Influencing Users Towards Better Passwords: Persuasive Cued Click-Points

Sonia Chiasson[1,2] Alain Forget[1,2] Robert Biddle[2] P.C. van Oorschot[1]

[1]School of Computer Science, [2]Human-Oriented Technology Lab
Carleton University, Ottawa Canada

chiasson@scs.carleton.ca, aforget@scs.carleton.ca, robert_biddle@carleton.ca, paulv@scs.carleton.ca

ABSTRACT

Usable security has unique usability challenges because the need for security often means that standard human-computer-interaction approaches cannot be directly applied. An important usability goal for authentication systems is to support users in selecting better passwords, thus increasing security by expanding the effective password space. In click-based graphical passwords, poorly chosen passwords lead to the emergence of hotspots – portions of the image where users are more likely to select click-points, allowing attackers to mount more successful dictionary attacks. We use persuasion to influence user choice in click-based graphical passwords, encouraging users to select more random, and hence more secure, click-points. Our approach is to introduce persuasion to the Cued Click-Points graphical password scheme (Chiasson, van Oorschot, Biddle, 2007). Our resulting scheme significantly reduces hotspots while still maintaining its usability.

Categories and Subject Descriptors

H.5.2 [Interfaces and Representation]: User Interfaces – *Graphical user interfaces*; K.6.5 [Computing Milieux]: Security and Protection – *Authentication*.

General Terms

Design, Experimentation, Security, Human Factors.

Keywords

Graphical passwords, persuasive technology, usable security.

1. INTRODUCTION

People select predictable passwords. This occurs with both text-based and graphical passwords. Users tend to choose passwords that are memorable in some way, which unfortunately often means that the passwords tend to follow predictable patterns that are easier for attackers to exploit. While the predictability problem can be solved by disallowing user choice and assigning passwords to users, this usually leads to usability issues since users cannot easily remember such random passwords.

An authentication system should encourage strong passwords while still maintaining memorability. We propose that users be *persuaded* to select more secure passwords. Our proposed system allows user choice while attempting to influence users to select stronger passwords. It also makes the task of selecting a weak password (easy for attackers to predict) more tedious, in order to discourage users from making such choices. In effect, our scheme makes choosing a more secure password the “path-of-least-resistance”. Rather than increasing the burden on users, it is easier to follow the system’s suggestions and create a more secure password; a feature that is lacking in other schemes.

We applied our approach to a click-based graphical password system and conducted an in-lab usability study with 39 participants. Our results show that our Persuasive Cued Click-Points scheme is effective at reducing the number of hotspots (areas of the image where users are more likely to select click-points) while still maintaining usability. While we are not arguing that graphical passwords are the best approach to authentication, we find that they offer an excellent environment for exploring strategies for helping users select better passwords since it is easy to compare user choices. Indeed, we also mention how our approach might be adapted to text-based passwords.

As an independent research contribution, we introduce and utilize a statistical approach for determining and comparing clustering in point patterns that arise in graphical passwords, by using spatial statistics typically used in earth sciences and biology.

The remainder of this paper is organised as follows. We first discuss background literature on usable security, graphical passwords, and persuasive technology. Next we describe our Persuasive Cued Click-Points system and methodology for the usability study. Finally we provide analysis and discussion of the results.

2. BACKGROUND

Designing user interfaces for authentication systems, and security applications in general, raises some interesting challenges. While the area of *usable security* [6] can draw from existing HCI knowledge, some fundamental differences must be taken into account. The properties of security systems that set them apart include:

- There is a second group of users, namely illegitimate users, who are actively trying to attack the system. Such attackers will

Human-Computer Interaction Conference (HCI 2008), Liverpool UK.

Published by the British Computer Society

exploit any information leaked by, or that can be extracted through, the interface. They will also leverage any way that the system can be misused or any means to spoof the interface to trick legitimate users. This makes providing helpful feedback difficult, as it may also help attackers.

- Security is typically a secondary task [28]; if it impedes users' primary goals, users will often try to circumvent security.
- Users have poor mental models of security [4, 28] and often misunderstand or underestimate the consequences of insecure actions. They may not even realise that their actions are insecure in the first place.
- Computer security suffers from the "barn door" property [28]: if information or a system is exposed even for a brief time, there is no guarantee that it has not been compromised in an irrecoverable way.

While these represent security concerns, they are all directly related to users of the system and as such, solutions must focus as much on the HCI aspects of the system as on the technical security components.

For example, authentication schemes have both a theoretical and effective password space. The former space includes the set of all (theoretically) possible passwords. User choices tend to fall into a much smaller subset of the full password space, known as the effective password space. To illustrate, 4-digit PINs offer 10000 possible combinations (0000 to 9999). However, some digit combinations are much more likely to be selected by users, such as years or patterns like 1234. Therefore, while the theoretical password space has a size of 10000, the effective password space is much smaller. We use PINs only to illustrate the concept of password spaces. As their small theoretical password space makes them inherently insecure, PINs are typically used in conjunction with a second authentication method such as providing an access card.

An important security goal is to design a system that maximises the effective password space. Since the effective password space is determined by user behaviour, such a design involves usability as well. The resulting usability goal is that users must be encouraged to select more secure passwords without sacrificing the usability of the system.

One of the challenges in measuring the effective password space is determining a proximity function (a measure of similarity between items). With text passwords, there is no single, obvious measure of what makes two passwords similar: Similar letters in the same positions? Common pet names or birthdays? Some other measure? Click-based graphical passwords however, have a natural proximity measure: the spatial distance between two points. As such, graphical passwords provide an excellent environment to explore and analyse user password choice, as well as approaches for enlarging the effective password space.

Usable authentication is an active research area but no method has yet emerged as the ideal solution. Text passwords are the most popular method of authenticating users in computer systems, but these suffer from security and usability problems. Improvements such as mnemonic passwords [18] and passphrases [17] have had limited success as they also suffer from predictability problems or their security has not been sufficiently studied. Biometric authentication systems [15] have also been proposed but these have a number of usability issues and privacy implications. For example, if an account is compromised in some way, it can be difficult to issue a new biometric to a user. Furthermore, it is difficult for users to create distinct identities for various parts of their life. Other methods of authentication include the use of tokens, such as smart cards, but these may be forgotten or stolen.

