The Role of Visual Coherence in Graphical Passwords

Ülkü Arslan Aydın[1] (ulku.arslan@gmail.com) **Cengiz Acartürk[1] (acarturk@metu.edu.tr)**

Kürşat Çağıltay[2] (kursat@metu.edu.tr)

[1]Cognitive Science, Informatics Institute
[2]Computer Education and Instructional Technology
Middle East Technical University, 06800, Ankara, Turkey

Abstract

Graphical password is an alternative method of authentication to alphanumerical passwords. From the perspective of research on human memory, it is yet another novel technology that introduces challenges on human memory components. In this study, we aim to investigate the previous findings in human visual memory in the domain of graphical passwords by analyzing the role of visual coherence in passwords. The results of an experimental study reveal that in terms of memorability, coherent images are better candidates as graphical password images than jumbled images.

Keywords: Graphical passwords; visual coherence, visual working memory, eye tracking.

Knowledge-Based Authentication Systems

The extended use of human computer interfaces in the past few decades has introduced several challenges on users' working memory. One such challenge is the requirement to memorize numerous passwords for security authentication. From the viewpoint of information security, user access to a security system is granted in three phases: identification, authentication and authorization (Figure 1). After identification, the user supplies the proof of her/his identity in the authentication phase. The proof of identity is usually accomplished by employing methods such as using a smartcard (token-based authentication), using biometric information such as fingerprints (biometric-based authentication), or entering an alphanumeric or a graphical password (knowledge-based authentication).

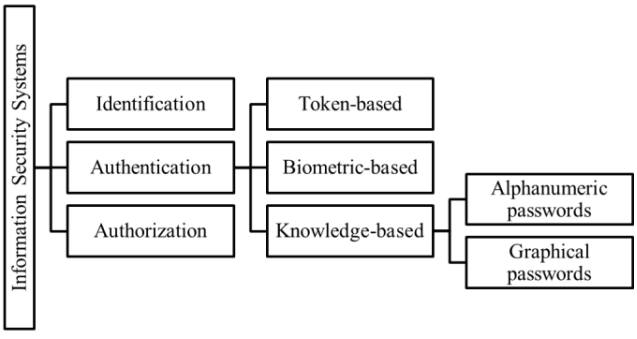

Figure 1: A taxonomy of authentication methods in information security systems.

Recently, knowledge-based authentication methods—in particular, text-based, alphanumerical passwords—are largely used for information access in information security systems (Herley et al., 2009). An alternative knowledge-based authentication method, namely graphical passwords, has been recently gaining an increased use.

Graphical Passwords as an Alternative Method to Alphanumerical Passwords

Graphical passwords were developed to overcome some of the security issues involved in the use of alphanumeric passwords (Dunphy et al., 2008). Graphical passwords are of different types, such as recall-based, recognition-based, and click-based (Figure 2). In a click-based graphical password system, a pixel-based image acts as a cue for activating user's memory. When creating a password in a click-based system, the user selects a sequence of number of (e.g., four or five) points on the presented image. After then, to login the system, the user reselects the points on the image in the same order by clicking on (or near to) them (Blonder, 1996; Wiedenbeck et al., 2005; Chiasson et al., 2007; Chiasson, 2008).

Figure 2: Sample graphical passwords (Dhamija & Perrig, 2000, Lashkari et al., 2009; Valentine, 1999).

From the end-user's point of view, the major motivation for the development of graphical passwords was to take the advantage of picture memorability over text while maintaining security (Wiedenbeck, et al., 2005), thus providing a solution to the security-usability dilemma.[1] The motivation for the development of graphical passwords finds its roots in early studies in cognitive psychology research, which revealed that humans have a tendency to

[1] The security- usability dilemma refers to the observation that "passwords are often either memorable-but-insecure or secure-but-difficult-to-remember" Chiasson, 2008, p. 3. Graphical passwords as a solution to the dilemma are beyond the scope of this study.

remember images longer and better than words (cf. the picture superiority effect, Nelson et al., 1976). Accordingly, images are usually expected to be "easier to remember and more secure than words" (e.g., Cranor, & Garfinkel, 2005; Kirkpatrick, 2002; Suo, Zhu, & Owen, 2005), thus leading to memorability advantages over alphanumerical passwords.

