Do Background Images Improve “Draw a Secret” Graphical Passwords?

Paul Dunphy
School of Computing Science
Newcastle University, UK
p.m.dunphy@ncl.ac.uk

Jeff Yan
School of Computing Science
Newcastle University, UK
jeff.yan@ncl.ac.uk

ABSTRACT

Draw a secret (DAS) is a representative graphical password scheme. Rigorous theoretical analysis suggests that DAS supports an overall password space larger than that of the ubiquitous textual password scheme. However, recent research suggests that DAS users tend to choose weak passwords, and their choices would render this theoretically sound scheme less secure in real life.

In this paper we investigate the novel idea of introducing background images to the DAS scheme, where users were initially supposed to draw passwords on a blank canvas overlaid with a grid. Encouraging results from our two user studies have shown that people aided with background images tended to set significantly more complicated passwords than their counterparts using the original scheme. The background images also reduced other predictable characteristics in DAS passwords such as symmetry and centering within the drawing grid, further improving the strength of the passwords. We estimate that the average strength of successfully recalled passwords in the enhanced scheme was increased over those created using the original scheme by more than 10 bits. Moreover, a positive effect was observed with respect to the memorability of the more complex passwords encouraged by the background images.

Categories and Subject Descriptors

D.4.6 Security and Protection, H.1.2 User/Machine Systems.

General Terms

Security, Human Factors.

Keywords

Usable security, graphical passwords, authentication, Draw a Secret.

CCS’07, October 29–November 2, 2007, Alexandria, Virginia, USA.
Copyright 2007 ACM 978-1-59593-703-2/07/0010...$5.00.

1. INTRODUCTION

The use of textual passwords (i.e. alphanumeric passwords) for user authentication is ubiquitous. However, this common practice has some well-known weaknesses. For example, many people find it difficult to remember strong passwords, and they tend to choose easily memorable passwords that are susceptible to guessing.

Many of the deficiencies of textual passwords arise from the limitations of human memory [17]. Numerous cognitive and psychological studies have revealed that people perform far better when remembering pictures rather than words: as the saying goes, a picture is worth a thousand words. This has inspired research into the design of graphical passwords systems in both the security and HCI communities in recent years.

Draw a Secret [7] is a representative graphical password scheme, in which a user’s password is a free-form drawing produced on an N × N grid. DAS is alphabet independent and so is accessible to users of all languages. This scheme is particularly suitable for handheld devices such as PDAs due to its ability to accept graphical input on the screen just as you would write on a piece of paper.

Amongst various graphical password schemes, DAS is of particular interest and worthy of extensive study for the following reasons. First, in theory, DAS boasts an overall password space (i.e. total number of possible passwords) larger than that of the textual scheme. Second, unlike many graphical schemes, DAS can be used not only for user authentication (i.e., to verify the claimed identify of a user), but also for key generation (i.e., to use a password to generate a long cryptographic key).

However, recent research suggests that DAS users might tend to pick weak graphical passwords that are vulnerable to the *graphical dictionary attack* [12]. That is, user choices of their own passwords would render the theoretically sound DAS scheme less secure in practice for both user authentication and key generation.

One proposed solution to harden the DAS passwords was to use a method called *grid selection* [13]: a user first selects an N x N drawing grid within a much larger *selection grid*. Then they zoom in and create the secret as per the original DAS scheme. The location of the chosen drawing grid adds an extra degree of complexity to the password, as there are thousands of possible drawing grids within the selection grid. This technique in theory could significantly increase the password space by adding up to 16 bits to the password space. To our knowledge no user study of grid selection has been carried out, so it is unclear whether this works as well in practice as expected. A well-known lesson on usable security, in particular for password schemes, is that what engineers expect to work and what users actually make work are two different things [17].

In this paper, we present our own solution to this problem, a novel variant of the DAS scheme which we call BDAS (Background Draw-a-Secret). In BDAS, instead of creating a password on a blank canvas overlaid with a grid, a user will first choose a background image to be overlaid by the grid, and then draw their secret as in DAS. We aim to study whether this variant will enhance the original scheme. Specifically, we are interested in exploring whether a background image would encourage users to choose more complicated passwords, which are usually less vulnerable to dictionary and other guess attacks. We are also interested in whether the background image could aid users to remember their passwords.

Our main motivation is as follows. In DAS, it is difficult for a user to reconstruct a complex secret on the drawing grid as, clearly, all the cells are identical. The cells on the perimeter of the grid appear much easier to distinguish, while it is much harder to do the same with the centremost cells due to them being surrounded on all sides by identical cells. This can cause problems particularly when drawing curved objects that must cut cells accurately. So what if re-creation of a secret can be aided by something that reduces the confusion, such as a background image?

The results of our empirical studies are very encouraging. For example, people using the BDAS scheme in our controlled experiments tended to set less predictable and significantly more complicated passwords than their counterparts using the original scheme. Background images also improved password memorability: although BDAS users had to remember more complicated passwords, their recall success rate was about the same as that with DAS users who created simpler passwords.

The rest of this paper is organized as follows. We briefly review the DAS scheme in Section 2. Next we discuss related work and elaborate the novelty of our BDAS scheme in Section 3. We discuss in Section 4 the choice of background images, with particular attention to its relevance to our experiment design. Then in Sections 5 and 6, we report two comparative user studies of both the original and extended DAS schemes. In Section 7, we discuss our experimental results. Finally we draw some conclusions and discuss future work in Section 8.

2. DRAW-A-SECRET: AN OVERVIEW

In DAS, a password is a picture drawn free-form on a grid of size N × N. Each grid cell is denoted by two-dimensional discrete coordinates $(x, y) \in [1, N] \times [1, N]$. A completed drawing, i.e., a secret, is encoded as the ordered sequence of cells that the user crosses whilst constructing the secret. Each time a user lifts the pen from the drawing grid surface, a "pen-up" event is encoded by a distinguished coordinate pair (N+1, N+1). **Two secrets are identical if the encoding is the same, not the drawing itself.** This allows some margin of error as the drawing created does not have to be re-created precisely. That is, **the encoding of a particular secret has a one-to-many relationship with the possible drawings it can represent**.

A secret is disallowed if it contains a cell crossing where it is difficult to ascertain which destination cell had been intended. This also extends to a construction coming so close to a grid line that it is not obvious what the intended route was. Figure 1 shows two routes which would be disallowed by the system.

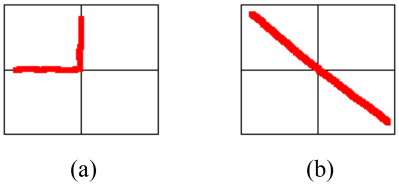


Figure 1. Illegal crossings due to *fuzzy boundaries*: (a) tracing grid lines (b) crossing through a cell corner.

Remembering the image itself is not good enough to login to a DAS protected system. To replicate a password, the user must cross the same cells, in the same order, lifting the pen in the same places, and not making any crossings that are difficult to encode.

The following terminology is critical for understanding the DAS scheme.

A *stroke* is a sequence of cell crossings bounded at both ends by pen-up events, exclusive of the pen-ups themselves. For example, the sequence *Pen-up, (1,2),(1,3),(1,4), Pen-up* defines a stroke: (1,2), (1,3), (1,4). The *length of a stroke* is the number of coordinate pairs it contains. Therefore, the above stroke has a length of 3.

Since a password in the DAS scheme is a sequence of strokes separated by pen-ups, the *length of a password* is the sum of the lengths of its component strokes, exclusive of pen-ups.