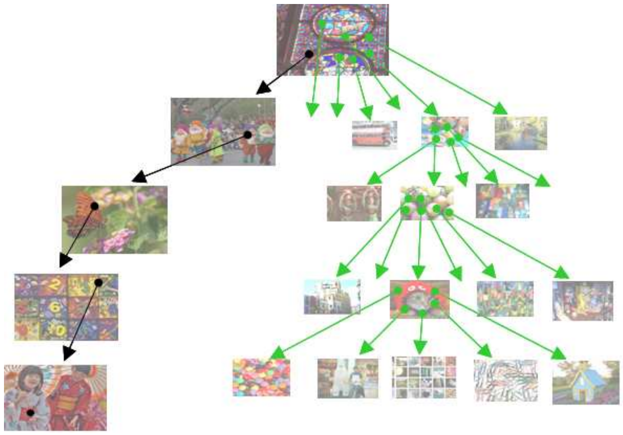


Figure 1: A user's navigation path through a sequence of images to form a password in CCP. Users click on one point per image and the current click-point determines the next image displayed.

2.1. Click-based graphical passwords

Graphical passwords offer an alternative to text-based passwords that is intended to be more memorable and usable because graphical passwords rely on our ability to more accurately remember images than text [20]. Several forms of graphical passwords have been proposed. Suo et al. [22] and Monrose and Reiter [19] offer overviews of various schemes and their design rationales. Of particular relevance is Jimini [23] where passwords are created by positioning a "template" over a background image so that the user's secret areas fall within the cut-out portions of the template. They found that users had difficulty remembering the position of their template and selected similar areas of the images.

We focus primarily on click-based graphical passwords. In PassPoints [29, 30], passwords consist of a sequence of five click-points on a given image. Users may select any pixels in the image as click-points for their password. To log in, they repeat the sequence of clicks in the correct order. Each click must be within a system-defined tolerance region of the original click-point. The usability and security of this scheme was evaluated by the original authors [9, 29, 30] and subsequently by others [3, 16, 25]. It was found that although relatively usable, security concerns remain. The primary security problem is hotspots: different users tend to select similar click-points as part of their passwords. Attackers who gain knowledge of these hotspots through harvesting sample passwords or through automated image processing techniques can build attack dictionaries and more successfully guess PassPoints passwords [9, 25]. A dictionary attack consists of using a list of potential passwords (ideally in decreasing order of likelihood) and trying each on the system in turn to see if it leads to a correct login for a given account. Attacks can target a single account, or can try guessing passwords on a large number of accounts in hopes of breaking into any of them.

To reduce the security impact of hotspots and further improve usability, we proposed an alternative click-based graphical password scheme called Cued Click-Points (CCP) [5]. Rather than five click-points on one image, CCP uses one click-point on each of a sequence of five images. The next image displayed is determined by the location of the previously entered click-point (Figure 1). The claimed advantages are that logging on becomes a true cued-recall scenario, wherein seeing each image triggers the memory of a corresponding click-point. Thus remembering the order of the click-points is no longer a requirement on users, as the system presents the images one at a time. CCP also provides implicit feedback claimed to be useful only to legitimate users. When logging on, if users suddenly see an image they do not recognise, they know that their previous click-point was incorrect. However, to an attacker without knowledge of the correct password, this cue is meaningless. Hotspots are still reported [5] in CCP, but because a very large pool of images can be used (as opposed to a single image per user in PassPoints), attackers must perform proportionally more work to gain useful information.

Visual attention research [31] shows that different people are attracted to the same predictable areas when looking at an image. This suggests that if users select their own click-based graphical passwords without guidance, hotspots will remain an issue. Davis et al. [7] suggest that user choice in all types of graphical passwords is unadvisable because users will always select predictable passwords. To the best of our knowledge, no research prior to the present paper exists on helping users select better graphical passwords, nor on how to avoid hotspots in click-based systems during password creation.

2.2. Persuasive Technology

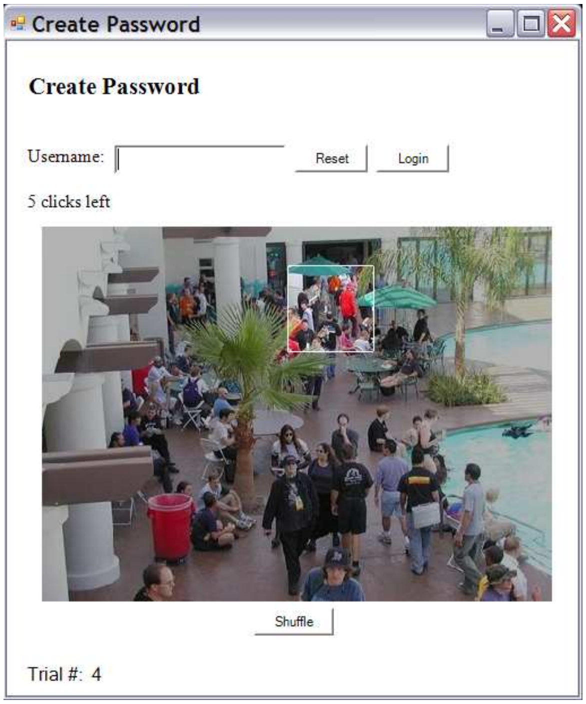
Persuasive Technology was first articulated by Fogg [11] as using technology to motivate and influence people to behave in a desired manner. He discusses how interface cues can be designed to actively encourage users to perform certain tasks. Forget et al. [12] propose how these may be condensed into a set of core persuasive principles for computer security.

An authentication system which applies Persuasive Technology should guide and encourage users to select stronger passwords, but not impose system-generated passwords. To be effective, the users must not ignore the persuasive elements and the resulting passwords must be memorable. As detailed in the next section, our proposed system accomplishes this by making the task of selecting a weak password more tedious and time-consuming. The path-of-least resistance for users is to select a stronger password (not comprised entirely of known hotspots or following a predictable pattern). As a result, the system also has the advantage of minimizing the formation of hotspots across users since click-points are more randomly distributed.

3. PERSUASIVE CUED CLICK-POINTS (PCCP)

Previous work [9, 16, 25] has shown that hotspots are a problem in click-based graphical passwords, leading to a reduced effective password space that facilitates more successful dictionary attacks. We investigated whether password choice could be influenced by persuading users to select more random click-points while still maintaining usability. Our goal was to encourage compliance by making the less secure task (i.e., choosing poor or weak passwords) more time-consuming and awkward. In effect, behaving securely became the path-of-least-resistance.