In addition to offering a more memorable solution for security system authorization, graphical passwords provide a naturalistic environment for research on visual memory in daily life tasks. Although the focus of research has been the security-usability dilemma from an information security point of view, there are many aspects that need further investigation from the perspective of cognitive science, such as the identification of the circumstances under which graphical passwords achieve better memorability. One such factor is visual coherence, as described below.

Visual Coherence in Graphical Passwords

Two major aspects of binding of objects in visual working memory are the binding of objects to perceptual features, such as color, shape and orientation, and the binding of objects to locations (Hollingworth & Rasmussen, 2010). In visual cognition, the concept of coherence has been studied by Biederman (1972) and Biederman, Glass and Stacy (1973), leading to research results which showed that the objects were recognized and identified more efficiently and quickly when the scene image was presented coherent rather than jumbled.[2] Mandler and colleagues have shown that the presence of a coherent background scene improves memory for both the location and the perceptual features of the object in the scene (e.g., Mandler & Parker, 1976; Mandler & Ritchey, 1977; Hollingworth, 2009). The facilitating effect of context in memory retrieval has been observed in both short-term time scale and long-term time scale (see Brady et al., 2011 for a review) (Brockmole et al., 2006; Foulsham et al., 2011). Those findings in visual working memory research suggest that visual coherence in graphical password images would improve memory for graphical passwords. In other words, when used as a graphical password image, a coherent image may reveal advantages over jumbled images. To test this hypothesis, we conducted an experimental study, in which the participants were shown how to create a click-based graphical password and how to login with the password, as described in the following section.

Experiment

In the practice session of the experiment, participants were guided by on-screen instructions about how to create a password. During the experiment, participants were presented a visual distraction task and then they were asked to login once. They were asked, however, to click on a black screen to login, instead of the previously presented graphical password image. This was the end of the first session. In the second session, three days after the first session, they were asked to login by using their password, again on the black screen. In both sessions, participants' login success and time were recorded. Participants' eye movements were recorded by a 50 Hz. non-intrusive eye tracker, integrated into 17'' TFT monitor.[3] The experiment was conducted in an office environment, with a developed application which simulated the interfaces of operating system that the participants were already familiar with. Overall, the experimental setting provided a relatively naturalistic environmental setting.

Participants, Materials and Design

Sixty-three participants (29 females, 34 males M= 32.1; SD= 0.73) participated in the experiment. All of the participants were employees at a governmental institution and the participation in the study was voluntary. The participants were divided into two groups, according to the type of the graphical password image they were presented in the password creation phase: (1) a coherent image or (2) a jumbled image. Each group was further divided into two groups according to the type of the image presented when participant failed to login on a black screen.: (1) the same image as the image presented in the password creating phase or (2) a shuffled version of the previous image. The base image for the graphical password was a high resolution (2362*2362) image taken in a professional setting. The image was converted into gray scale to reduce visual saliency effects due to color contrast. This image was used as the graphical password for the coherent-image group participants (henceforth, the coherent group). The jumbled image, which was used as the graphical password image for the jumbled-image group participants (henceforth, the jumbled group), was produced out of the coherent image by randomly jumbling the pieces of the coherent image, in the form of a 3x3 grid (Figure 3). There was at least one identifiable object in each cell of the grid.

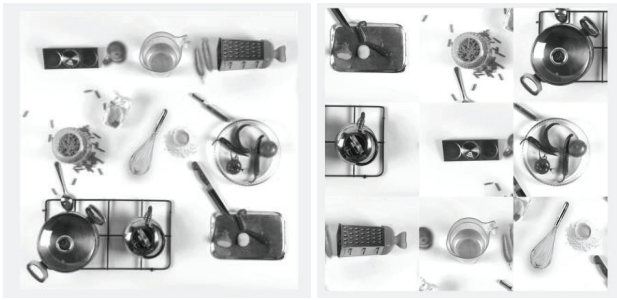

Figure 3: The images (600*600) for the graphical password in the coherent group (left) and the jumbled group (right).