The number of strokes (i.e. the *stroke count*) and the password length are important security metrics measuring the strength of a DAS password. A high number of strokes or a high password length usually provides a high level of security as such secrets reside in a more secure part of the password space.

The impact of stoke count, password length and drawing grid size on the DAS password space was examined in [13], and it was found that the stroke count matters the most amongst these factors. However, increasing the stroke count is not the only way of improving the security of DAS passwords. A low stroke count can be compensated for by increasing the password length [12,13]. For example, the password space for DAS passwords with a stroke count of 4 or less and a password length of 20 is about 2^57 on a 5×5 grid (See Table 2 in the extended version of [13]). This is larger than the number of textual passwords of 8 characters or less constructed from the printable ASCII codes (2^53).

Other traits of strong DAS passwords include few trends of symmetry and centering within the drawing grid [12,8].

3. RELATED WORK

The notion of a "graphical password" is due to Blonder [1]. In Blonder's scheme, a password is created by having the user click one or more predetermined positions ('tap regions') on an image. PassPoints [15] extended this idea by eliminating the predefined tap boundaries and allowing arbitrary images to be used. Schemes such as VisKey (developed by a German company) [19] and V-GO [20] (developed by Passlogix) use essentially the same idea as PassPoints. V-Go alters the idea slightly by, for example, asking the user to click a location on the image of a room to hide an object, or click the order of ingredients to make a cocktail.

Some intuitive observations of the effects of image choices were discussed in [16] for PassPoints. Also in [16], in order to compare the effect of different images in this scheme, four everyday images were used for creating passwords, and each image was tested with a different group of subjects. The results indicated that the choice of image did not make any significant difference in PassPoints, in terms of both user performance (e.g. in password creation and retention) and user perceptions. On the contrary, a recent study [3] found that the images used had a significant impact in usability. For example, the users experienced difficulty selecting and remembering points on images with few salient areas.

It might appear that the BDAS scheme is just a trivial combination of the DAS scheme and the Blonder-style systems such as Passpoints. However, this is not true, as shown by the following subtle arguments.

In all the Blonder-style systems known to us, how a user creates a password is directly influenced by the content of the image used in the system. All that the user must remember are positions clicked and the order of those clicks.

In BDAS, a background image has a different purpose. As shown in [6], users may be able to remember what their password drawings look like, but fail to replicate them in the correct location in the grid (since the cells in the grid all look alike). We use the image with the grid superimposed over it, expecting the contents of the image will help users remember the location of their passwords in the grid, thus enhancing both human memorability and password complexity.

In addition, Blonder-style systems use the notion of selecting a point of interest in an image. In BDAS this kind of approach is possible (and to some extent these dot strokes increase the security of secrets); however, on the whole no areas of the image are explicitly selected. If an area is of interest to a user of BDAS, it could be used in a number of ways instead of touching it. For instance a user could draw a part of their image next to it, could gauge the location of another part of the drawing by using its position, or could draw over it. We are certain there are other scenarios not mentioned that would come to light after more study.

Two empirical studies are of direct relevance. An informal user study of the DAS scheme was reported in [8]. In this study, 16 computer science students were recruited, and each asked to draw a memorable DAS password on paper in order to determine if there are any predictable characteristics in the graphical passwords that people choose. The findings include that 30% of passwords were symmetric, 80% were composed of 1-3 strokes, and 86% were centred or approximately centred within the grid. However, no recall test was conducted in this experiment to check whether the passwords chosen by the users were memorable. Goldberg et al [6] conducted a user study of Passdoodles, i.e. hand-drawn doodle passwords. This scheme is very similar to the DAS scheme except that 1) a doodle can be drawn in a number of colors, 2) a doodle may include multiple stokes placed anywhere on the drawing screen and 3) it appears that no drawing grid was provided. Thirteen subjects were recruited to participate in this study, using a paper prototype of the system. It was found that people could remember complete doodle images as accurately as alphanumeric passwords, but they were less likely to recall the order in which they drew a doodle than the resulting image.

More distantly related is work on user choice in two recognition-based graphical password schemes [4], which shows that permitting user selection of passwords in schemes like Passfaces [2,10] can yield passwords with entropy far below the theoretical optimum.

One final point - graphical passwords is an active research topic; due to space limit, some interesting work is not covered here, but can be found in a recent survey of this topic in [11].

4. BACKGROUND IMAGE CHOICE

Our idea is to introduce background images to the original DAS scheme so that both a background image and the drawing grid can be used to provide cued recall. But what images would be suitable for inclusion? It is possible, in our DAS variant, that a user's drawing will be influenced by the image that is placed before them. For example, the number of 'hot spots' (areas of interest) on a background image might influence the distribution of the passwords created, and impose bias in certain areas.

Unfortunately, there is very limited prior art on the choice of image that is highly indicative or relevant for our purposes. Therefore, we built our work largely upon our intuition and the image study reported in [16], which was built on existing psychological research but focused on the PassPoints scheme only.

The image choice study in [16] assumed that the level of detail of an image is just as important as how many hot spots are present. A user selects a point by going through the following procedure: firstly they must select an area they wish to work in, and then they must pick a point in this chosen area to use. If the image is of high detail, a spot can easily be selected; if a low-detail image, it is much more difficult to pick a memorable point.

Figure 2 shows a map along with a cross section. If one had to select a point in the style of PassPoints, it would not be difficult as there are an abundance of hotspots that could be completely captured by a pen-point.

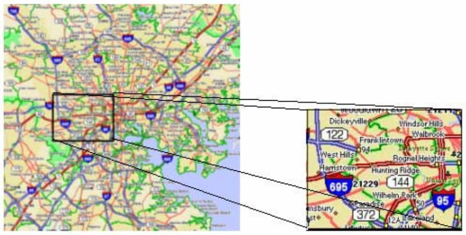


Figure 2. A cross section of a map (reproduced with permission from Collins Bartholomew)

Figure 3 is different in that, upon selecting an area of interest, the level of detail is low, making it difficult to select a pen-point sized area. Any focal points that were apparent in the main image (on the left) disappear when applying the point of the pen to the chosen area (the right image), due to the overall lack of detail.

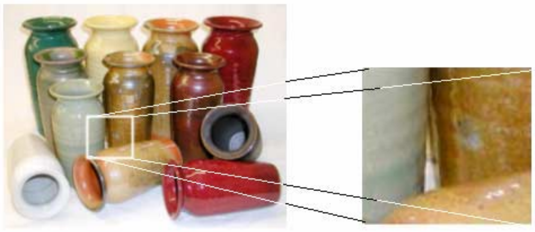


Figure 3. A cross section of some vases

The role of the background images in our variant of DAS and the role of images used in PassPoints may have a similar level of importance with respect to the security of the scheme. Thus it was decided to select images in much the same way, the main criteria being those that have meaningful content and are rich [16]. The images we planned to use were of an everyday theme that could be related to easily by everyone.