Using CCP [5] as a base system, we added a persuasive feature to encourage users to select more secure passwords, and to make it more difficult to select passwords where all five click-points are hotspots. Specifically, when users created a password, the images were slightly shaded except for a randomly positioned viewport (see Figure 2). The viewport is positioned randomly rather than specifically to avoid known hotspots, since such information could be used by attackers to improve guesses and could also lead to the formation of new hotspots. The viewport's size was intended to offer a variety of distinct points but still cover only an acceptably small fraction of all possible points. Users were required to select a click-point within this highlighted viewport and could not click outside of this viewport. If they were unwilling or unable to select a click-point in this region, they could press the "shuffle" button to randomly reposition the viewport. While users were allowed to shuffle as often as they wanted, this significantly slowed the password creation process. The viewport and shuffle buttons only appeared during password creation. During password confirmation and login, the images were displayed normally, without shading or the viewport and users were allowed to click anywhere.


Create Password

Create Password

Username:

5 clicks left

Trial #: 4

Figure 2: Screenshot of the PCCP Create Password interface with the viewport highlighting a portion of the image. (Pool image from [21])

Our hypotheses were:

1. Users will be less likely to select click-points that fall into known hotspots.
2. The click-point distribution across users will be more randomly dispersed and will not form new hotspots.
3. The login success rates will be similar to those of the original CCP system.
4. Participants will feel that their passwords are more secure with PCCP than participants of the original CCP system.

4. LAB STUDY

The methodology for the usability study was reviewed and approved by our university's ethics committee for psychological research. We tested Persuasive-CCP (PCCP) in a lab study with

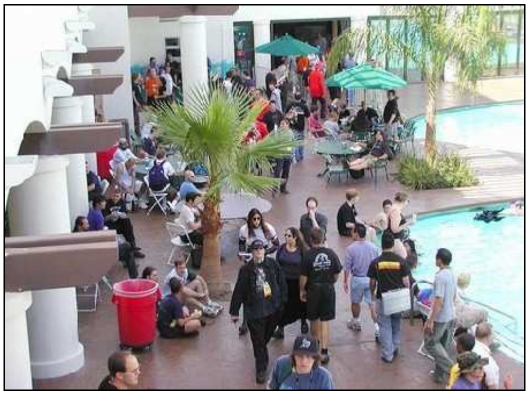


Figure 3: The Pool image [21]

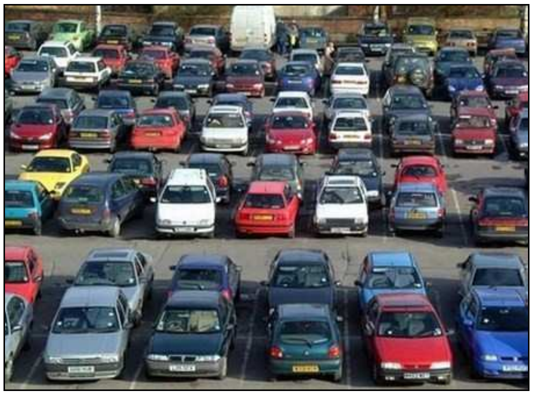


Figure 4: The Cars image [2]

39 participants who completed individual one-hour sessions. Participants ranged in age from 17 to 37. Most were university students from various fields. All were regular computer users who were comfortable with passwords and using a mouse. In total, data from 307 trials was collected. A trial consisted of a 5-step process that included creating, confirming, and logging on with a password.

The PCCP system was implemented in J# and ran on a Windows-based computer with a screen resolution of 1024x768. Consistent with previous PassPoints [3, 29, 30] and CCP [5] studies, the image dimensions were 451x331 pixels and the tolerance region was 19x19 pixels (the area around an original click-point accepted as correct since it is unrealistic to expect users to accurately target an exact pixel). We used the same set of 330 images as in the CCP study [5], including the 17-image subset used in the PassPoints lab study [3]. In our test system, the viewport was a 75x75 pixel square. System logs recorded the coordinates of the click-point on each image, the location of the viewport for each shuffle, and timestamps for each user action.

We used a between-participants design, with all participants from this study assigned to the viewport condition. For comparison, we used data collected from previous studies [3, 5] where participants created passwords without the viewport. The methodology, including instructions to participants, questionnaires, equipment, software (other than the addition of the viewport), and images were identical to those used for CCP. Both studies were conducted by the same researchers. Data collected from CCP can therefore be used as a control group against which to measure the effects of the viewport in PCCP.

Participants were first introduced to the system and told that they would be creating graphical passwords. They were further instructed to pretend these passwords were protecting their bank information, and thus should select passwords that were memorable but difficult for others to guess. They were told that the viewport was a tool to help them select more secure passwords, but that they could shuffle as many times as they wished to find a suitable click-point. Participants completed two practice trials (not included in the total count of 307 trials) to ensure that they understood how the graphical password system worked. They then proceeded to complete up to 10 further trials, as time allowed. A trial consisted of the following steps:

1. Create a password: Users selected one click-point on each of five different images. They could use the shuffle button to move the viewport until they found a desired click-point.
2. Confirm a password: Users re-entered their click-points. If they made an error, they could clear their clicks and try again. In cases where they absolutely did not know their password, they could reset, effectively returning to the first step.
3. Answer two questions: Users answered two on-screen questions about their current password, providing their opinion of how easy it was to create a password and how difficult it would be to remember it in a week.
4. Complete a Mental Rotations Test (MRT): Users spent at least thirty seconds completing an MRT puzzle [22]. This was primarily intended to simulate the passage of time and work as a distraction to clear visual working memory.
5. Log in: Users re-entered their password to log in. As with the Confirm phase, they could clear their click-points at any stage if they made a mistake or they could reset their entire password and return to the first step of the trial if they were unable to log in. If users were frustrated and could not use the given images, the interface allowed them to skip this trial and move to the next one.

Users also completed two questionnaires: a demographics questionnaire at the midway point and a final post-task questionnaire to complete the hour-long session.

5. DATA ANALYSIS & INTERPRETATION

To analyse PCCP's performance, we compared the data from this user study to the following three datasets collected in previous studies [3, 5]:

PassPoints-lab (PPLab): 43 participants tested a PassPoints system with 17 different images in a lab setting with the same methodology as this current study. At least 31 passwords (155 click-points) were collected on each image.

PassPoints-field (PPField): 376 participants used a PassPoints system for 7-9 weeks to access online notes for their class. Only the Pool (580 click-points) (Figure 3) and Cars (545 click-points) (Figure 4) images were used. These two images were selected from the set used in the PassPoints-lab study.