[2] The jumbled image was created by dividing the coherent image into multiple sections and manipulating the relative positions of the sections without rotating them.

[3] The participants were seated at a distance of approximately 60 cm to the monitor. Spatial resolution and accuracy of the eye tracker was about 0.25° and 0.50° degrees respectively.

The participants were instructed to choose passwords which they could remember but that would be difficult for others to guess. They were guided by the instruction screens. The experiment consisted of two sessions. The first session was divided into five phases: practice, password generation, questionnaires, mental rotation task and login. In the practice session, the participant was shown how to create a graphical password and how to login with the selected graphical password. After the practice session, the participants picked their passwords by clicking four click-points on the provided image. After then, they filled out a demographic questionnaire and a usability questionnaire. A 30 second mental rotation task was then administered to disrupt visual memory. In the last phase of the first session, the participants were asked to use the selected password to login the system. The second session was a login session only; it was administered three days after the first session.

Results

In the last phase of the first session, the participants were asked to login the system by clicking on a black grid screen, without the graphical password image on the screen. This screen consisted of nine black squares in the form of a grid. The motivation for using the black screen was to investigate participants' strategy for choosing the password items. If the participant chose password just by memorizing object properties, without memorizing the spatial locations, s/he would not be able to log in without seeing the graphical password image. This was a surprise task for the participants because they were not informed about the black screen beforehand. The results showed that, however, the participants achieved a very high success login ratio on the black screen: Fifty-seven of sixty-three participants were able to login on the black screen, before being presented the graphical image (i.e., in the first, the second and the third attempt). This finding suggests that the participants remembered very well the locations of the click points in the first session. The results also suggested that a comparative analysis between the coherent-group participants (who were presented a coherent image as the graphical password) and the jumbled-group participants (who were presented a jumbled image as the graphical password) would be possible, because the results were similar between the pair-groups and the further division according to the type of the image presented at the login phase (i.e., shuffled vs. same) was no more necessary. Accordingly, the analyses were performed in terms of the measures below

- The time to login, create and confirm the password
- Eye movement parameters (fixation count, duration, Levenshtein distance) and visual saliency
- Password creation strategies

All the analyses were performed on participants' performance on the black grid screen in the first session (i.e., the login test in the same day) and in the second session (i.e., the login test three days after the first session). Additional analyses were also reported below, on visual saliency and on answers to questionnaires. Overall, the results suggested that coherent-group participants exhibited better memory performance compared to the jumbled-group participants, as presented below.

Login Success

The participants were allowed to try to login three times on the black grid screen. We performed a comparative analysis for the login success of the 56 participants in the first login attempt only. A three-way loglinear analysis (Login Success x Session x Group Type) produced a final model that retained login success and group type effects. This indicated that the interaction between login success and group type was significant, independently from the session. $\chi^2(1) = 5.20$, $p = .02$ (Figure 4). Based on the odds ratio, the odds of success in the first attempt was 2.98 times higher for the coherent-group than the jumbled-group participants.

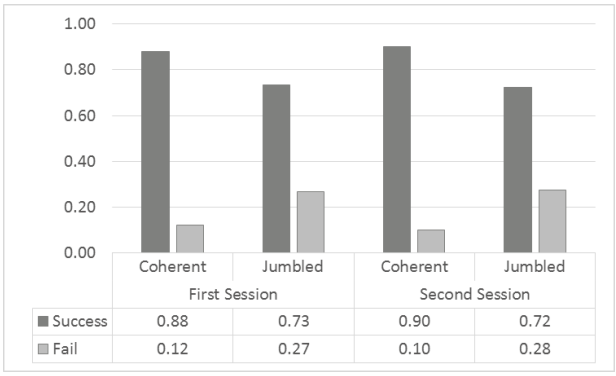

Figure 4: Login success of the participants in the first try (the numbers show the success and the failure ratios of the participants between 0 and 1)

The overall success ratio on the black screen, including the further attempts (up to three), revealed a similar finding, (i.e. the interaction between login success and group type was significant, independently from the session, $\chi^2(1) = 4.96$, $p = .04$), showing that coherent-group participants were more successful to login than the jumbled-group participants in the overall login attempts on the black screen.

