A Comprehensive Study of the Usability of Multiple Graphical Passwords

Soumyadeb Chowdhury, Ron Poet, Lewis Mackenzie

▶ **To cite this version:**

Soumyadeb Chowdhury, Ron Poet, Lewis Mackenzie. A Comprehensive Study of the Usability of Multiple Graphical Passwords. 14th International Conference on Human-Computer Interaction (INTERACT), Sep 2013, Cape Town, South Africa. pp.424-441, 10.1007/978-3-642-40477-1_26 . hal-01504899

HAL Id: hal-01504899

https://inria.hal.science/hal-01504899

Submitted on 10 Apr 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

A Comprehensive Study of the Usability of Multiple Graphical Passwords

Soumyadeb Chowdhury, Ron Poet, Lewis Mackenzie

School of Computing Science, University of Glasgow

{soumc, ron, lewis} @ dcs.gla.ac.uk

Abstract. Recognition-based graphical authentication systems (RBGSs) using images as passwords have been proposed as one potential solution to the need for more usable authentication. The rapid increase in the technologies requiring user authentication has increased the number of passwords that users have to remember. But nearly all prior work with RBGSs has studied the usability of a single password. In this paper, we present the first published comparison of the usability of multiple graphical passwords with four different image types: Mikon, doodle, art and everyday objects (food, buildings, sports etc.). A longitudinal experiment was performed with 100 participants over a period of 8 weeks, to examine the usability performance of each of the image types. The results of the study demonstrate that object images are most usable in the sense of being more memorable and less time-consuming to employ, Mikon images are close behind but doodle and art images are significantly inferior. The results of our study complement cognitive literature on the picture superiority effect, visual search process and nameability of visually complex images.

Keywords: usability, user authentication, multiple image passwords,

1 Introduction

Information security systems must permit only legitimate users to gain access to the system and use its resources. This is done by a two step process. *Identification* verifies the user's right to access the system and, once the user is identified, they have to prove their identity by *authentication*. In computer security mechanisms, people are often required to authenticate themselves by using a secret known as a *password or authenticator*. In current practice, alphanumeric passwords are the most widely used mechanism to authenticate users. According to the studies reported in [1, 2] most people find it difficult to remember these passwords. As the number of passwords per user increases, the rate of forgetting them also increases [3]. In order to cope with multiple passwords, users tend to adopt unsafe strategies, which include writing down the passwords, reusing the same passwords and sharing them with others [1, 3].

RBGSs are an alternative type of mechanism where images are used as passwords. The idea is that humans can remember images better than recalling alphanumeric text

© Springer-Verlag Berlin Heidelberg 2011

[4, 5] and so this may be a way of devising more memorable passwords. Given the need for alternative usable authentication systems and existing interest in image passwords [3, 6, 7, 8, 9, 10] as a potential solution, we identify an important limitation of existing work: most prior studies with images as passwords, except [11, 12], have focused on the usability of a single password. However, people will need to remember and use multiple image passwords in the same way that they currently use multiple alphanumeric passwords. The usability of multiple image passwords has not been explored except using faces and everyday objects (flowers, food, sculptures, nature etc.). So it is necessary to conduct experimental studies with other popular image types to compare their usability, when multiple image passwords are used. The motivation of our work is to investigate '*which image type (*s*) performs best in terms of usability, when multiple graphical passwords are used?*' We believe, unless such large-scale studies are done, image passwords will always look good (usable) on paper but fail in real life and cannot be considered as a viable alternative to alphanumeric passwords.

This paper presents a study with 100 participants, who used multiple image passwords over a period of 8 weeks. The study compares the usability of 4 image types: Mikon, doodle, art and everyday object, when used as graphical passwords. In this paper: (1) we identify the need to study and compare the usability of multiple image passwords using different image types; (2) we show that the image types used as passwords by the participants significantly affect the effectiveness as well as efficiency of RBGSs, when multiple image passwords are used; (3) We discuss the implications of our findings, regarding the attempts to record image passwords motivating a need for future studies examining the vulnerability (guessability) of these passwords to descriptions given by the users.

2 Background Work

Existing studies have reported the usability of faces [3, 7], Mikons [6, 10], doodles [6, 9], abstract art [8] and pictures of daily objects [7]; however, these studies have focused only on the use of single passwords.

We are aware of only two pieces of prior work that studied multiple image passwords: [12] studied the use of multiple facial passwords and [11] compared multiple picture passwords to that of multiple PIN's.

The work reported in [12] studied the effects of frequency of facial password usage, the effects of interference resulting from the use of multiple graphical passwords and the effects of different patterns of access while training with multiple facial passwords. The study also demonstrated that long term memorability of the multiple facial passwords issued by the system is low.

The work reported in [11] compared the memorability of multiple graphical passwords to that of multiple alphanumeric passwords. The images used in this study were photographic and were drawn from different categories such as food, music, sports etc. In the study, there was a single challenge set which contained 4 authenticator images and 6 decoy images. However, this limited set may not be secure and may

make the authentication procedure somewhat too easy. In the experimental framework, the participants used the password during training session just after completion of the registration and then three retention tests were carried out with a gap of two weeks between each of them. In such a scenario, the encoding of password information in the long term memory may not be entrenched enough to aid memorability. This is evident from the login success rate reported in the paper, which just after training is almost three times higher than the success rate after 4 weeks. We also find that passwords are issued by the system rather than chosen by the user themselves, and this may hamper the usability of the system.