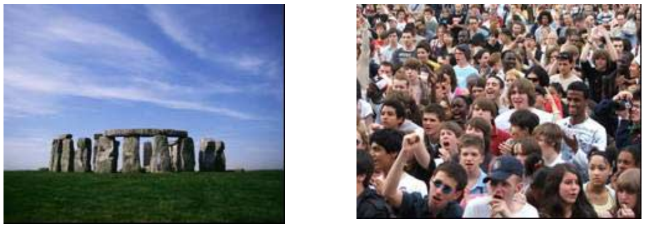


Figure 4. (a) A low-detail image (b) A busier image (reproduced with permission from Kodak and *http://www.lovemusichateracism.com*, respectively)

The issue of background images inducing bias is important from a security perspective. Such problems exist in the original DAS scheme, where users are inclined to incorporate symmetry about an axis (e.g. the center) in their drawings. When beginning to draw a symmetric image, the user must pick an 'anchor' point or region around which to develop their secret. As users of DAS are already inclined to create secrets influenced by the drawing grid alone, it is important to select images that do not introduce a possibility of imposing further bias based on their content. Figure 4(a) is an example of such an image that would almost certainly encourage the creation of predictably placed secrets, as the only focal point is the famous Stonehenge site. Figure 4(b) however contains many more focal points that could distribute the user's attention widely.

With these examples in mind, we decided that the most desirable images should contain many areas of interest and be bursting with content. Each person who views such an image may take an interest in a different part. This would increase the variation of the secrets created by users, and make it more difficult to predict how a particular background image would influence a user. (Any trends discovered with respect to a particular background image would enable an attacker to piece together a graphical dictionary of likely responses.)

It is worthwhile to note the following differences between the DAS and PassPoints schemes. First, there are fundamental differences in creating a free-from drawing as a password and clicking a number of locations on an image as a password.

Second, a user can make use of the background aids in the DAS scheme in any of the following three ways: 1) The user has a secret in mind to begin with and draws using points from the background to map various features of their drawing, 2) The user's choice of secret is affected by various characteristics of the image, and their secret shares traits exhibited by the image, and 3) A mix of the two. However, we believe that it is unlikely for a user of the PassPoints scheme to make use of the image aid in the first way.

We decided that we would also include an image of low detail in our experiment to gauge the effect of such a background choice.

As well as the content of background images, the format was also an important consideration. The displays of most PDAs are of low resolution, thus ideal images would have to mirror this setup. Distorted images are of course detrimental to the experiment as they are more difficult to focus on.

Another concern in our experiment design was how many images would be presented to the user to select from. Since people taking part in our experiment would be doing so on a voluntary basis, an excessive amount of images to look through would reduce their enthusiasm. The chosen approach was to gather images from a limited number of categories.

5. USER STUDY 1: A PILOT STUDY

5.1 Experiment Stationery

We developed paper stationery to capture input to our experiment and organise it so that information could be extracted as efficiently as possible. The main stationery contained two sets of drawing grids, one to 'set' the secret and another to login after a distraction period. This setup was printed onto both ordinary A4 paper and transparencies. The paper grid was given to participants who would use the original DAS scheme whilst the transparencies would be given to those testing our variant with background images. The size of the drawing grid was set at 3.7" (corner to corner) replicating the screen size of a top-selling PDA at the time. In addition, previous research [12] suggests that a 5×5 grid would provide a happy medium between usability and security. Therefore, we took this dimension as default in this study.

Aside from the stationery given to the user, a sheet was developed so that the experiment facilitator could make notes in a structured way whilst observing each participant. This too included two drawing grids so that they could trace the route of the image being created in front of them. It also included sections to record any interesting traits exhibited by the participant as well as the characteristics of the secret, e.g. length and stroke count.

An A4 template known as the "Practice grid" was also created, consisting purely of drawing grids, 6 in total. This was to give the participant ample opportunity to practice the drawing of a secret that abided by the rules of DAS. In order for participants using the BDAS scheme to practice, scribbling was allowed with a non-permanent marker upon the transparencies.

The background images themselves were printed on a sheet of A4 high quality paper, in a layout that mirrored the location of the grids printed on the transparency. This enabled the grid to easily overlay the desired image.

For those using a transparency, colored marker pens were available. This was to cover the possibility that the color of a background image clashed with the standard black marker.

5.2 Procedure

A total of 21 subjects were recruited, 15 male and 6 female. None of them had previously used DAS passwords. The subjects were then asked for their consent. The majority of the subjects (15) were in the age range of 18-30, while 5 subjects were in the age range of 31-50, and 1 in the age range of 50+.

Participant backgrounds and occupations were also collected to see if the DAS scheme was as accessible to people from a technical background as those from a non-technical background. 14 subjects were of non-technical backgrounds and the remaining 7 subjects technical. People in the technical category included students of science and engineering disciplines. Occupations found in the non-technical category include teachers, merchandisers and solicitors. These are types of jobs that would not normally be associated with technical expertise. It was important to get a good selection of people with these non-technical backgrounds as any new password scheme must be accessible to all, like the textual password scheme. It is the results from the non-technical people that will really show how usable a new password scheme is.

The subjects were randomly assigned to one of the following two experimental groups.

Control group (or the DAS group). People in this group were given the stationery that represented the original DAS scheme.

Background image group (or the BDAS group). People in this group were given the special stationery printed on transparencies to simplify the addition of a background image. Each subject in this group was presented with five images, and instructed to choose one as the background image. We allowed users to choose their own backgrounds, because we would like to gauge what kind of images would be popular. Any answer to this will certainly have implications to the security of BDAS passwords.

It was the intention to examine the characteristics of DAS passwords when the user is aided by a background image as opposed to the standard drawing grid. Our hypothesis was that users will set more complicated passwords with the help of background images to map the construction of their secret. Also we wanted to see whether this aid enables a user to remember their password over an extended period of time.

In advance of the experiment, each subject was given an information sheet. This provided information of their activities on the day of the experiment and most importantly explained the intricacies of a DAS password and the rules of creating one. A small number of example drawings were also included.

On the day, participants were given time to practice with their assigned scheme. Upon indication they had practiced enough, we began the experiment and asked them to create a secret on the grid in a slow and steady fashion. After a short delay we asked them to repeat what they had initially drawn. While all this was happening, a facilitator documented the route of their secret on the results sheet so that it could be compared to the route of their repeated secret. Then one week after the initial results were collected, the subjects were asked to try and re-create what they had drawn on the first occasion.

5.3 Background Images Used

The images used in our experiment are shown in Figure 5. The images were taken from various categories: space, nature, crowd, map and an image of low detail.

The stars image would test the temptation of users drawing lines that simply 'join the dots', the dots obviously being the stars. Due to the fact the stars are evenly distributed on the image, it would also be interesting to see to what extent this increases the distribution of the secret about the grid.

The crowd image came to mind almost immediately when thinking in terms of images that completely populate the space given to them. Many examples can be found of crowds at concerts and soccer matches where the content spills out beyond the frontiers of the image. This was considered to be beneficial for our purposes as there is no empty space.

The map category shares characteristics with both the crowd and space categories: snapshots of crowds usually constitute a snap shot of a bigger picture, and with space pictures there could be a tendency for the "join the dots" effect to occur with points of interest on the map.