Login Duration

The participants spent time to create the password and then to login on the first day of the experiment (i.e., the first session). In the first session, no difference was observed between the jumbled-group ($M = 25.2$ seconds, $SD = 15.7$) and the coherent-group participants ($M = 22.4$ s, $SD = 11.6$) in creating the password, $t(61) = 0.81$, $p = .42$, $r = .10$. Moreover, the time to login was not different between the jumbled-group ($M = 9.75$ s, $SD = 6.32$) and the coherent-group participants ($M = 7.85$ s, $SD = 3.82$). Although the participants spent approximately the same time to login between the first session ($M = 8.79$ s, $SD = 5.25$) and the second session ($M = 8.56$ s, $SD = 5.69$), the time spent to login in the second session was different between the jumbled-group ($M = 10.1$ s, $SD = 7.19$) and the coherent-

group participants (M = 7.07 s, SD = 3.19), $t(57) = 2.02$, p = .048, with a small effect size of r = .26. To sum up, the analysis of login durations showed that, in the second session of the experiment which was conducted three days after the first phase, the coherent-image group spent less time to login compared to the jumbled-group participants.

Fixation Counts

In this study, the term fixation count is used for describing the number of fixations on the black grid screen. The fixation counts were analyzed for a comparison between the jumbled group and the coherent group. The results were similar to the results obtained for login duration: there was no significant difference between the jumbled group (M = 18.3, SD = 12.3) and the coherent group (M = 14.4, SD = 8.83) in the first session. In the second session, however, the difference in fixation counts between the jumbled group (M = 18.6, SD = 16.6) and the coherent group (M = 11.1, SD = 6.57) was significant, $t(56) = -2.00$, p =.05, with an effect size of r = .26. There was also a significant main effect of the session in fixation counts, $F(1, 56) = 10.16$, p = .002, showing that the participants produced more fixation counts in the first session than they did in the second session. As the final step of the fixation count analysis, we investigated whether each fixation location belonged to the password (i.e., a pass item) or it did not belong to the password (i.e., a non-pass item). The participants in both groups spent more fixations on their pass items than their non-pass items, both in the first session, $F(1, 61) = 79.9$, p = .00, and in the second session, $F(1, 56) = 111.3$, p =.00. Moreover, in the second session, the coherent group spent less fixations on the non-pass items (M = 0.47, SD = 0.51) compared to the jumbled group (M = 1.63, SD = 2.07), $t(56) = 3.17$, p =.00. To sum up, in the second session, the jumbled group produced more frequent fixations compared to the coherent group. Moreover, the coherent group focused more efficiently on their pass items compared to jumbled group, who were focusing on non-pass items as well as pass items.

Fixation Durations

The term fixation duration is used in this study for the mean duration of single fixations on the black grid screen. There was no significant difference between the coherent group (M = 429.4 ms, SD = 95.2) and the jumbled group (M = 433.1 ms, SD = 145.7) in the first session. On the other hand, in the second session, the participants in the jumbled-image group had shorter mean fixation duration on the black grid (M = 453 ms, SD = 171) than the participants in the coherent-image group (M = 496 ms, SD = 125). This difference was significant $t(56) = -2.12$, $p < .04$ and it did represent small-sized effect r = .27. As the final step of the mean fixation duration analysis, we investigated whether each fixation location belonged to a pass item or it belonged to a non-pass item. The participants in both groups made longer fixations on their pass items than their non-pass items, both in the first session, $F(1, 56) = 30.5$, p =.00, and in the second session, $F(1, 56) = 50.8$, p =.00. Moreover, in the second session, mean fixation duration on pass items were similar for both the coherent group (M = 562 ms, SD = 259.8) and the jumbled group (M = 486 ms, SD = 152). On the other hand, the coherent group made shorter fixations on non-pass items (M = 228 ms, SD = 173) compared to the jumbled group (M = 373 ms, SD = 250), $t(56) = 3.33$, p =.00 in the second session. To sum up, the analysis of fixation durations revealed that in contrast to the similarities between the two groups in the first session, the jumbled group exhibited shorter fixations compared to the coherent group in the second session. Moreover, in the second session, the coherent group exhibited shorter fixations on the non-pass items.

