.2 Appendix to the interfaces of the application used in the main experiment

B.2.1 Interface of the Login page

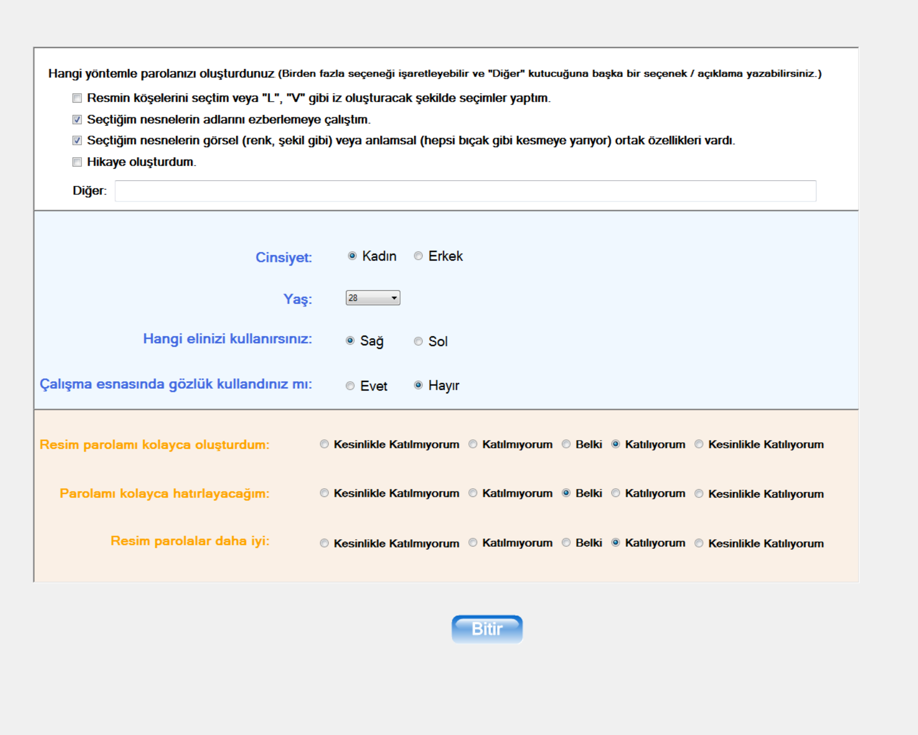


kullanici_001

Parola

Figure B-2 Login page

B.2.2 Interface of the Questionnaire page



Hangi yöntemle parolanızı oluşturdunuz (Birden fazla seçeneği işaretleyebilir ve "Diğer" kutucuğuna başka bir seçenek / açıklama yazabilirsiniz.)

- ☐ Resmin köşelerini seçtim veya "L", "V" gibi iz oluşturacak şekilde seçimler yaptım.
- ☑ Seçtiğim nesnelerin adlarını ezberlemeye çalıştım.
- ☑ Seçtiğim nesnelerin görsel (renk, şekil gibi) veya anlamsal (hepsi bıçak gibi kesmeye yarıyor) ortak özellikleri vardı.
- ☐ Hikaye oluşturdum.

Diğer:

Cinsiyet: ● Kadın ○ Erkek

Yaş: 28

Hangi elinizi kullanırsınız: ● Sağ ○ Sol

Çalışma esnasında gözlük kullandınız mı: ○ Evet ● Hayır

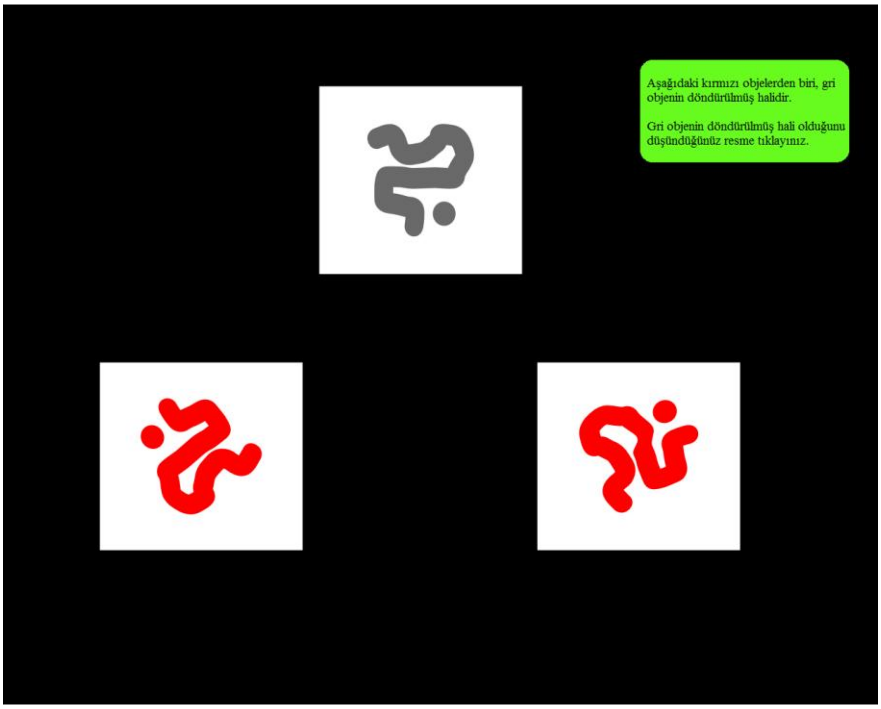
Resim parolamı kolayca oluşturdum: ○ Kesinlikle Katılmıyorum ○ Katılmıyorum ○ Belki ● Katılıyorum ○ Kesinlikle Katılıyorum

Parolamı kolayca hatırlayacağım: ○ Kesinlikle Katılmıyorum ○ Katılmıyorum ● Belki ○ Katılıyorum ○ Kesinlikle Katılıyorum

Resim parolalar daha iyi: ○ Kesinlikle Katılmıyorum ○ Katılmıyorum ○ Belki ● Katılıyorum ○ Kesinlikle Katılıyorum

Figure B-3 Questionnaire page

B.2.3 Interface of the MRT page

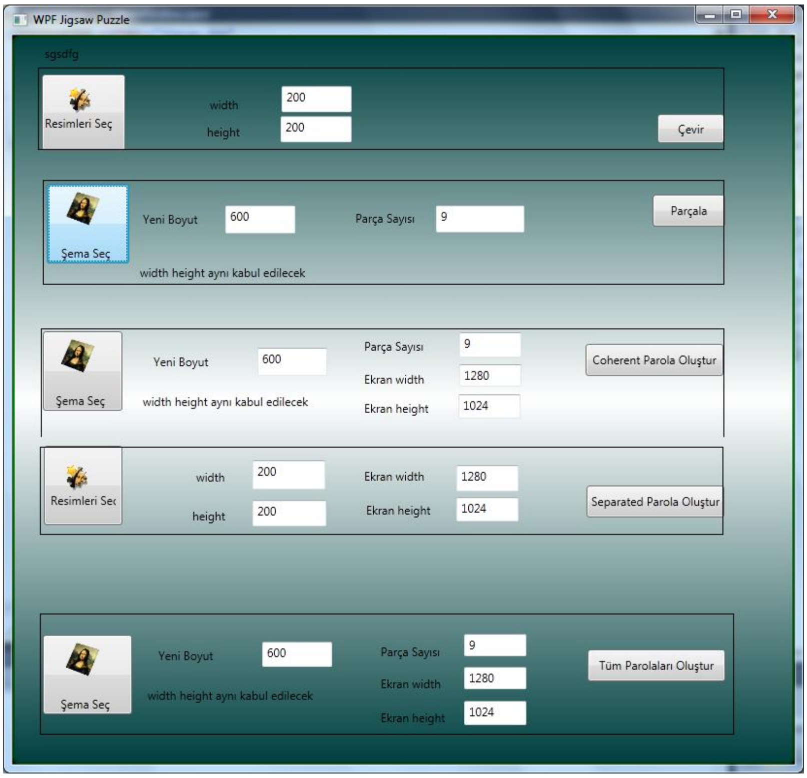


Aşağıdaki kırmızı objelerden biri, gri objenin döndürülmüş halidir.