Cued Click-Points (CCP): 57 participants tested a Cued Click-Points system with the same set of 330 images and same methodology to this current study. 32 to 39 click-points were collected on each of the 17 core images from the PassPoints-lab study. Data was also collected on the remaining 313 images, but

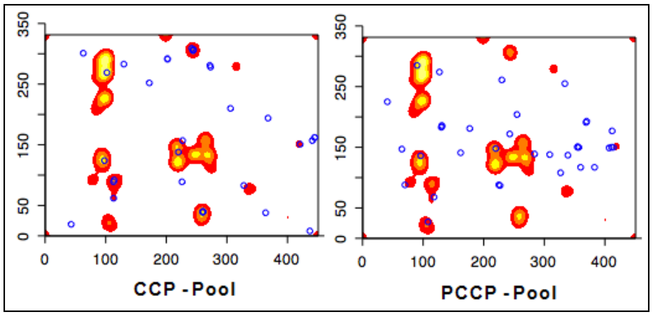


Figure 5: Displays individual click-points from CPP and PCCP respectively for the Pool image. The base heat map shows the location of known hotspots derived for the PassPoints-field dataset and thus is identical on both plots. [2] **(Best viewed in colour).**

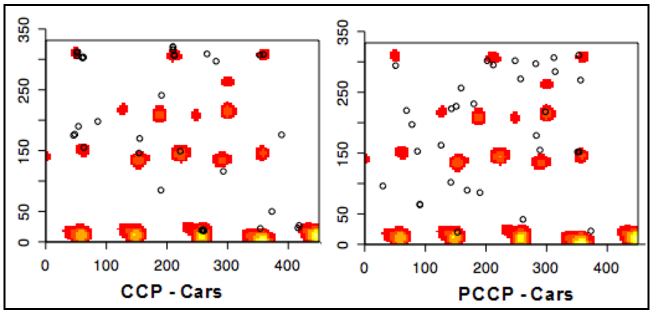


Figure 6: Displays individual click-points from CPP and PCCP respectively for the Cars image. The base heat map shows the location of known hotspots derived for the PassPoints-field dataset and thus is identical on both plots. [2] **(Best viewed in colour).**

since these were randomly displayed and only a small subset was seen by each participant, limited data was available.

We had the most data available for the two images used in the field study: the Pool image (Figure 3) and the Cars image (Figure 4). In most cases, the click-points collected in the PassPoints-field study will be used as the reference dataset since they were gathered in a realistic usage scenario and included the most samples.

Our data analysis examines several aspects of the system in order to address each of our previously stated hypotheses. We first look at the general usability of PCCP, then focus on the issue of hotspots, and finally discuss users' perception of the system.

5.1. Success rates and timings

As shown in Table 1, participants were able to successfully use PCCP. Success rates were calculated as the number of trials completed without errors or restarts, over all trials. As in earlier studies with click-based graphical passwords [3, 5], participants had some difficulty during confirmation while learning their password, but had little problem logging on afterwards. The success rates in Table 1 were calculated using the most stringent criteria: only passwords that were entered correctly on the first attempt without pressing the reset/clear button were considered successful. With a broader interpretation of "success", there are only 3 instances (99% success) where users were unable to eventually log in correctly and had to create a new password.

In comparison, CCP's [5] reported confirmation and login success rates were 83% and 96% respectively. We suspect that PCCP participants had more difficulty initially learning their password because they were selecting click-points that were less obvious than those chosen by PassPoints and CCP participants. However, PCCP participants were ultimately able to remember their passwords with a little additional effort. The login success rates of CCP and PCCP are not significantly different ($\chi^2(1,N=564)=0.07$, p=.796)[1], thus suggesting that the gain in security (reduced hot-spotting, as shown in Section 5.3) was not at the expense of usability.

Password creation was the longest of the three phases (Table 2). Users were progressively quicker with each re-entry. This is consistent with the pattern seen in the previous graphical password studies. We report the total time taken to complete a phase: from the time the first image was displayed to the time that they pressed the Login button, which included time spent thinking about their password. We also report the "click-time": the time taken from the first click-point to the fifth click-point. This represents the time taken to actually enter their password.

Table 1: PCCP success rates compared to CCP [5]

	Create	Confirm	Login
PCCP Success rate	305/307 (99%)	211/307 (69%)	278/307 (91%)
CCP Success rate	251/257 (98%)	213/257 (83%)	246/257 (96%)

Table 2: PCCP completion times for each phase (in seconds)

	Create	Confirm	Login
Total time: mean	50.7	29.9	16.2
Total time: median	41.4	18.9	14.0
Click-time: mean	36.3	24.9	10.6
Click-time: median	28.5	11.6	7.8

The CCP study [5] reports a median login click-time of 6.0 seconds which is faster than PCCP's 7.8 seconds. This difference is likely due to the slightly steeper learning curve from memorising a password that is not comprised of hotspots. However, PCCP participants did get progressively quicker and we speculate that comparable login times may be achievable with a few more login attempts.

[1] Results of the Chi-square (χ^2) test and other tests of statistical significance used within this paper are considered statistically significant when p < .05, indicating that the groups being tested are different from each other with at least 95% probability.

[2] The heat map is included to illustrate how many of the CCP and PCCP click-points fall near or within known hotspots.

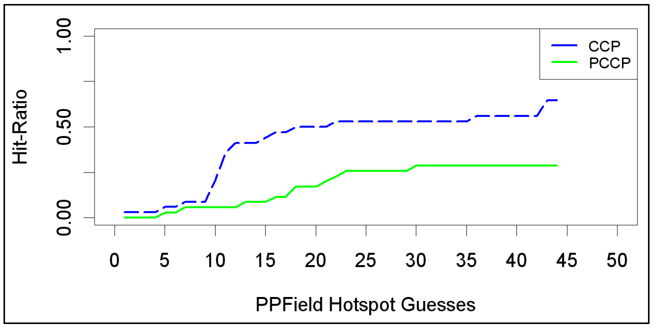


Figure 7: Individual click-points "guessable" using hotspots from the PassPoints-field study on the Pool image

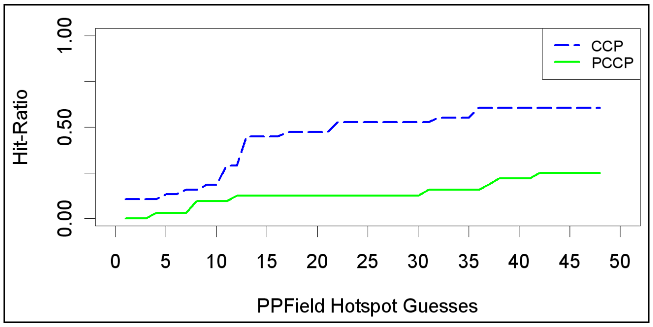


Figure 8: Individual click-points "guessable" using hotspots from the PassPoints-field study for the Cars image

5.2. Shuffles

The shuffle button was used moderately during password creation (Table 3). 63% of trials had 5 or fewer shuffles across all 5 images within a password (i.e., an average of at most 1 shuffle per image). We found that users who shuffled a lot had higher login success rates than those who shuffled little but the difference was not statistically significant (t(305)=1.89, p = .06).