Given the significant interest of researchers in image passwords, it is clear that additional work is needed using other image types to enhance knowledge of usability, when multiple image passwords are employed. Therefore we designed a usability study to explore the potential of 4 different images types (graphical passwords) Mikon, doodle, art and everyday objects with: (1) Common experimental settings including user tasks for all graphical passwords; (2) Same evaluation parameters; (3) Same design of the authentication system interfaces ; (4) Testing in both conditions i.e. passwords frequently used and less frequently used.

3 User Study overview

The main aim of the study was to compare the usability of multiple image passwords in RBGSs. A three stage study was designed, which consisted of: (1) a pre-study questionnaire evaluating the participant demographics and current password strategies; (2) an 8-week online study of participants using multiple image passwords; and, (3) a post-study questionnaire regarding participant experiences. In our usability study we used the following image types (See Appendix): (1) Mikon: These are icon-like images which have been drawn by users in studies reported in [6, 10] using a tool called the Mikon engine developed by Mikons.com. (2) Doodle: These images are drawn by a user using pen on paper, in studies reported in [6, 9]. (3) Art: These images were collected from a range of free websites and comprised of paintings from different styles like cubism, abstract, modernism etc. (4) Object: These images comprised of pictures of food and drinks, sculpture and buildings as well as sports and leisure activities, collected from a range of free websites.

3.1 Experimental conditions and Images

We used independent measures (between subjects) style of experimental design with four conditions (equal number of participants in each condition) namely Mikon (m), doodle (d), art (a) and objects (o). Each participant was assigned to only one of the conditions randomly. The participants in the Mikon condition used Mikon images as their password. They created 4 Mikon passwords and authenticated using them. Each password comprised of 4 Mikon images. So the participants had to remember 16 images in total. Similarly participants in the doodle, art and object conditions used the respective types as passwords. The participants were given a task information sheet

which contained the information on the steps to register with the system i.e. select four images to create a password and steps to login to the system after successful registration. However, the participants were neither given any instructions regarding the strategy they should use to select their password images nor the strategy they should employ to remember them.

Each condition was associated with a mock online study website that had a distinct logo and address. Each website had 4 hyperlinks (each link corresponding to one password, See Fig 1 step 1): My jokes; My movies; My news; My status. All hyper-links except My status had a collection of 150 different images of the same image type (Table 1 and Appendix). My status had a collection of 150 images (50 images each from My jokes, My movies and My news respectively). Each hyper link had a distinct name and background color. The participants could post information/content in the link upon successful authentication: This made sure that the participants had a context to use in differentiating their multiple passwords.

Table 1. Image categories in each link

	My jokes (MJ)-150	My movies (MM)-150	My news (MN)-150	My status (MS)-150
Mikon	Colorful images (No tags)	Black and white. (No tags)	Colorful as well as Black and white images with tag	50 images each, from first three links
Doodle	Black and white (No tags)	Black and white - not same as MJ (No tags)	Black and white with tags	50 images each, from first three links
Art	Abstract paintings	Paintings different from MJ (cubism/ modernism etc)	Paintings different from MJ and MM (cubism/ modernism etc)	50 images each, from first three links
Object	Food and drinks	Sculpture and buildings	Sports and leisure	50 images each, from first three links

3.2 Tasks

Registration

The online study website had 4 links. Each link would correspond to one password (made up of 4 images). The participants could register in each of the links by entering a username and selecting 4 images as their password from the given collection. Figure 1 shows the registration screens of the RBGS developed for our experiment. Each link

had a collection of 150 images, presented on the screen as six sets of 25 images in the form of 5X5 grids. The participants could browse from one set to other using the 'change set' button on the web page. The participants could choose all the 4 images from a single set or each image from a different set. The experiment was designed such that each of the participants would use a different collection of the same image type while registering for each of their passwords. For example, each participant in the Mikon condition created 4 passwords, with 1 password selected from the image collection of the first link (My jokes), one from the second link (My movies) and so on. The image archive in each link of the Mikon website consisted of the same image type (Mikon).