Gri objenin döndürülmüş hali olduğunu düşündüğünüz resme tıklayınız.

Figure B-4 MRT page

B.2.4 Interface of the WPF Application



width 200
height 200

Yeni Boyut 600
Parça Sayısı 9
width height aynı kabul edilecek

Yeni Boyut 600
width height aynı kabul edilecek
Parça Sayısı 9
Ekran width 1280
Ekran height 1024

width 200
height 200
Ekran width 1280
Ekran height 1024

Yeni Boyut 600
width height aynı kabul edilecek
Parça Sayısı 9
Ekran width 1280
Ekran height 1024

B.3 Appendix to the images used in the main experiment

Test phase images;

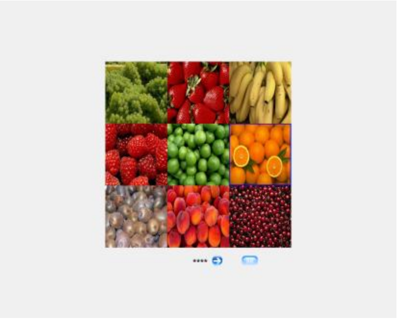
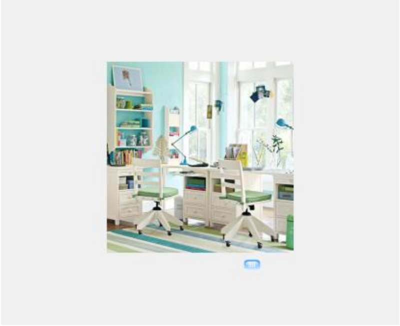
	

Figure B-5 Test phase image for G1 and G2	*Figure B-6* Test phase image for G3 and G4.

The main experiment images;