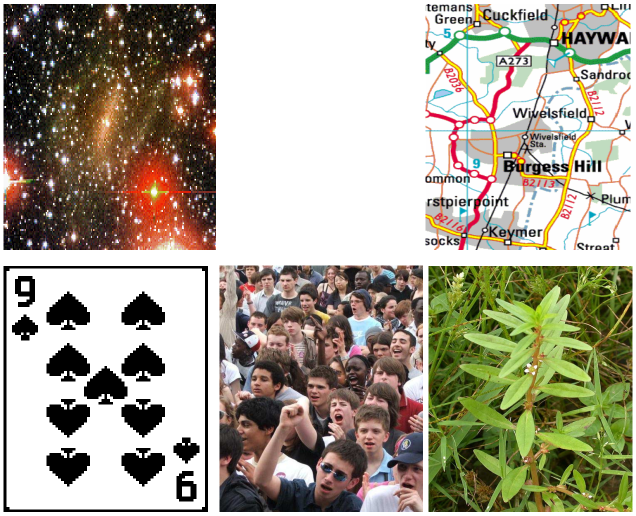


Figure 5. Background images used in the experiment (all reduced in size for this paper) (Map image reproduced with permission from Collins Bartholomew)

The image of the plant was chosen as, despite the number of hotspots in the image (leaves, blades of grass), the hotspots are of low detail. It is hoped that if selected, we will see how a user would react to this sudden lack of detail, when at first glance it appears to be rich with content.

For the low detail image a playing card was chosen, the 9 of spades. This doesn't share any traits with the previous categories as it is of low detail. It is hoped to gauge if this type of image would be popular and would have an adverse effect on the secret set.

5.4 RESULTS

5.4.1 Background Image Choices

The distribution of background image choices made by BDAS participants is shown in Figure 6.

The playing card was the most popular choice, and it was selected 50% of the time. The nearest to the playing card in popularity was the flower image, having 30% of selections. The rest of the images were very 'busy' and were not as popular. It was anticipated that images dense with content (such as the crowd scene and the map) would be the most popular, due to the many points of interest within. This was clearly contradicted.

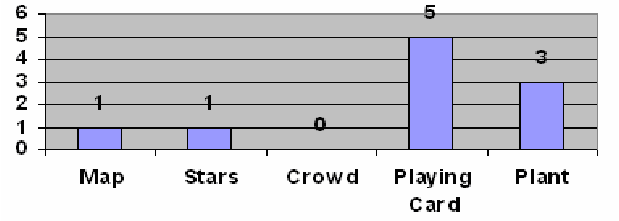


Figure 6. The distribution of background image choices (Study 1)

5.4.2 Secret Quality

We collected 10 valid secrets in each of the experiment groups.

Complexity. We compare secret complexity in the two groups in Table 1. The average password length with background images was 26.6 and without, 17.5. A t-test yields a result of t=2.377, p < .05 (one tail), indicating that the password length for the BDAS group was significantly longer than for those in the DAS group. The background images did also increase the stroke count of passwords on average, but not to a statistically significant level.

Table 1. Complexity of secrets in each group (Study 1)

Group	Strokes				Password length			
	Avg.	S.d.	Max	Min	Avg.	S.d.	Max	Min
BDAS	5.8	1.8	8	2	26.6	9.0	38	10
DAS	4.9	3.6	13	2	17.5	8.0	37	9

As such, we consider that our hypothesis that the background images did encourage people to choose secrets with more complexity was supported.

Symmetry. When examining all secrets, we found that 9 out of 10 (90%) secrets created in the DAS group had global symmetry, i.e. all the strokes of a drawing were symmetric about the same axis [12]. This figure is much higher than the percentage observed by Nali and Thorpe (30%) [8]. However, only 5 out of 10 (50%) secrets in the BDAS group possessed global symmetry.

According to [13, 12], global symmetry leads to weak DAS passwords that are vulnerable to dictionary attacks, and should be avoided. Therefore, the above observation appears to suggest that background images automatically aided the users in avoiding global symmetry, further improving the quality (i.e. security) of the resulting secrets.

Centering. In addition, 9 out of 10 (90%) secrets in the DAS group were centred or approximately centred within the drawing grid. This figure is close to the percentage observed in [8] (86%). However, only 7 out of 10 (70%) secrets in the BDAS group were centred or approximately centred within the grid. This appears to suggest that background images can also reduce to some extent another predictable characteristic in DAS passwords, making the resulting passwords even more secure.

5.4.3 Recall – Five Minutes After

A recall test was conducted five minutes after a secret was created. This can be regarded as a simulation of the common practice of password confirmation. Often, when a newly chosen (textual) password is set, it has to be repeated correctly for confirmation. In a textual password setting, many users might input their confirmation within half minute. However, since it usually takes much more time to either create or memorise a DAS password than a textual one, and since people were not so familiar with graphical passwords, we have assumed that on average the users repeat the password for confirmation after a longer period.

Table 2 compares the successful recalls in each group (Each participant was allowed three attempts). The recall success rate in the BDAS group was the same as in the DAS group. This appears to suggest that when a background image was available, people would at least perform the same in their password confirmation as in the original scheme, maintaining a smooth process of password setting.

Table 2. Recall results (5-minute test, Study 1)

Group	Total responses	Correct	Percentage
BDAS	10	8	80%
DAS	10	8	80%

Table 3 compares the complexity of secrets that were successfully recalled in two groups. Those secrets in the BDAS group on average did contain more complexity. A t-test yields a result of t=2.948, p < .01 (one tail), indicating that the passwords that were successfully recalled in the BDAS group were significantly longer than those in the DAS group. Furthermore, only 37.5% (3/8) of recalled secrets in the BDAS group had global symmetry, compared to 100% (8/8) in the DAS group; 50% of (4/8) recalled secrets in the BDAS group were centered, compared to 87.5% (7/8) in the DAS group.

Table 3. Complexity of successfully recreated secrets (5-minute test, Study 1)

Group	Strokes				Password length			
	Avg.	S.d.	Max	Min	Avg.	S.d.	Max	Min
BDAS	5.5	1.9	8	2	25.5	9.6	38	10
DAS	5	4.1	13	2	14.6	4.2	23	9

Estimated with Table 2 in the extended version of [13], the average strength of memorable passwords in the DAS group was between 48.7 and 50.9 bits in this recall test, less secure than that could be achieved by 8-character textual passwords (about 53 bits). On the contrary, the average strength of memorable passwords in the BDAS group was larger than 61.8 bits[1]. Therefore, although the BDAS scheme did not significantly increase the average stroke count, it significantly increased the security of memorable passwords by increasing the password length, and the effective password space for memorable BDAS passwords was increased by a factor of more than 1024.

5.4.4 Recall – A Week Later

One week later, each participant was asked to recall their secret. Each participant was allowed three attempts. Unfortunately not all the participants were available as some were in other parts of the country and Europe at the time. Among 8 valid responses in the BDAS group and 7 in the DAS group, four secrets were successfully recalled in each group. These secrets were also successfully recreated in the earlier 5-minute recall test. Table 4 below shows that the recall success rate was higher without the background images. However, it is useful to note that in practical terms, this was only a difference of one incorrect response, and due to low subject numbers, must be treated with caution.

Table 4. Recall results (1-week test, Study 1)

Group	Total Responses	Correct	Percentage
BDAS	8	4	50%
DAS	7	4	57%

Table 5 puts this into perspective by comparing the complexity of secrets that were successfully recalled in this phase.

[1] We believe this is an underestimate, as discussed later on in this paper.

Table 5. Complexity of successfully recalled secrets (1-week test, Study 1).

Group	Strokes				Password length			
	Avg.	s.d.	Max	Min	Avg.	s.d.	Max	Min
BDAS	5.75	2.6	8	2	20.25	10.5	35	10
DAS	2.25	0.5	3	2	14.5	2.4	18	13

On average, the secrets recalled with the background images showed much more complexity in terms of both stroke count and password length than those without. Furthermore, only 50% (2/4) of recalled secrets in the BDAS group had global symmetry, compared to 100% (4/4) in the DAS group; 25% of (1/4) recalled secrets in the BDAS group were centered, compared to 100% (4/4) in the DAS group.