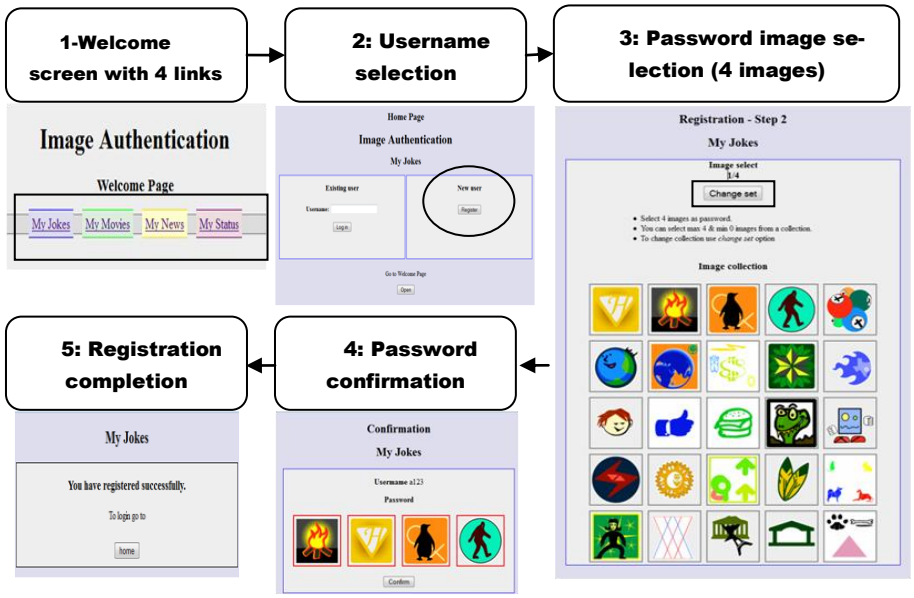


Fig. 1. Steps to register with a password in RBGS

Authentication Procedure

Upon accessing the online study website, the participants had to select the correct password images (the ones they selected during registration) from a sequence of 4X4 grids at each step of a four-step procedure (Figure 2). All the images other than the password image in the sequence are known as *decoy images*. The grid consisting of 15 decoy images and 1 password image is called a *challenge set*. The decoy images for each step were fixed during registration to ensure that the intruder would not be able to guess the correct password image, merely refreshing the web page. The decoy images for each password image were chosen randomly from the collection of 150 images. The decoy images for each of the 4 target images were distinct and never repeated. If the participants at any step during the login procedure selected wrong password image, then they would never get any of their registered password images in the subsequent steps. So in this case, 16 decoy images (without the password image), different from the origi-

nal challenge set were displayed to the users. The participants were given feedback on the result of the login only after the last step of the procedure. In case of three continuous failed login attempts, the participants were automatically reminded of their password. The reminder was given only for the first 20 login attempts (week-1) of each password. Once the participants logged in with the correct password images, they could post any information which could be seen by others using the system, who could like it.

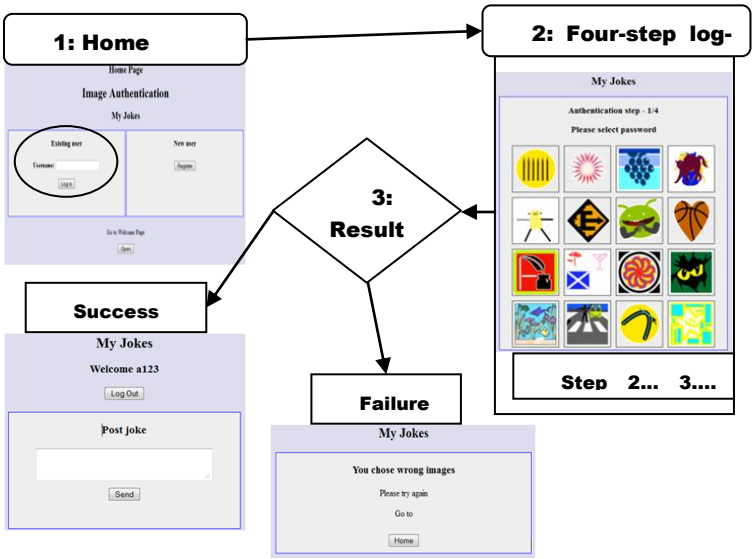


Fig. 2. Steps to authenticate in RBGS

Upon successful authentication, the main task of the participants was to post some content in each of the links. The participants were not forced to use the system and the experimental procedure was designed to allow flexibility in the tasks. All sorts of ethical approvals were taken from to conduct the experimental study.

4 Experimental Framework and Protocol

The essential component of the experimental design is the 8 week online study. We used email prompts to inform the participants about the experimental procedure and tasks. Emails were sent to participants on Day 1, 3 and 5 of each week to notify the progress that they have already made and they are expected to make, to complete the weekly tasks. These emails helped them to keep track of the tasks and made sure they followed the experimental procedures.

There is no standard procedure to design experiments for studying multiple image passwords. In our experiment we tried to vary the frequency of login at regular intervals to simulate a real life scenario. We designed the study, considering the following.

- Ecological validity: participants had a task and a context. A pre questionnaire was conducted for two reasons: (1) to decide the number of multiple graphical passwords to be used for the experiment; (2) to examine, if our participant sample has the same text password behaviour as those of other participants, reported in existing studies [1].
- Confounding variable: the authentication systems (user interface, design), tasks and frequency of login were the same for all conditions.

4.1 Instructions

The participants had to first register with the system. Each participant in each condition had to create 4 passwords, each password comprising of 4 images. Once they finished registration, they had to practice logging in using the password images. The participants were instructed as follows: Week 1: They must create 4 passwords and login 20 times with each of the passwords (total 80 logins in the week). Week 2: They must login 20 times with each of the passwords (total 80 logins in the week); Week 3-4: They must login with each of the passwords 10 times (total 40 logins in the week); Week 5: In this week participants were instructed to login with each of the passwords 2 times (total 8 logins in the week); Week 6: In this week participants were instructed to login with each of the passwords 4 times (total 16 logins in the week); Week 7: In this week participants were instructed to login with each of the passwords twice (total 8 logins in the week); Week 8: In this week participants were instructed to login with each of the passwords thrice (total 12 logins in the week); The participants were instructed to distribute their login sessions over a period of time, instead of finishing them simultaneously. The participants who did not follow the experimental procedure were left out of the experiment.