Table 3: Effect of shuffles on success rates for 307 trials

Shuffles	# of trials	Login Success Rate
Low (0-5)	194 (63%)	89%
High (>5)	113 (37%)	94%

Most participants devised a shuffling strategy and used it throughout their session. They either consistently shuffled a lot at each trial or barely shuffled during the entire session. Those who barely shuffled selected their click-point by focusing on the section of the image displayed in the viewport, while those who shuffled a lot scanned the entire image, selected their click-point, and then proceeded to shuffle until the viewport reached that area. When questioned, participants who barely shuffled said they felt that the viewport made it easier to select a secure click-point. Those who shuffled a lot felt that the viewport hindered their ability to select the most obvious click-point on an image and that they had to shuffle repeatedly in order to reach this desired point.

5.3. Hotspots

The primary goal of PCCP was to increase the effective password space by guiding users to select more random passwords. To gauge our success, we therefore needed to determine whether PCCP click-points were more randomly distributed across the image and whether they successfully avoided known hotspots from previous studies.

To begin our analysis, we represented the click-point data graphically on the images themselves. The *PassPoints-field study* involving the Pool and Cars images yielded a large volume of data about where users clicked. We used a Gaussian kernel smoothed intensity function to summarise this data for each image [8]. We then created heat maps to depict this summary on the image area, using several colour bands to represent varying intensities of click-point concentration. The most intense areas thus correspond to hotspots. This heat-map of hotspots was used as the basis for comparing whether PCCP was better at avoiding known hotspots than CCP.

Figure 5 shows the heat map for the *PassPoints-field* click-points on the Pool image. White areas are the least click-point intensive and cover most of the image area. The five colour bands from red to yellow indicate progressively more intense areas thus revealing severe hotspots. The figure shows the same heat map twice: on the left, overlaid with the individual click-points (shown as small circles) from the CCP study (34 click-points), and on the right for our PCCP study (35 click-points). Figure 6 shows the corresponding information for the Cars image. Visually, it appears that PCCP click-points are more randomly distributed across the image, and not as concentrated on the heat map hotspots. As described below, we further tested to see whether this was true by conducting a dictionary attack on the click-points and by conducting some spatial statistics tests which confirm that PCCP click-points are more randomly distributed on the images.

To determine whether PCCP helped users avoid hotspots, we used the data from the earlier PassPoints-field study [3] to compile a list of hotspots for the Pool and Cars images. The PassPoints-field datasets included 580 click-points for Pool and 545 click-points for Cars. The hotspots were determined by finding the number of neighbouring click-points that were within tolerance of each click-point, sorting in decreasing order on this number of neighbours, then greedily assigning each click-point to the largest hotspot for which it was within tolerance. The result was a list of hotspot coordinates sorted in decreasing order by number of click-points they encompass.

We compared these hotspots to the click-points gathered for PCCP and CCP. Figure 7 and Figure 8 show the cumulative percentage of individual click-points that were "guessable" (i.e., the click-point fell within tolerance of a hotspot) for the Pool and Cars images respectively. PCCP click-points were much less likely to fall within hotspots than CCP's. For example, in the dataset for the Pool image (Figure 3) the 12 largest hotspots correctly identify 40% of CCP click-points but only 8% for PCCP. It should be noted that these are individual click-points, not passwords. An attacker would need to correctly identify all five of a user's click-points and images in order to successfully guess a password. For a more detailed discussion of security, see [5].

Due to the large set of images used in PCCP and CCP, we currently do not have hotspot information on all images and thus could not build an attack dictionary for entire passwords.

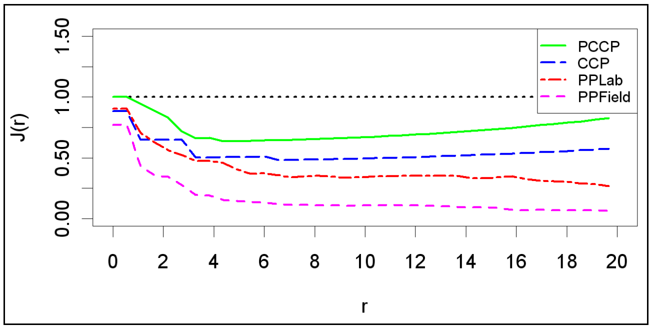


Figure 9: J-function showing amount of clustering at different radius values measured in pixels for PCCP, CCP, PassPoints-lab, and PassPoints-field on the Pool image.

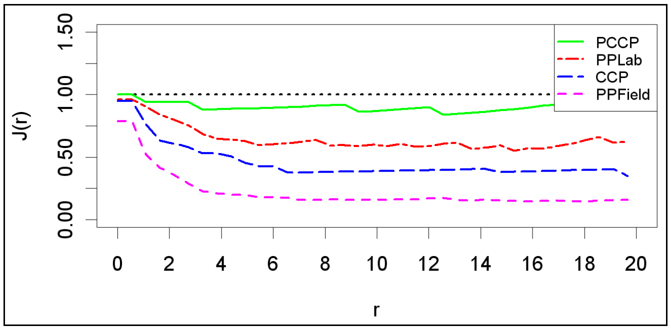


Figure 10: J-function showing amount of clustering at different radius values measured in pixels for PCCP, CCP, PassPoints-lab, and PassPoints-field on the Cars image