In this recall test, the average strength of memorable passwords in the DAS group was less than 41.9 bits (estimated with avg. stroke count = 2.25 and avg. password length = 14.5), less secure than could be achieved by 8-character textual passwords. On the contrary, the average strength of memorable passwords in the BDAS group was larger than 61.8 bits (estimated with avg. stroke count = 5.75 and avg. password length = 20.25). Therefore, it appears that the BDAS scheme also significantly increased the security of memorable passwords by increasing the password length in longer term.

It is worthwhile to note that previous research assumed, based on psychology studies, that users would struggle to set secrets containing greater than 4 strokes [12]. Our recall results in the DAS group appeared to confirm this assumption. The average stroke count of secrets that were successfully recalled is merely 2.25, and the maximum stroke count 3. On the contrary, the average stroke count of the secrets that were successfully recalled in the BDAS group is 5.75, with the maximum stroke count being 8.

Table 6. Complexity of secrets failed to be recalled (1-week test, Study 1)

Group	Strokes				Password length			
	Avg.	s.d.	Max	Min	Avg.	s.d.	Max	Min
BDAS	5.25	0.96	6	4	27.75	2.63	30	18
DAS	4.67	1.53	6	3	24.67	11.59	37	14

In Table 6 we also compare the complexity of secrets that the participants failed to recall in the two groups. It is not surprising that on average, those secrets in the BDAS group were more complicated than those in the DAS group, in terms of both the stroke count and the password length.

By examining Tables 5 and 6, two interesting patterns can be observed. First, secrets that were not successfully recalled in the DAS group had a much larger average stroke count (4.67 vs. 2.25) and a much higher password length (24.67 vs. 14.5) than those that were successfully recalled. Second, secrets that were successfully recalled in the BDAS group had a slightly higher average stroke count (5.75 vs. 5.25) but a much lower average password length (20.25 vs. 27.75) than those that were not successfully recalled.

Furthermore, it is worthwhile to note that, unlike the DAS group, those who had the aid of the background image (including those who could not repeat their drawing successfully) all started from the correct cell. Each of them put this down to the aid of the background image.

5.5 Interpretation of Study 1 Results

With the aid of a background image, people tended to construct more complicated passwords (e.g. with a larger length or stroke count) than their counterparts using DAS. The background image appears also to reduce other predictable characteristics such as global symmetry and centering within the drawing grid that led to weak DAS passwords. A possible explanation is that DAS users employed centering and symmetry -- natural techniques of drawing -- as ways to help remember their secrets; BDAS users however had more options of placement and shape as they could use features of the image, thus reducing the predictable characteristics seen in DAS secrets.

There was also evidence that the background image improved the memorability of passwords. For example, no short-term detrimental memory effect was observed for using more complex passwords encouraged by background images. However, the question of whether the background image in the longer run helps the users remember the password better than the original DAS scheme does is more difficult to answer. No definite trends were encountered that suggested either way, although the users were more likely to remember their starting cell with the aid of the background images. In our view, this is largely due to the limited number of responses we received in the experiment, and a larger scale empirical study will give a clear answer.

6. USER STUDY 2

Results from our first study were very encouraging. However, the number of responses we received in the 1-week recall test was not large enough to do meaningful statistical tests. Therefore, we decided we run a second experiment on a larger scale. In this experiment, transparency was introduced to the DAS group as well, and a video camera was used to capture the drawing process of each participant, but everything else (including experiment stationery, procedure and background images) remained the same as in the first study.

We recruited 46 participants for this study, 32 male and 14 female. All the participants were undergraduate students at the time of the experiment, but none of them had previously used DAS or BDAS passwords. The typical age range of subjects was 18-25 with 1 participant in the group 50+. 20 subjects had technical backgrounds (majoring in computer science, engineering and etc.), and the remaining 26 subjects non-technical (majoring in language, business and others).We randomly assigned each participant to the DAS or BDAS group. Experimental results are discussed as follows.

6.1 Background Image Choices

Figure 7 summarises the distribution of background image choices in the BDAS group. The playing card was still the most popular image, and was selected 33% of the time. This was closely followed by the plant image with 30% of selections. The map and crowd images were the least popular choices.

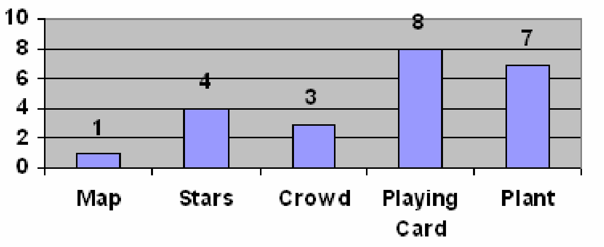


Figure 7. The distribution of background image choices (Study 2)

6.2 Secret Quality

We collected 23 secrets in the DAS group and 23 in the BDAS group, and all of them were valid.

Complexity. We compare secret complexity in two groups in Table 7. The average stroke count using BDAS was 7.22 as opposed to 5.30 with DAS. The standard deviation of stroke count with BDAS was 2.21 compared to 2.44 with DAS. A t-test with respect to the stroke count of secrets yields a result of t=2.78 p < 0.01, indicating that the background images significantly increased the stroke count of secrets.

Background images also increased the password length on average (21.43 vs. 18.26), but not to a statistically significant level. The standard deviation with respect to password length was again lower with BDAS (7.76 vs. 9.19).

Table 7. Complexity of secrets in each group (Study 2)

Group	Strokes				Password length			
	Avg.	S.d.	Max	Min	Avg.	S.d.	Max	Min
BDAS	7.22	2.21	12	4	21.43	7.76	37	6
DAS	5.30	2.44	10	1	18.26	9.19	42	6

The average strength of all passwords in the BDAS group is estimated to be of larger than 70.2 bits. The corresponding figure for the DAS group would be less than 60 bits. That is, around a 10 bit difference.

Symmetry. We found that 43% of secrets (10/23) created with BDAS exhibited global symmetry, compared to about 57% (13/23) in the DAS group. Symmetry was reduced with BDAS, but not to the extent we observed in the first study.

Centering. We found that 43% (10/23) secrets with BDAS exhibited centering, while 87% (20/23) using DAS exhibited this too. To a large extent centering was reduced with BDAS, as the background image content gave the users different options for image placement.

6.3 Recall – Five Minutes After

Table 8 summarises the results of recall tests conducted five minutes after a secret was created in each group. In this test, only one person in the BDAS group was unable to repeat their secret, and the recall success rate is quite close in two groups.

Table 8. Recall results (5-minute test, Study 2)

Group	Total responses	Correct	Percentage
BDAS	23	22	96%
DAS	23	23	100%

Table 9 compares the complexity of secrets that were successfully recalled in two groups. The average stroke count of a drawing created with BDAS was 7.45 as opposed to 5.09 with DAS. The standard deviation was also lower with BDAS (2.26 vs. 2.41). A t-test gives t=2.9 p< 0.01, indicating that the stroke count of the successfully recalled BDAS secrets was significantly larger. The length of recalled BDAS passwords was also larger on average (21.7 vs. 17.71), though not to a statistically significant level. Furthermore, only 40.9% (9/22) of recalled secrets in the BDAS group had global symmetry, compared to 56.5% (13/23) in the DAS group; 40.9% (9/22) of recalled secrets in the BDAS group were centered, compared to 87.0% (20/23) in the DAS group.