4.2 Measures – Usability criteria

We evaluated the usability of the different image types used as password in different conditions using the four measures given below:

1. Effectiveness : This examined the *average/mean successful login percentage (S)* for each of the conditions calculated as,

$$S = \frac{\textbf{Total number of successful login in the condition}}{\textbf{Total number of login in the condition}} \times 100$$

2. Efficiency: It examined the average/mean registration time (R), and average/mean login time of successful login (L). The registration time for each of the passwords is the time taken to go from screen 3 to screen 6 of the registration process as shown in Fig 1. The average registration time (R) for each condition is calculated:

$$\frac{1}{4}\sum_{i=1}^{4} \textbf{Registration time for password } (i), \qquad \textbf{where } i \textbf{ denotes password } 1, 2, 3, 4$$

The login time for a password is the time taken to go from screen 2 to the success notification screen of the authentication process shown in Fig 2. The average time of successful login (L) for each condition is calculated as given below, z represents total number of successful login.

$$\frac{1}{z}\sum_{n=1}^{z} \textbf{\textit{Login time for successful login}}\ (n)$$

3. Satisfaction: This dimension was assessed from the ratings (1- 5, 1 being highly dissatisfied to 5 being highly satisfied) given by the participants to the different aspects in the post study questionnaire- (sat1) Ease to register; (sat2) Ease to authenticate; (sat3) Meaningfulness/nameability of the image; (sat4) satisfaction with the type of image used as password. These aspects were based on some of the items in SUS (System Usability Scale) questionnaire [13].
4. Stress: This dimension was assessed from the ratings (1- 5, 1 being least stressful to 5 being highly stressful) given by the participants to the different aspects in the post study questionnaire- (str1) level of mental stress; (str2) level of physical stress; (str3) amount of effort required to choose images during registration; (str4) amount of effort required to successfully login. These aspects were based on the items in the NASA Task Load Index questionnaire [14].

5 Results

5.1 Participants

115 undergraduate participants, 30 female and 85 male, of age 20-24 took part in our experiment. They were studying different undergraduate courses: Mechanical Engineering - 22, Electrical Engineering- 19, Aerospace Engineering- 25, Computer Science- 24, Electronics and Communication Engineering- 25. Of the 115 participants who took part, 10 participants had a very low participation rate (did not follow the experimental procedure) and 5 of the participants had to withdraw due to some personal circumstances. So the participation rate was 86.9%.

5.2 Pre study questionnaire results

Given the number and quality of problems associated with alphanumeric passwords, we conducted an online survey to obtain information about password construction techniques and different issues related to them. A total of 150 participants took part in the survey, which included the 115 participants of our usability study. A web based questionnaire was used to obtain data on different aspects of user behavior and perceptions in context of the use of alphanumeric passwords. The framework of the survey was developed using Grounded theory [15]. The framework provided a step by step methodology (Fig 3): (1) identifying the key points of the survey; (2) categorizing the key points depending on the factors influencing them; (3) parameters or concepts to be examined under each category; (4) explanation of the results (qualitative data) to draw conclusions.

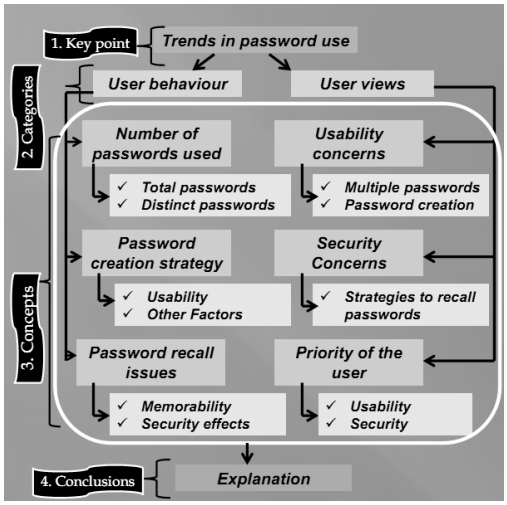


Fig. 3. Grounded theory framework for pre-study survey

The results of the survey revealed that participants used as many as 4-6 passwords in their day to day lives. This result enabled us to make the choice of using 4 graphical passwords with our sample users for the usability study. In context to the password creation strategy, of almost 600 responses: 71.2 % reported using similar passwords for all web accounts; 54.4 % used passwords that could be linked to their personal likings. The users felt that having similar passwords across all accounts aids their memorability. The results also demonstrated that 80% of the participants forget their passwords, either due to the strategy used to aid memorability, or constraints imposed by the system while creating alphanumeric passwords. The results of the survey are in line with the findings of similar research reported in [1]. Thus we can confirm that the sample population used for usability studies has the same password behaviour as reported in other studies. Hence, the sample would represent an accurate reflection of the general population.

5.3 Effectiveness results

Our first planned analysis compares the effectiveness of the different image types used as password. We analyzed the data from week 2 to week 8. We eliminated the data from week 1, as it was considered a training week, where participants would get used to the system. The dependant variable average/mean login success percentage (for 7 weeks) for each of the conditions (image types) was normally distributed as assessed by the Shapiro-Wilk test. Levene's test indicated that the assumption of homogeneity of variance has not been violated ($F(3, 96) = 2.083$, $p = 0.108$). Given the use of the independent measures (between subjects) experimental protocol with four conditions and the normal distribution of the data, we chose One-way independent

measures ANOVA to examine statistical significance. The results of the ANOVA shows that there is a statistically significant difference between the conditions F (3, 96) = 129.659, $p<0.01$. These indicate that the type of images used as password by the participants significantly affect the average successful login percentage. The effect size ($r = 0.83$, large effect) indicates that the effect of the images used as authenticator on the average successful login percent is substantial. The results of the Tukey post hoc tests revealed significant differences between all conditions ($p < 0.001$ for all tests) except between Mikon and Object ($p = 0.059$). Thus there was no significant difference between the effectiveness of Mikon and Object images.