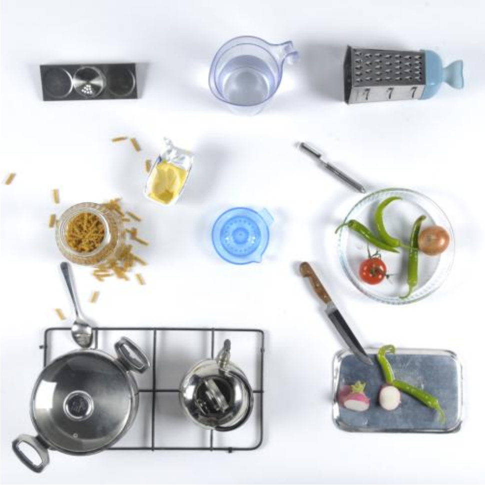


Figure B-7 The original picture

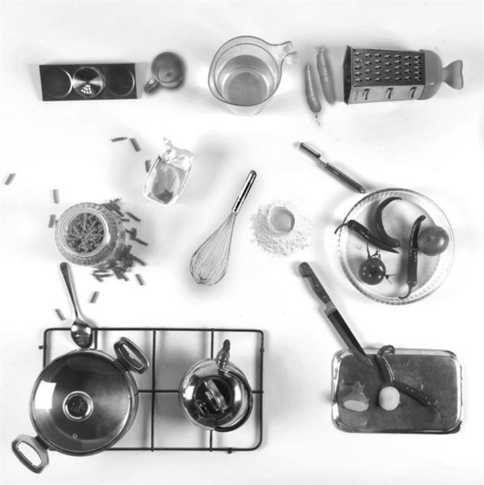
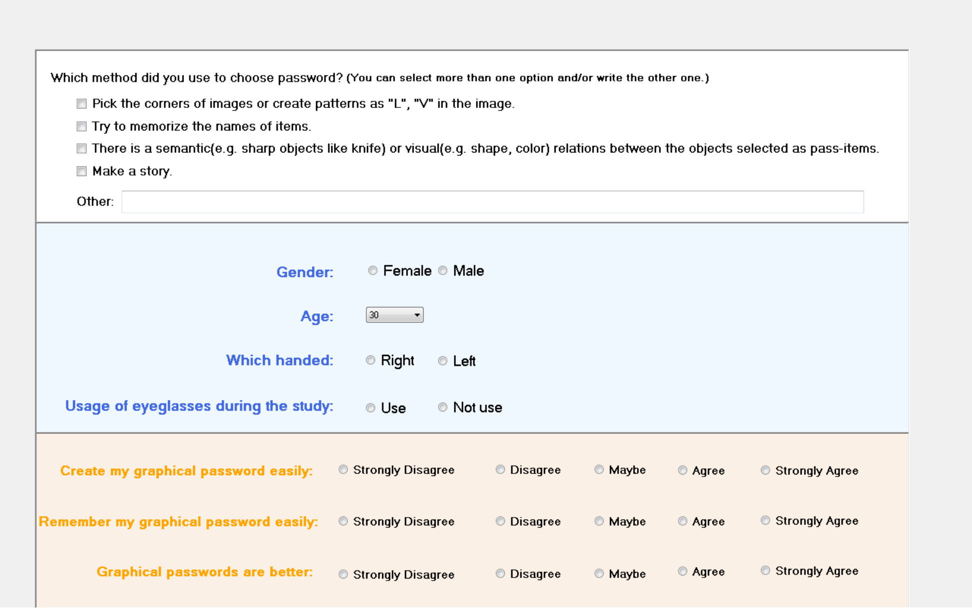


Figure B-8 Manipulated picture (converted to gray scale then modified via Photoshop)

B.4 Appendix to the questionnaires of the main experiment

It is presented after the password creation task, original one is in Turkish.



Which method did you use to choose password? (You can select more than one option and/or write the other one.)

□ Pick the corners of images or create patterns as "L", "V" in the image.

□ Try to memorize the names of items.

□ There is a semantic(e.g. sharp objects like knife) or visual(e.g. shape, color) relations between the objects selected as pass-items.

□ Make a story.

Other:

Gender: ○ Female ○ Male

Age: 30

Which handed: ○ Right ○ Left

Usage of eyeglasses during the study: ○ Use ○ Not use

Create my graphical password easily:	○ Strongly Disagree	○ Disagree	○ Maybe	○ Agree	○ Strongly Agree
Remember my graphical password easily:	○ Strongly Disagree	○ Disagree	○ Maybe	○ Agree	○ Strongly Agree
Graphical passwords are better:	○ Strongly Disagree	○ Disagree	○ Maybe	○ Agree	○ Strongly Agree

Figure B-9 Questionnaire (in English)

Appendix C: Analyses and Results of the Main Experiment

This appendix is composed of two parts. The first part C.1 is an appendix for the results of the comparison of gender in terms of login success in the black-grid-screen. In the second part C.2, AOIs of each image used in the login phase is presented.

C.1 Appendix to the analysis of gender difference in login success in the main experiment

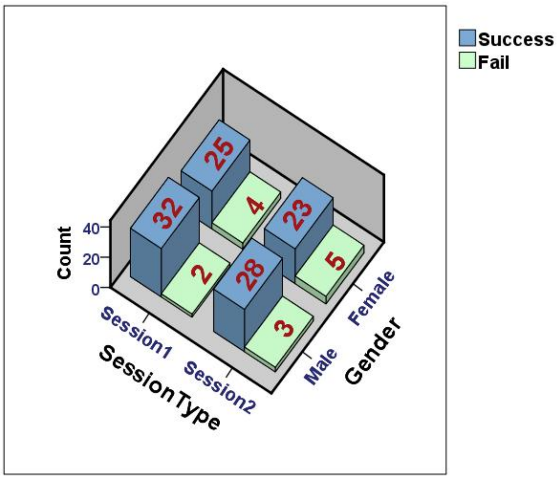


Figure C-1 Login success of users in the black-grid-screen. (Success*Gender*Session)