Table 9. Complexity of successfully recalled secrets (5-minute test, Study 2)

Group	Strokes				Password length			
	Avg.	S.d	Max	Min	Avg.	S.d.	Max	Min
BDAS	7.45	2.26	12	4	21.7	8.31	37	6
DAS	5.30	2.44	10	1	18.26	9.19	42	6

In this recall test, the average strength of memorable passwords in the BDAS group was greater than that of all passwords in the same group (stroke count: 7.45 vs. 7.21, password length: 21.7 vs. 21. 43), but the average strength of memorable passwords in the DAS group was weaker than that of all passwords in the same group (stroke count: 5.09 vs. 5.30, password length: 17.71 vs. 18.26). Therefore, the difference of the strength of memorable passwords in two groups was larger than 10 bits.

Additionally we examined the number of attempts users needed to recall a secret successfully. The performance of DAS users was marginally better than BDAS: 18 DAS users succeeded in recalling their secrets first time, compared to 16 with BDAS; 5 users in both schemes needed 2 attempts, and 1 BDAS user required 3 attempts; no DAS user needed more than 2 attempts. The figures returned in this analysis are similar, so users performed about the same repeating more complex secrets on average than DAS, implying that BDAS did not hinder user performance.

6.4 Recall – A Week Later

One week later, participants were asked to recall their own secrets created the week before. Not all people returned, and 2 participants were lost in each group.

We received 42 responses in this test and 2 people (one from each group) were unable to repeat their secrets within the 3 attempt limit. Table 10 compares the successful recalls in each group in this test, and the recall success rate is the same in two groups.

Table 10. Recall results (1-week test, Study 2)

Group	Total responses	Correct	Percentage
BDAS	21	20	95%
DAS	21	20	95%

Table 11 compares the complexity of secrets that were successfully recalled in two groups. The complexity of secrets recalled in the BDAS group was again greater than in the DAS group. A t-test gives a result of t=2.96 p<0.01, indicating that the stroke count of passwords that were successfully recalled was significantly larger. The length of recalled BDAS passwords was also larger on average (20.9 vs. 17.45). Furthermore, only 45.5% (10/22) of recalled secrets in the BDAS group had global symmetry, compared to 59.1% (13/22) in the DAS group; 45.5% (10/22) of recalled secrets in the BDAS group were centered, compared to 86.4% (19/22) in the DAS group.

The average strength of memorable passwords with BDAS in this recall was larger than 70.2 bits, but the corresponding figure for the DAS group would be less than 60 bits. That is still around a 10 bit difference. Although the average stroke count for memorable passwords in DAS group was 5, numerous secrets in this set had a stroke count of below 5. On the other hand, the average strength of memorable passwords in each group in this test was slightly weaker than that in the 5-minute recall. This is reasonable, since memory fades as time goes by.

We observed in the first study that all BDAS users (including those who failed to recall their secrets in the 1-week test) could all remember to start from the correct cell. However, this was not observed in the current study.

Table 11. Complexity of successfully recalled secrets (1-week test, Study 2)

Group	Strokes				Password length			
	Avg.	S.d.	Max	Min	Avg.	S.d.	Max	Min
BDAS	7.1	2.16	12	4	20.9	7.71	37	6
DAS	5	2.44	10	1	17.45	7.63	37	6

We collected the number of attempts each user needed to recall their secret. Again, DAS users performed marginally better, repeating 16 secrets correctly first-time compared with 13 using BDAS. 5 BDAS users needed 2 attempts compared with 4 from DAS, and only 1 person - a BDAS user - required 3 attempts. The same observation as before could be applied here: although secrets of more complexity were encouraged, BDAS would not hinder user performance in authentication.

6.5 Interpretation of Study 2 Results

Study 2 has confirmed that BDAS did significantly increase password complexity by increasing stroke count and password length, and it also made passwords less predictable by reducing global symmetry and centering. It appears that the explanation we made in the pilot study that centering and incorporating global symmetry are coping techniques employed by DAS users to make secrets memorable on an un-memorable grid was also confirmed.

Furthermore, the BDAS passwords were just as memorable as the weaker DAS passwords created. This outcome is important as any added gains in terms of complexity are useless without a sufficient gain in memorability. Our two studies are in fact consistent in these results. Due to the number of participants, we believe that the second study is even more reliable.

7. FURTHER DISCUSSIONS

7.1 On Background Image Choices

An interesting point to note in terms of image choice from both studies is that the playing card was the most popular followed by the plant image, although it was anticipated that the images dense with content would be the most popular.

The feedback we received suggests the popularity of the playing card was due to the image being much easier to digest than the others. Also participants felt the spades on the card appearing in a symmetric fashion made them feel it would be easier to construct a secret around them without being distracting.

The feedback from participants choosing the plant image included that the plant divided the picture well and provided a good focal point around which to construct a drawing without being too distracting.

In fact the word 'distracting' was a recurring one. Our own guesses at what makes a BDAS image distracting includes a number of contrasting colors and few elements of symmetry. Future study will determine if these observations hold in general.

The remaining 3 images show much less of structure but contain the most color. This is possibly why they were chosen so few times. The participants selecting the *space* and *crowd* images commented that they simply liked the image as a whole.

The security implications of varying background image choices made by the BDAS users are as follows. First, allowing users a free choice of background image in the real world might reduce the password space. Second, if both popular and unpopular images are available for users to chose, attackers who know the difference of these images can make use of this knowledge to their advantage.

7.2 On Estimating BDAS Password Strength

In Sections 5 and 6, the strength of BDAS passwords was estimated using the method designed for non-background grid drawings, i.e. DAS secrets. As discussed earlier, background images in principle could introduce bias to BDAS passwords, reducing their security. Would this imply that the observed increase in BDAS password complexity may not in fact indicate increased password security?

Our estimate of BDAS password strength did ignore possible negative biases that could be introduced by background images. However, what these negative biases are is an open problem, and so is the following:

- How these biases would aid attackers?
- Whether security reduction caused by these biases can be compensated by reduced symmetry and centering?

At the time of preparing the camera-ready version of the present paper, we became aware of two recent studies [14,5] that demonstrate the effects of hotspots could cause serious security downgrade in the PassPoints scheme. The weakness highlighted by these studies is a concern, however at present we have little knowledge on the implications of these results in the different BDAS environment.

Nonetheless, it is worthwhile to note:

- No significant bias that could impact security, other than reduced symmetry and centering, was observed in our experiments (though we agree it might not be the case in a larger scale study).
- Strength figures we used for BDAS passwords, for convenience, were just the minimum estimates directly quoted from Table 2 in the extended version of [13], while off scale estimate was not calculated.
- Our estimate did not consider either symmetry or centering reduced by BDAS at all. If considered, both would increase our password strength figures.

Another interesting question is how a system designer would determine and evaluate the possible negative impacts on password entropy imposed by a given background. In fact, evaluating the effect of different types of images on BDAS passwords is our future work. For example, we plan to identify hot spots in BDAS background images with eye tracking, and examine the correlation between those hotspots and BDAS passwords created by users. Eventually we will establish what effects these hotspots have on the system, then using this knowledge to propose guidelines for background image choices for BDAS.

7.3 Artistic Skills

The secrets created in each group very much depended on the artistic ability of the participant. On a number of occasions during the experiment, participants in both studies expressed concern at their lack of artistic skills and so their secrets tended to involve writing characters in grid squares in a pattern they could remember. Figure 8 shows three secrets produced by subjects who professed to be bad artists. The first drawing was from Study 1, and resulted in a weak secret featuring a high number of dots and with a relatively low password length. The remaining two were from Study 2.