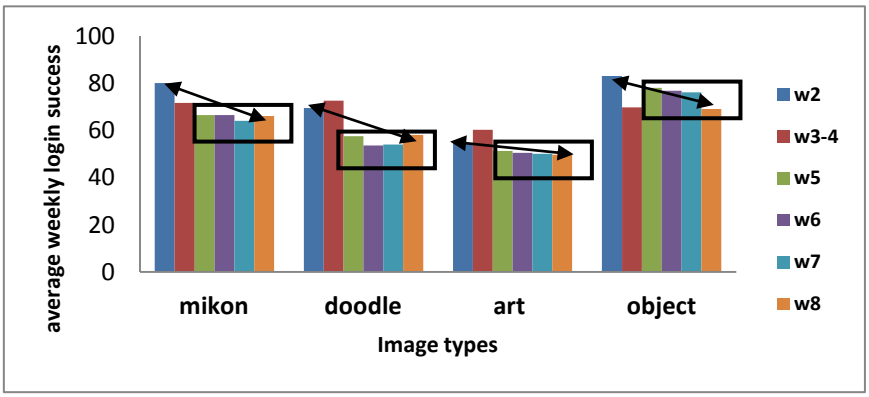


Fig. 4. Weekly login success percentage

We also analyzed the weekly login success percent for each of the image types (Fig 4). We found that the login success percentage for each of the image type falls from week 2 to week 8, as the frequency of usage of the passwords decreases. If we compare the average weekly login success percentage in week 2 and week 8 for each image type (shown by bidirectional arrows in Fig 4), it is found that they fall by 11.44 % in case of Mikon, 12.55 % in case of doodle, 7.74 % in case of art and 14% in case of object. The results of the study show that the average weekly login success percentages for Mikon, doodle and art passwords remain almost the same after week 4 (indicated using black rectangles). In case of object passwords the average weekly login success percentage remains almost same from week 5-7 but drops in week 8. Thus the decrease in memorability for the best performers i.e. Mikon and object are almost the same. Similar characteristic are shown by doodle passwords. In case of art passwords (lowest average login success percent), the difference is comparatively low which clearly suggests that people had problems remembering them in the initial as well as final stages of the study. This reflects that once the participants had used the passwords for a considerable amount of time i.e. at least for a month, memory interference does not hamper the memorability of the image passwords. The memorability of multiple image passwords after a considerable amount of time would depend on the encoding of the passwords in the human memory and frequency of their usage. A two-way ANOVA was conducted with week and image types as the two independent

variables. The dependant variable was average weekly login success percent. The week x image interaction was significant, ($F(15,576) = 6.102$, $p < 0.001$). This indicated that the average weekly login success significantly varied for each of the image types in different weeks.

5.4 Efficiency results

Efficiency is a measure of convenience: since a time consuming process will be a barrier, to the repeated use of the authentication system by the user.

Registration time

The mean registration time of the passwords (in seconds) for each condition is as follows: Mikon (mean: 72.18, SD: 5.48, SE: 1.17), doodle (mean: 75.40, SD: 4.27, SE: 0.88), art (mean: 84.44, SD: 4.91, SE: 0.99), object (mean: 70.61, SD: 3.84, SE: 0.76). We also find that the registration time decreases as users get used to the system in each of the conditions (registration time decreases from p1- first registered password to p4- last registered password) (Fig 5).

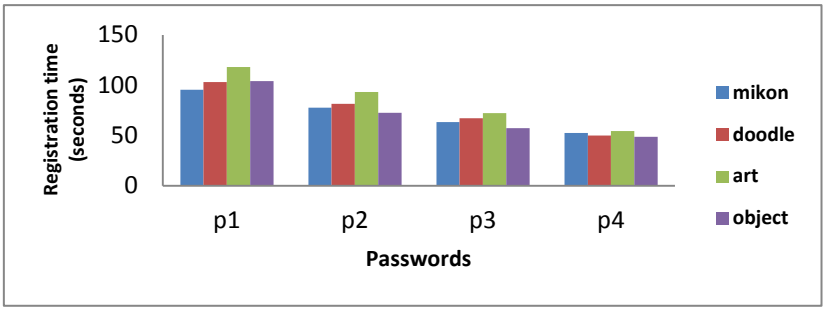


Fig. 5. Registration time for each password created in each condition

The mean registration time (of 4 passwords)) for each of the conditions (image types) was normally distributed as assessed by the *Shapiro-Wilk test*. Levene's test indicated that the assumption of homogeneity of variance has not been violated ($F(3, 96) = 1.968$, $p = 0.127$). We used One-way independent measures ANOVA as the statistical test. The result of the ANOVA shows that there is a significant difference between the conditions ($F(3, 96) = 41.277$, $p<0.001$). This indicates that the type of images used as a password by the participants affect the average registration time significantly. The effect size ($r = 0.78 > 0.5$ represents large effect) indicates that the effect of the images used as authenticator on the average registration time is substantial. The results of the *post hoc* tests revealed significant difference between all groups ($p < 0.05$) except Mikon-Doodle ($p=0.091 > 0.05$) and Mikon -Object ($p = 0.658 > 0.05$). In other words, the differences between the average registration time of Mikon and doodle as well as Mikon and object passwords is not significant.