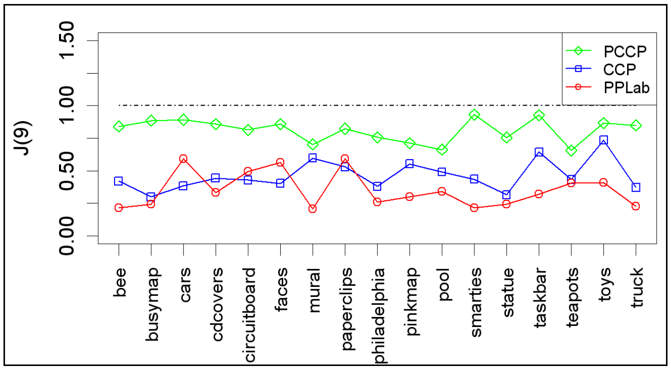


Figure 11: J-function at r=9 pixels for the set of 17 core images

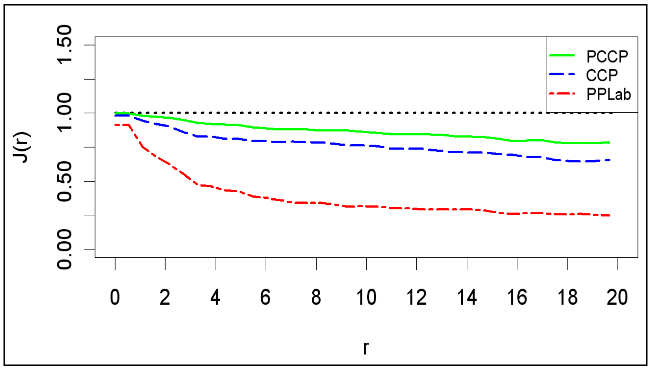


Figure 12: Cross J-function comparing PCCP, CCP, and PassPoints-lab to PassPoints-field reference dataset for the Pool image. PCCP is most dissimilar.

However, we can use the same method used in the CCP study [5] as an estimate. For CPP, the top 30 hotspots on an image cover approximately 50% of click-points (see Figure 7 and Figure 8). Assuming that a password consists of 5 click-points, the probability that a given password is found in an attack dictionary built from these hotspots would be 0.5^5 = 3%. For PCCP, the top 30 hotspots cover between 12% and 25% of click-points on the Pool and Cars images, so using an estimate of 20%, the probability that a password is in the same attack dictionary becomes 0.2^5 = 0.03%.

Standard statistical methods were inappropriate for this analysis because of the 2-dimensional nature of the click-point data. We instead applied point pattern analysis from spatial statistics [8] to measure the occurrence of hotspots and to evaluate whether click-points from the current PCCP study largely avoided hotspots established in the PassPoints-field study. We used the R programming language for statistical analysis and the *spatstat* package [1] to conduct our analysis.

To measure the level of clustering of click-points within datasets (the formation of hotspots), we used the *J-function* [26] statistic from spatial analysis. The J-function combines nearest-neighbour calculations and empty-space measures for a given radius r in order to measure the clustering of points. A result of *J* closer to 0 indicates that all of the data points cluster at the exact same coordinates, *J = 1* indicates that the dataset is randomly dispersed, and *J > 1* shows that the dataset is uniformly distributed. Ideally, we want the results to be near 1, indicating that the click-points are nearly indistinguishable from randomly generated points. Figure 9 and Figure 10 show that click-points on the Pool and Cars images are more randomly dispersed for PCCP than the other three datasets, indicating that the persuasive viewport was successful at guiding users to select more random click-points.

We further looked at the J-function measures at r = 9 pixels for the set of 17 core images. A radius of 9 approximates the size of the tolerance squares (19x19 pixels) used to determine whether a click was correct during password re-entry. Figure 11 shows that PCCP approaches complete spatial randomness for all 17 images (near J = 1) and is much more random than the CCP (t(15) = 9.85, p < .0001) and PassPoints-lab (t(15) = 11.70, p < .0001) datasets. A line graph was used for clarity, but in reality these are discontinuous points.

The *Cross J function* [27] is a multivariate summary statistic measuring the interaction between two spatial datasets. We use it as a measure of whether the PCCP click-points differ from those collected in previous click-based graphical password studies. Cross J close to 0 indicates that the two datasets are taken from the same population, Cross J = 1 shows that the datasets are distinct, and Cross J > 1 means that the datasets "repulse" each other. Figure 12 shows the Cross J values comparing each of the lab studies to PassPoints-field for the Pool image. The values for

PCCP are approaching 1, indicating that the PCCP dataset is distinct from the PassPoints-field reference set. Similar results were found for the Cars image. As results for PCCP are closest to 1, the Cross J function supports the assertion that the PCCP dataset is most dissimilar (among the three lab datasets) to our reference dataset of PassPoints-field.

5.4. User opinion and perception

A subset of the final questionnaire is reported here. The selected 10-point Likert-scale questions correspond to those reported in the previously cited studies [3, 5]. Users rated PCCP favourably (Table 4), with all median responses neutral or higher. They felt that PCCP passwords were easy to create and quick to enter, but they remained impartial on their preference between text and graphical passwords. Some of the questions were inverted to avoid bias (identified with a *). The scores for those questions were reversed prior to calculating the means and medians, thus higher scores always indicate more positive results for PCCP in Table 4

Table 4: Questionnaire responses. Scores are out of 10. The statements in parentheses provide the equivalent meaning for the reversed statement)

Question	Mean	Median
1. I could easily create a graphical password	8.0	8.0
2. ***** Someone who knows me would be better at guessing my graphical password than a stranger (i.e., when reversed: “someone who knows me would not be any more likely to guess my password than a stranger”)	7.0	8.0
3. Logging on using a graphical password was easy	6.4	7.0
4. Graphical passwords are easy to remember	6.0	6.0
5. ***** I prefer text passwords to graphical passwords (i.e., when reversed: “I like graphical passwords at least as much as text passwords”)	4.9	5.0
6. ***** Text passwords are more secure than graphical passwords (i.e., when reversed: “Graphical passwords are at least as secure as text passwords”)	6.2	6.0
7. I think that other people would choose different points than me for a graphical password	7.2	7.0
8. With practice, I could quickly enter my graphical password	8.3	8.0

We compared the two security-related questions (2 and 6) to the previous CCP responses to see if PCCP participants felt that their passwords were more secure. A Mann-Whitney (U) test was used to compare the sets of Likert-scale responses since they are comprised of ordered categorical data. The responses show that PCCP participants felt that their password would be equally difficult to guess for strangers or someone who knew them, while CCP participants were unsure (mean = 5.5, median = 5.0) (U = 675, p < .005). This may indicate that PCCP participants felt that their password did not contain personally identifiable characteristics, Also, PCCP participants felt that graphical passwords were at least as secure as text passwords while CCP participants were unsure (mean = 5.1, median = 5.0) (U= 723, p<.05).