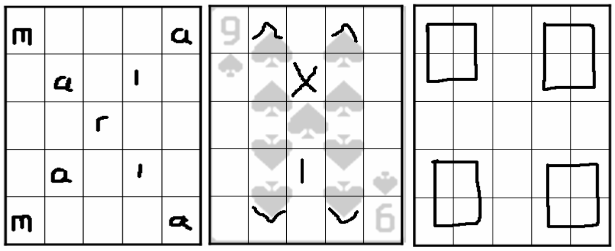
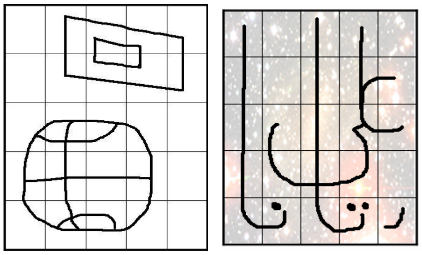


Figure 8. Secrets created by participants claiming to be bad artists

Of the participants who attempted to draw something, the most popular creations in the first experiment were simple, everyday objects such as cars, cups, and in particular houses, all of which accounted for 38% of the secrets created. Experiment 2 had a slightly different outcome where the most common creations were random constructions of lines and shapes with no significant meaning to the participant. The participants associated with these creations were often short of drawing ideas when starting the experiment. Such constructions accounted for 30% of drawings while the aforementioned group of everyday objects accounted for 20%. The most interesting creations were that of a basketball and a backboard, and one participant's name written in very complex Persian script (see Figure 9). Both were successfully repeated in the recall tests and were of high complexity.



(a) (b)

Figure 9. Memorable complex secrets (a) Basketball and backboard, (b) Persian Name.

At the end of both experiments participants were asked their opinions of the schemes to which they were assigned. The participants who were more capable artists expressed contentment using both BDAS and DAS. Those who were not so capable remarked they were happier using passwords rather than DAS or BDAS. Those using DAS were then introduced to BDAS and most claimed they would have found BDAS to be more stimulating to use, especially if they were able to use their own images as a background.

7.4 Weak Passwords

It is not surprising that there were very weak passwords in both groups in each user study. For example, in the first user study, one subject in the BDAS group created a secret with a stroke count of 2 and a password length of 10, and another subject in the DAS group created a secret that had a stroke count of two and a password length of 13. Their corresponding maximum password spaces, according to [13], were about 26.9 and 32.7 bits, respectively.

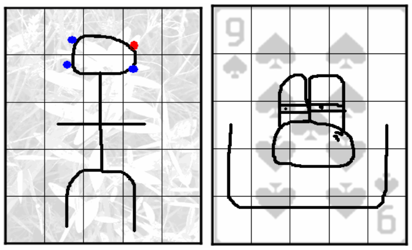
We did not expect that the BDAS scheme would eliminate all weak password choices. Rather, low-quality passwords in both the DAS and BDAS schemes should be disallowed by proactive password checking [18]. As BDAS increases secret complexity, the likelihood that a user secret would pass proactive checking is increased. A usable BDAS implementation in combination with proactive checking could potentially be a usable authentication solution. We believe that proactive checking would be more obstructive when applied to DAS due to our experience of people setting comparatively weaker passwords.

We have implemented, GraphiCheck, a proactive checker for DAS and BDAS graphical passwords. An empirical evaluation of such a tool is ongoing work, and is beyond the scope of the present paper.

7.5 Recall Errors

In the majority of cases where participants made an incorrect recall, the overall image of their secrets was at least roughly remembered. The nature of many recall failures was down to either mixing up the order of the strokes in a secret, or forgetting the starting point of a symmetrical shape such as a circle or square.

For example, the only recall failure in the 5-minute test of Study 2 was for the secret shown in Figure 10(a), a stick man. The problem encountered was not remembering the overall secret, but remembering the starting point of the head. The original starting point is indicated by the red dot in the image, his attempts are shown by the blue dots.



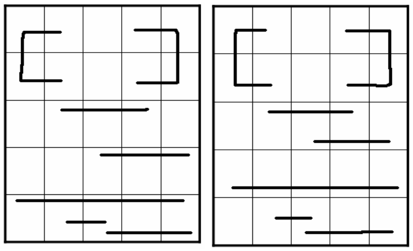
(a) (b)

Figure 10. Secrets failed to recall (a) in the 5-minute test, Study 2, (b) in the 1-week test, Study 2

Recall failure in the 1-week test occurred to the participant using a secret shown in Figure 10(b). Visually, the participant correctly repeated the appearance of his drawing each time; however it proved to be too difficult to remember the stroke order, especially where intricate work was required round the eyes.

We spotted similar cases in our first study. One participant created a car including detail such as the headlights and doors (strokes: 8, password length: 36, background: PLANT). In the 5-minute recall, the participant was overwhelmed by remembering the order in which everything was drawn, despite remembering the features of the image perfectly.

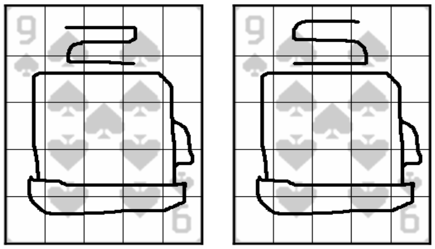
The above cases all echo the observation by Goldberg et al that people have difficulty in remembering the order of strokes although they can memorise the drawings [6]. In fact, stroke order being an important determinant of correct secret recall was met with exasperation by many participants in our experiments. Those who did not initially understand this tended to create more complex drawings in the practice runs. Then when this understanding dawned upon them, their creations were changed to appear much simpler.



(a) **(b)**

Figure 11. Memory decaying (a) the first creation (b) a week later.

Another interesting failure case occurred in the 1-week test of Study 2. A participant created a secret as in Figure 11(a). He had no problem in recreating it in the 5-minute test, but a week later he could not do better than producing the secret as in Figure 11(b). As can be seen the overall proportions of the secret were remembered correctly. However the details of positioning were not. It is interesting to note that the participant admitted defeat after 6 attempts, 3 more than allowed. Upon introduction to BDAS after the experiment, he thought he would have been assisted greatly by the background image. The image would have enabled him to map his lines to part of a meaningful background, rather than white space.



(a) (b)

Figure 12. Memory decaying (a) the first recall (b) a week later

Another interesting case of memory decaying over a week occurred in our first study. One participant successfully recalled his secret, a tea cup drawing with (1, 4) as the starting cell (see Figure 8(a), in the 5-minute test. However, the steam represented by a number 2 like shape as shown in Figure 12(a) manifested itself as an 'S' shape as in Figure 12(b) a week later. The rest of the cup was reproduced correctly up to that point.

7.6 Starting Cells

We also observed that BDAS increased the distribution of cells people would use to start their drawing. Observing DAS in both studies we found that right handed participants were most likely to start drawing between the centre of the grid and the far left. In the first study we found that 100% (9/9) of right-handed participants followed this trend, while in study 2 this was 78% (14/18). BDAS bettered this as 75% (6/8) behaved this way in study 1 followed by 48% (10/21) in study 2.

The small number of left-handed participants for both studies means that a similar conclusion could not be reached. A study more evenly populated with left-handed participants would be required to validate any trend.

7.7 Implementation Considerations

Observations in the two user studies suggest the following enhancements to implementation of DAS and BDAS.