Authentication time

The average authentication time of the successful login (in seconds) for each of the conditions is as follows: Mikon (mean: 19.52, SD: 3.60, SE: 0.72), doodle (mean: 22.16, SD: 3.75, SE: 0.75), art (mean: 24.56, SD: 4.8, SE: 0.96), object (mean: 18.28, SD: 2.84, SE: 0.59). The average login time (of 7 weeks) for each of the conditions (image types) was normally distributed as assessed by the Shapiro-Wilk test. Levene's test indicated that the assumption of homogeneity of variance has not been violated ($F(3, 96) = 1.791$, $p = 0.124$). The results of the one way ANOVA shows that there is a statistically significant difference between the conditions ($F(3, 96) = 13.199$, $p < 0.001$). This indicates that the type of images used as password by the participants affect the average login time significantly. The effect size ($r = 0.61 > 0.5$ represents large effect) indicates that the effect of the images used as password on the average login time is substantial. The results of the *post hoc* tests revealed significant differences between all groups ($p < 0.05$) except Mikon-Doodle ($p=0.091 > 0.05$) and Mikon -Object ($p = 0.658 > 0.05$). In other words, the differences between the average login time of Mikon and doodle as well as Mikon and object passwords is not significant.

5.5 Satisfaction results

According to the box plot in Fig 6, the satisfaction scores of objects range from 13-15 which is better than Mikons ranging from 12-14.5. The doodles have a range of 10.5-12 whereas art have a range of 8-10.5. Hence objects have the best satisfaction score distribution followed by Mikons. But the doodle and art images have inferior satisfaction scores.

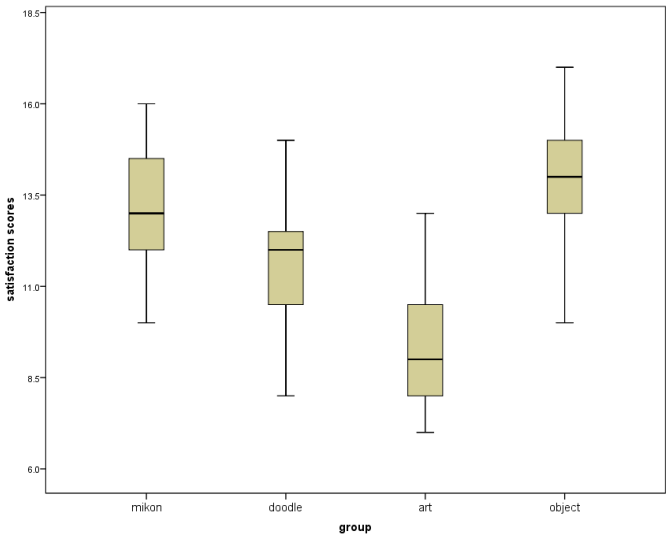


Fig. 6. Box plot for satisfaction scores

Given the ordinal scale of data (user ratings) and the independent measures (between subjects) experimental protocol with four conditions, we used Kruskal-Wallis test to establish statistical significance of data. The result shows that the satisfaction scores for each of the conditions were statistically significant [$H(3) = 52.37$, $p < 0.001$). In other words, the satisfaction of the participants was significantly affected by the type of images used as password. We conducted a Mann-Whitney test to follow up our findings by applying a Bonferroni correction, to report all the effects at a 0.008 level of significance. The results reveal that the satisfaction scores were significantly different in all conditions ($p < 0.008$ for all tests) except for the Mikon-Object ($p = 0.156 > 0.008$). So we conclude that users are most satisfied with object and Mikon passwords (no significant difference), followed by doodles and least satisfied with art passwords.

5.6 Stress results

According to the box plot in Figure 7, the stress scores of objects range from 11-13 which is same as that of Mikons. The doodles have a range of 12-14, whereas art have a range of 13-16. Hence the art images have the highest stress score distribution closely followed by doodles. The object and Mikon images have the lowest stress score distribution.

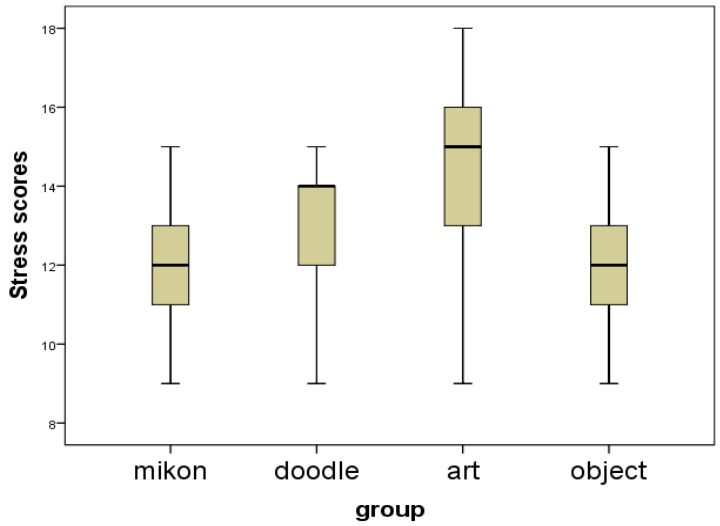


Fig. 7. Box Plot for stress scores

Given the ordinal scale of data (user ratings) and the independent measures (between subjects) experimental protocol with four conditions, we used Kruskal-Wallis test. The result demonstrates that the stress scores for each of the conditions were statistically significant [$H(3) = 23.314$, $p < 0.001$). In other words, the stress scores given by the participants were significantly affected by the type of images used as password. We conducted a Mann-Whitney test to follow up our findings by applying

a Bonferroni correction, to report all the effects at a 0.008 level of significance. The results show that the stress scores were significantly different in all conditions ($p < 0.008$ for all tests) except for Mikon-Object ($p = 0.32 > 0.008$). Hence we conclude that art images are most stressful to use, followed by doodles, whereas Mikons and objects (no significant difference) are least stressful to use.