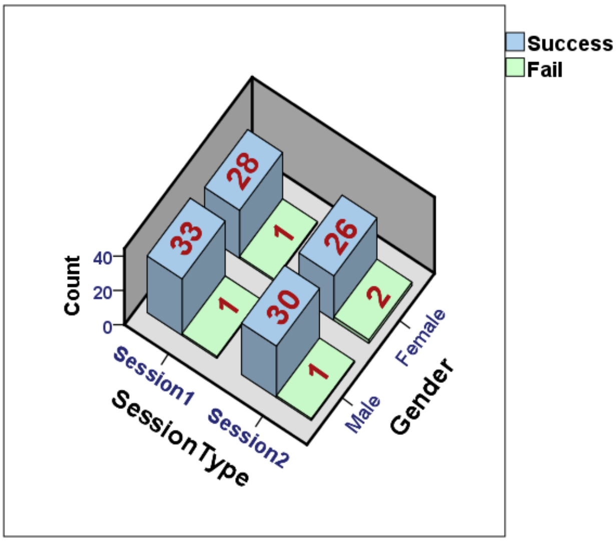


Figure C-2 Overall (i.e., in and after the black-grid-screen) login success of users. (Success*Gender*Session)

C.2 Appendix to the AOIs

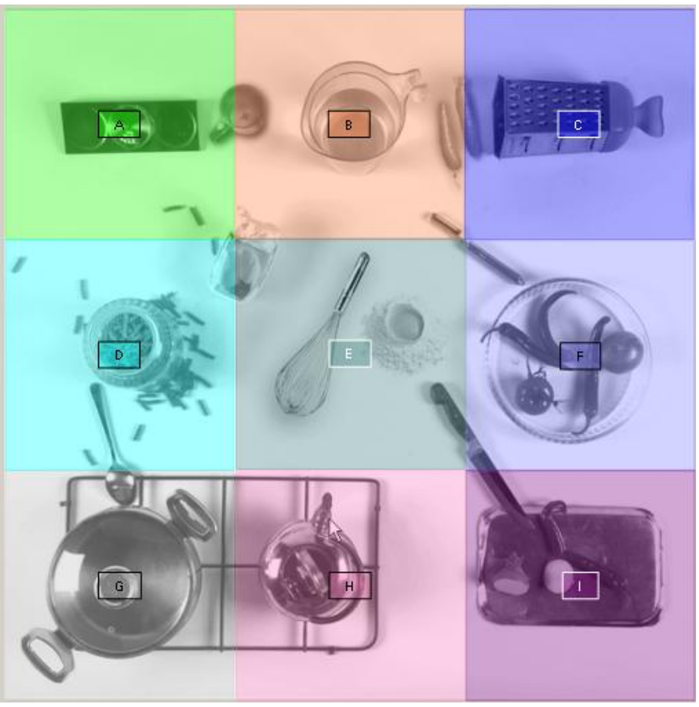


Figure C-3 Coherent-image separated in to 9 AOIs named A to I.