It appears that users were aware that the viewport was helping to create more secure passwords and that the passwords were more random (i.e., less based on personal user choice). Several commented during the session that they were avoiding certain points because they were too obvious or too likely to be chosen by someone else and that the viewport was useful for helping them select a better click-point than they would have selected on their own. We speculate that in these cases users may be forming a more accurate mental model of the graphical password system and learning how to create stronger passwords. More research is needed to confirm this shift in users’ mental models.

5.5. Validation of hypotheses

We now revisit our hypotheses to evaluate whether to accept or reject them in light of the data analysis.

1. Users will be less likely to select click-points that fall into known hotspots when using the persuasive viewport.

   Hypothesis supported: This was confirmed by using known hotspots from the PassPoints-field data to attack the PCCP and CCP datasets. Click-points were significantly less predictable for PCCP (recall Figure 7 for Pool and Figure 8 for Cars), indicating that they did not fall within known hotspots.

   The Cross J-function results also provide statistical evidence that the PCCP dataset is distinct from the PassPoints-field dataset.

2. The click-point distribution across users will be more randomly dispersed and will not form new hotspots.

   Hypothesis supported: The results of the J-function tests show that the PCCP dataset is more random (less clustered) than the previous PassPoints-lab, PassPoints-field and CCP datasets.

3. The login success rates will be similar to those of the original CCP system.

   Hypothesis partially supported: The login success rates are slightly lower with PCCP, but the difference is not statistically significant. It may be that PCCP click-points require slightly more practice before being successfully memorised. Given that they avoid hotspots, it intuitively makes sense that less obvious areas of an image may require more attention to memorise. It may also be that since the image is initially dimmed during password creation, users had less chance to initially memorise the location of their point in reference to the remainder of the image. However, the learning curve appears acceptable as 99% of trials eventually ended with a successful login.

4. Participants will feel that their passwords are more secure with PCCP than participants of the original CCP system.

   Hypothesis supported: The questionnaire results show that PCCP participants felt that graphical passwords were at least as secure as text passwords and felt that their password was less personal because they believed that someone who knew them was no more likely to guess their password than a stranger.

6. DISCUSSION

Graphical passwords have some drawbacks as a form of authentication. They are susceptible to shoulder-surfing (i.e., when it is possible to observe or record someone entering their password to gain some or all of the details necessary to log in to their account). There is also some concern about interference [3] when users have to remember multiple graphical passwords. However, graphical passwords do offer an excellent environment for exploring strategies for helping users select better passwords since it is easy to compare user choices.

A common goal in authentication systems is to maximise the size of the effective password space. When user choice is involved, this also becomes a usability issue since users will be responsible for selecting their password. We have shown that it is possible to allow user choice while still increasing the effective password space. Furthermore, tools such as PCCP's viewport are only used during password creation so they cannot be exploited during an attack on an existing account. We could further deter users from selecting obvious click-points by limiting the number of shuffles allowed during the creation of a password or by progressively slowing system response in repositioning the viewport with every shuffle past a certain threshold. These approaches present a middle-ground between insecure but memorable user-chosen passwords and secure system-generated random passwords that are difficult for users to remember. While user choice is constrained with PCCP, the low number of shuffles indicates that users were willing to accept the system's suggestions and we believe that this design decision is justified by the increased security it offered and the apparently minimal usability drawbacks.

Providing instructions on how to create secure passwords, using password managers, or providing tools such as strength-meters to gauge the strength of a password have had only limited success [10]. The problem with such tools is that they require additional effort on the part of users who are creating passwords and often provide little useful feedback to guide the user's actions. In PCCP, creating a more secure password (by selecting a click-point within the first system-suggested viewport position) is the easiest course of action and requires little cognitive effort. Users still make a choice but they are constrained in their selection. Simplification and creating a path-of-least-resistance are both recommended strategies in Persuasive Technology for encouraging users to behave in the desired manner. PCCP demonstrates one possible application of Persuasive Technology [11,12] but other strategies could also be applied, even for graphical passwords.

The idea of guiding users during password selection can be extended beyond graphical passwords and we have some evidence that it would be useful in increasing the effective password space of text passwords as well. An analogous system to PCCP for text password might use a "hangman" or "Wheel-of-Fortune" strategy where new passwords are seeded with a few randomly assigned characters and users must fill in the remaining characters. For example, the system could offer

_ _ ! _ 9Q _ _

as a starting point. Here, the !, 9, and Q are fixed characters and users must choose the remaining characters of their password. Users could shuffle to get new randomly positioned and chosen characters if they were unable to create a password using the current suggestion. Such a system would reduce the occurrence of weak passwords consisting solely of dictionary or common words and would limit password re-use since any new password would also contain random characters. We expect that these passwords would be more memorable than system-assigned passwords since the user could personalise the password to some extent and would be engaging in its creation, which should help with memorisation. Initial pilot testing of such a system revealed that this particular approach may make it too difficult for users to create their passwords. They resorted to predictable patterns such as repeating the system-assigned characters. For example, a likely password for _ _ ! _ 9Q _ _ would be !!!99QQQ [13].

Instead, we allowed users to create their password normally then the system inserted a few random characters in random positions within the password. For example, if their original password was "fluffy", the strengthened password may become "f2luffRy". Users could shuffle to find a combination that seemed suitable, but again shuffling required time and effort. Users saw their modified password and re-entered it with the additional characters. Lab results indicate that this may be a viable approach [13] because the passwords are mostly user-created and the extra random characters increase their security. We speculate that users were able to visualize and remember their password in "chunks" with the inserted characters in between these chunks [14]. However, the more interesting question is whether the resulting passwords would be sufficiently memorable for long-term practical use. We cannot at present answer this question.

Another often cited goal of usable security is helping users form accurate mental models of security. Through questionnaires and conversations with participants in authentication usability studies, it is apparent that in general, users have little understanding of what makes a good password and how to best protect themselves online. Furthermore, even those who are more knowledgeable usually admit to behaving insecurely (such as re-using passwords or providing personal information online even though they are unsure about the security of a website) because it is more convenient and because they do not fully understand the possible consequences of their actions.

We believe that guiding users in making more secure choices, such as using the viewport during graphical password selection, can help foster more accurate mental models of security. Rather than providing vague instructions such as "pick a password no one will guess", we are actively showing users how to select a more random password as they perform the task.