5.7 POST STUDY QUESTIONNAIRE RESULTS

Rate images: In the post study questionnaire the participants were asked to rate the image types on a scale: will use, not sure and never use. We find most of the Mikon (17/25) and object (20/25) users would like to use these images as passwords in the future. The art users (17/25) disliked these images to be used as password and the doodle users had a split opinion.

Strategy used for password creation: The participants were asked to provide information on the strategy/method they used to create their passwords (Fig 9). The results reveal that most Mikon and doodle users either used a story/pattern to remember their passwords or they chose passwords according to their personal likings (Fig 8). Most art users chose passwords either based on their personal likings- favorite color, objects, scene or visual and aesthetic quality of the images i.e. attractiveness. But, these strategies for creating memorable passwords may either make them guessable to an intruder who knows the user quite well, or could be disclosed and thus shared with ease.

Fig. 8. A Mikon password created using pattern strategy.

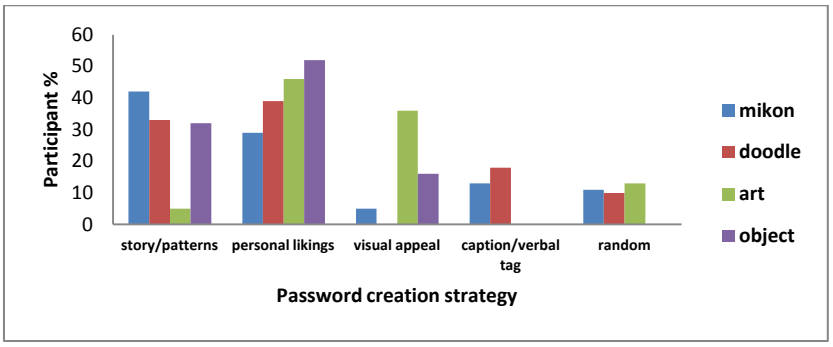


Fig. 9. Strategy employed to create image passwords

Recording passwords: None of the participants said they made an attempt to record their passwords but almost 84 % of the participants who used art passwords, 68 % of doodle users, 36% of Mikon users and 32% of object users reported that they might use screen captures, sketches or notes (written descriptions) to store their image passwords.

6 Discussion

This is the first study that compared the usability of multiple image passwords using 4 different images types- Mikon, doodle, art and objects in RBGS. The use of different experimental framework, dependant variables and image types in [11, 12] makes it difficult to allow systematic comparison of our results with them.

Our findings show that the effectiveness of graphical passwords is significantly affected by the type of images used. In this context, the results show that the mean login success percentage is highest for objects, closely followed by Mikons, then doodles and lowest for art images. According to the cognitive studies, dual coding theory [5] and guided search process [16], an elaborative encoding of an image in human memory makes it memorable. Thus, an image which is easily associated with a name (nameable) or can be interpreted in a meaningful way is likely to be more memorable due to superior encoding in human memory. In this context, mean scores of the sat 3 parameter (meaningfulness of the image) is: highest for objects (3.65/5); closely followed by Mikon (3.5/5); then doodles (2.8/5); lowest for art images (2.14/5). Thus we find that the results of the satisfaction parameter are in line with the mean successful login percentage. Hence the higher memorability of the object and mikon images can be attributed to the fact that users find these images meaningful and easy to associate with something. The doodle images are black and white line drawings and do not convey much meaning to aid memorability. Hence these images may not be encoded in an elaborate way in the human memory. The art images were very difficult to remember because according to the users, it was not only difficult to associate them with something meaningful but they were visually complex, containing a lot of information and color which would lead to information overload in memory. This complements the work reported in [17] which suggested that visual complexity of an image is linked to the ease of associating it with a name. The work has indicated that it is difficult to assign names to visually complex images. The results also reveal that the mean registration time is: lowest for objects; closely followed by Mikons; then doodles; highest for art images. In this context, the mean scores of the sat 1 parameter (ease to register) are: highest for objects (3.24/5); closely followed by Mikons (3.12/5); then doodles (2.85/5); lowest for art images (2.45/5). Thus the qualitative data obtained from the participants through the questionnaire complement the mean registration time obtained from the online study. These results can be attributed to the fact that users find it difficult to choose meaningful images in case of doodle and art, which they could use as passwords. The authentication time follows the same trend as that of the registration time: lowest for objects; closely followed by Mikons; then doodles; highest for art images. The mean scores for the sat-2 parameter

(ease to authenticate): highest for objects (3.56/5); closely followed by Mikons (3.36/5); then doodles (2.95/5); lowest for art images (2.45/5). Thus the mean scores of the sat 2 parameter complement the results of the login time data obtained from the online study. The above discussion suggests that the effectiveness and efficiency results complement each other. So we can conclude that images which are meaningful or can be associated with something easily are: effective in the sense of being memorable; efficient i.e. less time consuming to employ. This conclusion is also supported by the mean satisfaction scores obtained through the questionnaires: highest for objects (13.91/20); closely followed by Mikons (13.16/20); then doodles (11.8/20); lowest for art images (9.24/20) and mean stress scores: lowest for objects (11.87/20); closely followed by Mikons (12.2/20); then doodles (13.12/20); highest for art images (14.66/20).