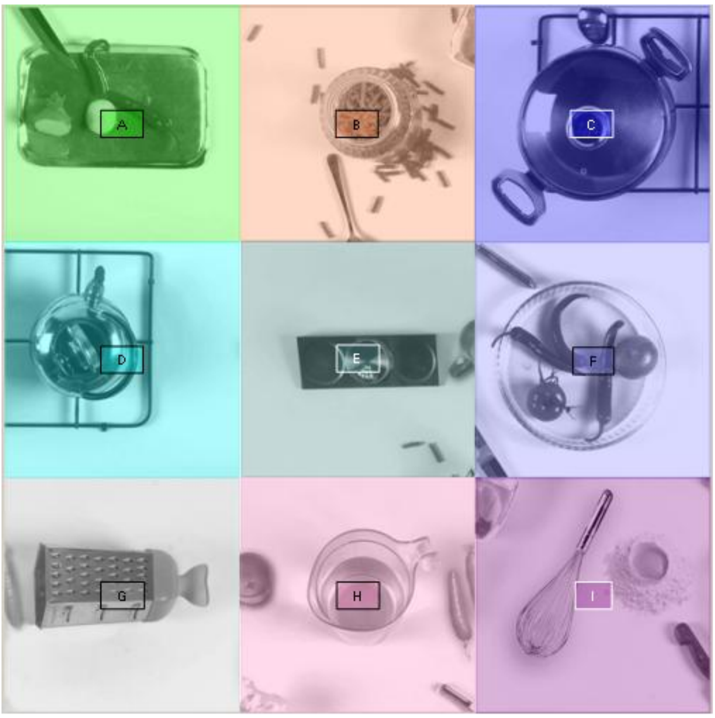


Figure C-4 Jumbled-image separated in to 9 AOIs named A to I.

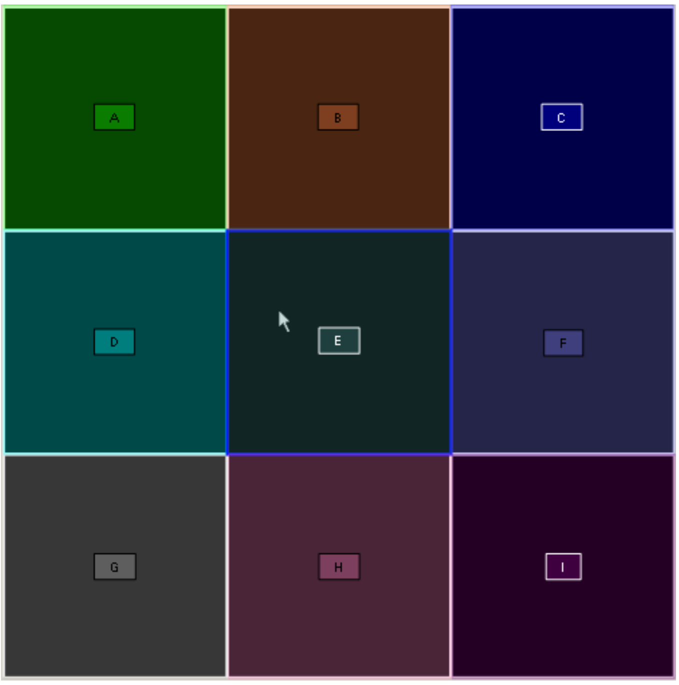


Figure C-5 Black grid separated in to 9 AOIs named A to I.

METU LIBRARY

TEZ FOTOKOPİ İZİN FORMU

ENSTİTÜ

Fen Bilimleri Enstitüsü ☐

Sosyal Bilimler Enstitüsü ☐

Uygulamalı Matematik Enstitüsü ☐

Enformatik Enstitüsü ☒

Deniz Bilimleri Enstitüsü ☐

YAZARIN

Soyadı : ARSLAN AYDIN

Adı : ÜLKÜ

Bölümü : BİLİŞSEL BİLİMLER

TEZİN ADI (İngilizce) : **THE EFFECTS OF COHERENCE OF THE IMAGE USED IN THE GRAPHICAL PASSWORD SCHEME IN TERMS OF USABILITY AND SECURITY**

TEZİN TÜRÜ : Yüksek Lisans ☒ Doktora ☐

1. Tezimin tamamı dünya çapında erişime açılsın ve kaynak gösterilmek şartıyla tezimin bir kısmı veya tamamının fotokopisi alınsın. ☐

2. Tezimin tamamı yalnızca Orta Doğu Teknik Üniversitesi kullancılarının erişimine açılsın. (Bu seçenekle tezinizin fotokopisi ya da elektronik kopyası Kütüphane aracılığı ile ODTÜ dışına dağıtılmayacaktır.) ☐

3. Tezim bir (1) yıl süreyle erişime kapalı olsun. (Bu seçenekle tezinizin fotokopisi ya da elektronik kopyası Kütüphane aracılığı ile ODTÜ dışına dağıtılmayacaktır.) ☒

Yazarın imzası Tarih: 28.08.2012