Although these initial results are promising, further work is needed to test the long-term memorability of PCCP passwords, test the effect of interference when users must remember multiple passwords, and observe user behaviour in a real-world setting. A field study where participants use PCCP passwords instead of text passwords to access online resources over a few months (similar to [3]) would provide insight into these issues.

7. CONCLUSION

An important usability and security goal in authentication systems is to help users select better passwords and thus increase the effective password space. We believe that users can be persuaded to select stronger passwords through better user interface design. As an example, we designed Persuasive Cued Click-Points (PCCP) and conducted a usability study to evaluate its effectiveness. We obtained favourable results both for usability and security.

Graphical passwords provide a useful environment for testing such approaches because it is easier to determine the similarity of passwords and hence test for characteristics such as the occurrence of hotspots. However, we believe that these ideas could be adopted for text passwords as well, helping to increase the effective password space by encouraging users to behave more securely.

PCCP encourages and guides users in selecting more random click-based graphical passwords. A key feature in PCCP is that creating a secure password is the "path-of-least-resistance", making it likely to be more effective than schemes where behaving securely adds an extra burden on users. The approach has proven effective at reducing the formation of hotspots and avoiding known hotspots, thus increasing the effective password space, while still maintaining usability.

8. ACKNOWLEDGMENTS

We thank the participants of our lab study for their time and valuable feedback. We also thank the reviewers for their valuable feedback.

9. REFERENCES

1. Baddeley, A. and Turner, R. spatstat: An R package for Analyzing Spatial Point Patterns. Journal of Statistical Software, v12(6), 2005.
2. Britton, Ian. http://freefoto.com Last accessed Feb. 2007.
3. Chiasson, S., Biddle, R., and van Oorschot, P.C. A Second Look at the Usability of Click-Based Graphical Passwords. Symp. on Usable Privacy and Security (SOUPS) 2007.
4. Chiasson, S., van Oorschot, P.C., and Biddle, R. A Usability Study and Critique of Two Password Managers. 15th USENIX Security Symposium, 2006.
5. Chiasson, S., van Oorschot, P.C., Biddle, R. Graphical Password Authentication Using Cued Click-points. ESORICS 2007.
6. Cranor, L.F. and Garfinkel, S. (eds). *Security and Usability: Designing Secure Systems that People Can Use.* O'Reilley Media Inc, Sebastopol, CA, 2005.
7. Davis, D., Monrose, F., Reiter, M.K. On User Choice in Graphical Password Schemes. USENIX Security Symp. 2004.
8. Diggle, P.J. *Statistical Analysis of Spatial Point Patterns*. Academic Press: New York, NY, 1983.
9. Dirik, A.E., Memon, N., and Birget, J.C. Modeling user choice in the PassPoints graphical password scheme. Symp. on Usable Privacy and Security (SOUPS) 2007.
10. Florencio, D. and Herley, C. A Large-scale Study of WWW Password Habits. Proceedings of WWW 2007.
11. Fogg, B.J. *Persuasive Technologies: Using Computers to Change What We Think and Do*. Morgan Kaufmann Publishers, San Francisco, CA, 2003.
12. Forget, A., Chiasson, S., and Biddle, R. Persuasion as Education for Computer Security. AACE E-Learn 2007.
13. Forget, A., Chiasson, S., van Oorschot, P.C., and Biddle, R. Persuasion for Stronger Passwords: Motivation and Pilot Study. 3rd Int. Conference on Persuasive Technology, 2008.
14. Forget, A. and Biddle, R. Memorability of Persuasive Passwords (poster). ACM SIGCHI Student Research Competition, 2008.
15. Jones, L.A., Anton, A.I., and Earp, J.B. Towards understanding user perceptions of authentication technologies. ACM Workshop on Privacy in Electric Society, 2007.
16. Golofit, K. Click Passwords Under Investigation. ESORICS 2007. LNCS 4734, 343-358, 2007.
17. Keith, M., Shao, B., and Steinbart, P.J. The usability of Passphrases for authentication: An empirical field study, Int. Journal of Human-Computer Studies, 65(1), 17- 28, 2007.
18. Kuo, C., Romanosky, S., and Cranor, L. F., Human Selection of Mnemonic Phrase-based Passwords, Symp. on Usable Privacy and Security (SOUPS), 2006.
19. Monrose, F. and Reiter, M. Graphical Passwords. Chapter 9 in [6].
20. Nelson, D.L., Reed, U.S., and Walling, J.R. Pictorial Superiority Effect. Journal of Experimental Psychology: Human Learning and Memory 2(5), 523-528, 1976.
21. PD Photo. http://pdphoto.org/ Accessed August 2007.
22. Peters, M. Revised Vandenberg & Kuse Mental Rotations Tests: forms MRT-A to MRT-D. Technical Report, Department of Psychology, University of Guelph, 1995.
23. Renaud, K. and De Angeli, A., My password is here! An investigation into visuo-spatial authentication mechanisms, Interacting with Computers 16(6), 1017-1041, 2004.
24. Suo, X., Zhu, Y., and Owen, G.S. Graphical Passwords: A Survey. ACSAC 2005.
25. Thorpe, J. and van Oorschot, P.C. Human-Seeded Attacks and Exploiting Hot-Spots in Graphical Passwords. USENIX Security Symp. 2007.
26. van Lieshout, M.N.M. and Baddeley, A.J. A nonparametric measure of spatial interaction in point patterns. Statistica Neerlandica 50(3), 344-361, 1996.
27. van Lieshout, M.N.M and Baddeley, A.J. Indices of Dependence Between Types in Multivariate Point Patterns. Scandanavian Journal of Statistics 26, 511-532, 1999.
28. Whitten, A. and Tygar, J.D. Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0. 8th USENIX Security Symp. 1999.
29. Wiedenbeck, S., Birget, J.C., Brodskiy, A., and Memon, N. Authentication Using Graphical Passwords: Effects of Tolerance and Image Choice. Symp. on Usable Privacy and Security (SOUPS) 2005.
30. Wiedenbeck, S., Waters, J., Birget, J.C., Brodskiy, A., and Memon, N. PassPoints: Design and longitudinal evaluation of a graphical password system. Int. Journal of Human-Computer Studies 63, 102-127, 2005.
31. Wolf, J. Visual Attention. In Seeing, 2nd edition. K.K. De Valois (ed.). Academic Press, 335-386, 2000.