7 Conclusions and Future work

We have presented the first study with 100 participants over a period of 8 weeks to compare the usability: mean success percentage for 7 weeks; mean registration and login time; the qualitative data i.e. participant's opinion (satisfaction as well as stress), of multiple image passwords using 4 different image types- Mikon, doodle, art and objects. The results of the study revealed that object and Mikon passwords performed best in each of the usability criteria compared to doodle and art passwords. So we conclude that meaningful or easily nameable image types are most usable when multiple graphical passwords are used. The experimental design in our study is: valid, it answers our research question through the data we collected for each measure; reliable, it can be reproduced by the research community; most importantly, such a study for the stated research problem has not been conducted in the past. In terms of improvement, the same study could be reproduced with a user-group other than students. The post study questionnaire results demonstrate that most people chose password images either by making a pattern/story or something which is related to them. These results underscore the need to examine, whether passwords created by using patterns aid memorability, when multiple graphical passwords are used and assess the ease of employing such a strategy these different image types. The results suggest that though meaningful images would aid usability when multiple graphical passwords are used, users may engage in insecure coping mechanisms like recording them through digital or non digital media. So in our ongoing work, we are investigating the vulnerability of image passwords to user descriptions (non digital attempt to describe the images by writing them or verbally communicating them). We also find that the statistical analysis alone does not unambiguously identify the most suitable image type to be used for graphical password. In our ongoing work, we are developing a framework that would help quantify the usability value for each image type, taking into account qualitative as well quantitative data obtained for all criteria from the experiment.

Acknowledgement

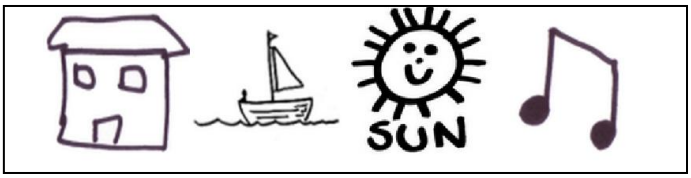
We are thankful to SICSA (Scottish Informatics and Computer Science Alliance) for funding this research. We are also thankful as well as grateful to Dr. Balvinder Shukla (Vice Chancellor Amity University, India) for helping us get the participants for the study and permitting us to perform the experiments with them.

References

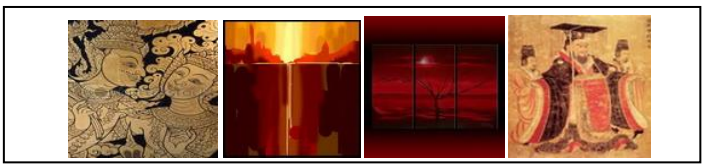
1. Adams A, Sasse M.A (1999) Users are Not the Enemy. Communications of the ACM, 40-46
2. Florencio D, Herley C (2007) A large-scale study of web password habits. In proceedings of 16th International Conference on World Wide Web, University of Calgary, pp. 657-666
3. Brostoff S, Sasse MA (2000) Are passfaces more usable than passwords? A field trial investigation. In proceedings of HCI on People and Computers XIV, pp. 405-424
4. Madigan S. Picture Memory (1983) Imagery, Memory, and Cognition: Essays in Honor of Allan Paivio. In: Yuille, J (ed), Hillsdale: Lawrence Erlbaum Associates
5. Paivio A (1986) Mental Representataions:A Dual Coding Approach, Oxford Press, UK.
6. Chowdhury S, Poet R (2011) Comparing the usability of doodle and Mikon images to be used as authenticators in graphical authentication systems. In proceeding of Conference on User science and Engineering, pp. 54-58.
7. Davis D, Monrose F, Reiter M (2005) On user choice in graphical password schemes. In proceedings of the Conference on USENIX Security Symposium, vol.13, pp. 11-11
8. Dhamija R, Perrig A (2000) Déjà vu: A user study using images for authentication. In proceedings of the Conference on USENIX Security Symposium, vol. 9, pp. 4-4
9. Renaud K (2005) A Visuo- Biometric Authentication Mechanism for Old Users. In proceeding of British HCI, pp. 167-182
10. Renaud K (2009) Web authentication using Mikon images. In proceedings of World Congress on Privacy, Security, Trust and the management of E-Business, pp. 1-10
11. Moncur W, Leplâtre G (2007) Pictures at the ATM: exploring the usability of multiple graphical passwords. In proceedings of SIGCHI Conference on Human Factors in Computing Systems, pp. 887-894
12. Everitt K, Bragin T, Fogarty J, KohnoT (2009) A comprehensive study of frequency, interference and training of multiple graphical passwords. In proceedings of Conference on Human Factors in Computing Systems, pp. 889-898
13. SUS questionnaire. Accessed 19th January 2013 http://www.usabilitynet.org/trump/methods/satisfaction.htm.
14. NASA Task Load Index. Accessed 19 January 2013 http://humansystems.arc.nasa.gov/groups/TLX/downloads/TLXScale.pdf
15. Strauss A, Corbin J (1990) Basics of qualitative research: Grounded theory procedures and techniques. Sage, Newbury Park
16. Wolfe M (1994) Guided search 2.0 a revised model of visual search. Psychonomic Bulletin & Review 1(5), pp. 202-238
17. Szekely A, Bates E (1999) Objective visual complexity as a variable in picture naming. CRL Newsletter Center for Research in Language, University of California, pp. 3-33

Appendix

Sample Mikon password



Sample doodle password



Sample art password

Sample object password