Step-by-step undo. An overlying theme across both studies was that participants from both groups made errors and wanted to remove particular strokes from the grid. This suggests that a user-friendly DAS/BDAS implementation would ideally support functionalities such as *step-by-step undo*.

Fuzzy boundaries. Another observation was that most users initially had difficulty constructing a secret that obeyed the restriction of grid crossings that traced grid lines, or entered the so-called 'fuzzy boundaries'. This effect was reduced with BDAS although not eradicated. Users often found the grid lines to be the most natural point of reference when drawing their secret. This was a major factor in distributing practice grids in our experiment. Therefore, the same observation by Nali and Thorpe [8] is confirmed by our experiment. This kind of user error could be dealt with by the step-by-step undo discussed above, but details of such an event must be communicated effectively so as to avoid the inevitable confusion of why the user cannot make such a seemingly elementary move. On the other hand, this could also be an indication that *an enhanced drawing encoding system* that embraces interaction with the gridlines is more feasible.

Enrolment. Initially, some participants found the process of creating a valid secret difficult. However many were confident that practice would ultimately improve their technique. We believe that any usable implementation of DAS/BDAS (and other graphical password systems) must introduce a period of enrolment, so that the user can practice and commit to memory the drawing they have created.

We also observed in our studies that the attempts of participants in both groups were much more hesitant in re-creating their drawings in the 1-week recall test. Participants often held their pen down in the same location for an extended period of time while they were thinking. It is likely that such an effect would also be reduced by an effective enrolment process.

Learning by repetition is a weak way to remember anything. An ideal enrolment process in graphical password systems should focus towards the user being led through a process that prompts a more reflective [9] encoding in memory.

7.8 Others

One interesting observation was that users with non-technical backgrounds had much more difficulty understanding the DAS scheme than those of a technical disposition. This appears to suggest that "*easy to describe*" and "*easy to understand*" might be important concerns when designing new graphical password schemes.

8. CONCLUSIONS AND FUTURE WORK

Two comparative user studies have shown that it is an effective enhancement to introduce background images to the original DAS scheme. With the aid of a background image, people tended to construct significantly more complicated passwords than their counterparts using the original scheme, and other predictable characteristics such as global symmetry and centering within the drawing grid that led to weak DAS passwords were also reduced. The background image also improved the memorability of passwords. Although people aided by a background image had to recall significantly more complicated passwords than their counterparts using DAS, the former performed just as well as the latter in terms of recall success in tests conducted 5 minutes and 1 week after the passwords were created.

We believe that BDAS is more effective than DAS both for user authentication and for key generation. The most exciting feature of BDAS is that, with a simple enhancement, it provides significantly enhanced usability and security simultaneously.

We believe that this work provides a significant extension to the study of security and usability of graphical passwords. Results obtained in our experiments are thought-provoking and pave the road to numerous further studies.

A number of experimental improvements are as follows. First, although this has been the largest user study of the DAS scheme so far, the experiment could be carried out on an even larger scale.

Second, the participants taking part in our studies had no incentive to perform as if protecting or accessing anything of real-life value to them. Development of a scenario where the user has some real benefit from performing desirably and some cost from not, would be a useful advance to research when studying user password security.

Also, the next logical step would be to move the experiment out of the laboratory as it were and test it in an environment that mirrors more closely the real process of logging into a system. Many new ideas work well in the laboratory but when put into mainstream existence are completely debunked.

We have implemented the BDAS scheme on PDAs, and we plan to run this study with this actual implementation on a larger scale in the near future. We will also carry out detailed studies to compare the effect of different background image choices on password complexity and memorability, to ascertain what will make good background images, and to compare the memorability of passwords in this enhanced DAS scheme and ordinary textual passwords. Not only would this larger scale study be more ecologically valid, but would also allow for more extensive and thorough statistical testing.

Furthermore, as in other settings, shoulder-surfing and interference between multiple passwords are concerns for BDAS too. We have plans to investigate these issues. For example, we have designed a BDAS variant that provides shoulder-surfing resistance, and a user study evaluating its effectiveness and usability is our ongoing work.

ACKNOWLEDGEMENTS

We thank David Greathead, Brian Randell and Lindsay Marshall for many useful discussions, constructive feedbacks and proofreading, and thank anonymous referees and our shepherd Lorrie Cranor for their valuable comments.

REFERENCES

1. G. Blonder. Graphical passwords. US Patent 5559961, 1996.
2. S. Brostoff and M. A. Sasse. Are Passfaces^{TM} more usable than passwords? A field trial investigation. Proc. of HCI, 2000, pp 405–424
3. S Chiasson, R Biddle and PC van Oorschot. A Second Look at the Usability of Click-Based Graphical Passwords. Symposium on Usable Privacy and Security, July 2007, CMU, USA. ACM Press.
4. D. Davis, F. Monrose, and M. K. Reiter. On user choice in graphical password schemes. Usenix Security, 2004.
5. A E Dirik, N Memon and J-C Birget. Modeling User Choice in the PassPoints Graphical Password Scheme. SOUPS'07.
6. J. Goldberg, J. Hagman, and V. Sazawal. Doodling Our Way to Better Authentication, Extended Abstracts CHI'02, 2002.
7. I. Jermyn, A. Mayer, F. Monrose, M. K. Reiter, and A.D. Rubin. The Design and Analysis of Graphical Passwords, Proc. USENIX Security Symposium, 1999.
8. D Nali and J Thorpe. Analyzing User Choice in Graphical Passwords, Technical Report TR-04-01, School of Computer Science, Carleton University, 2004
9. D Norman. Things That Make Us Smart: Defending Human Attributes in the Age of the Machine. Addison Wesley, 1994.
10. Real User Corporation. The Science Behind Passfaces. Revision 2, Sept 2001. Available at http://www.realuser.com/published/ScienceBehindPassfaces.pdf.
11. X. Suo, Y Zhu and GS Owen. Graphical Passwords: A Survey. ACSAC, 2005.
12. J. Thorpe and P. C. van Oorschot. Graphical Dictionaries and the Memorable Space of Graphical Passwords. Proc. USENIX Security Symposium, 2004.
13. J. Thorpe and P. C. van Oorschot. Towards secure design choices for implementing graphical passwords. ACSAC, 2004. An extended version available at http://www.scs.carleton.ca/~jthorpe/extendedStrokes.pdf.
14. J Thorpe and PC van Oorschot. Human-Seeded Attacks and Exploiting Hot-Spots in Graphical Passwords. Usenix Security, Aug 2007.
15. S Wiedenbeck, J Waters, JC Birget, A Brodskiy and N Memon. PassPoints: design and longitudinal evaluation of a graphical password system. Int'l J. of Human Computer Studies. vol. 63, pp.102-127, 2005.
16. S Wiedenbeck, J Waters, JC Birget, A Brodskiy and N Memon. Authentication using graphical passwords: effects of tolerance and image choice. SOUPS'05, CMU, USA. ACM Press.
17. J. Yan, A. Blackwell, R. Anderson and A. Grant. Password Memorability and Security: Empirical Results. IEEE Security & Privacy, Vol. 2 No. 5, 2004.
18. J. Yan. A Note on Proactive Password Checking. ACM New Security Paradigms Workshop, New Mexico, USA, 2001.
19. VisKey, http://www.sfr-software.de/cms/EN/pocketpc/viskey/index.html, last accessed in Feb, 2007.
20. V-GO, http://www.passlogix.com/, last accessed in Feb, 2007.