Levenshtein Distance (LD)

The Levenshtein Distance (LD) is a specific application of the string editing analysis, where the distribution of fixations on certain locations (in our case, the grid cells) is coded by letters. The letter strings of each participant are then compared with the password of the participant for similarity. The LD defines the number of modifications (i.e., insertions and deletions) on one string that is necessary to make it the same as the other.

In our study, LD was used a specification of the similarity between the two groups of participants. The results of the LD analysis revealed that, the participants in the first session (M = 8.22, SD = 7.18) exhibited longer LD compared to the participants in the second session (M = 6.97, SD = 8.24), $F(1, 56) = 9.69$, p =.00. In addition, in the first session, no significant difference was obtained in LD between the jumbled-group participants (M = 9.35, SD = 7.96) and the coherent-group participants (M = 7.08, SD = 6.25). In the second session, the difference between the jumbled-group participants (M = 9.64, SD = 10.2) and the coherent-group participants (M = 4.30, SD = 4.29) was significant, $t(56) = 2.57$, p =.01, with a medium-size effect of r = .32, indicating more search effort in the jumbled-group participants compared to the coherent-group participants.

Visual Saliency Analysis

The saliency maps of the coherent image and the jumbled image were computed by using the algorithm provided by Walther and Koch (2006).[4] Based on this, the percentage distribution of the saliency of each cell in the 3x3 grid was calculated. The resulting distribution provided the relative saliency distribution over the password image. The distribution of participants' pass-items was also calculated by analyzing the selected graphical passwords in the experiment. For the jumbled image, no relation was obtained between the saliency values and the ratio of being pass-item, r = -0.11, p =.34 (Figure 5).

[4] SaliencyToolbox library, http://www.saliencytoolbox.net

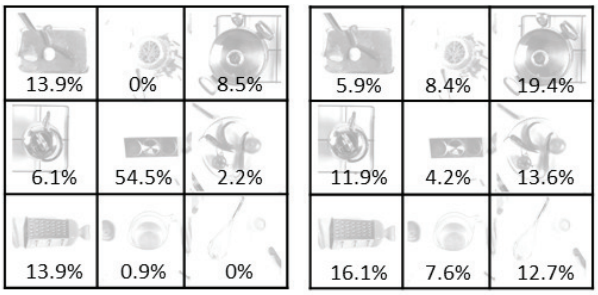

Figure 5: The saliency distribution in the jumbled graphical password (left) and the percentage of being selected as a pass-item by the participants (right).

A similar analysis was conducted for the coherent image. Again, no significant relation was obtained between the saliency values and the ratio of being pass-item, $r = 0.30$, $p = .47$ (Figure 6).

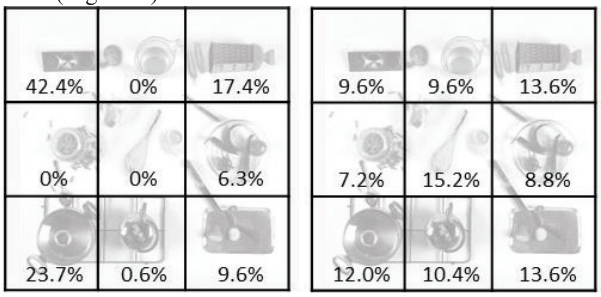

Figure 6: The saliency distribution in the coherent graphical password (left) and the percentage of being selected as a pass-item by the participants (right).

In summary, visual saliency analyses revealed no useful results to account eye movement behavior characteristics in graphical passwords, in line with the findings in relevant domains of visual cognition (Tatler & Vincent, 2009)

The Analysis of Password Creation Strategies

After the participants created the password in the first session of the experiment, they filled in a questionnaire about their strategy for creating the password. Four choices were presented to the participants: (a) *I created a pattern that looked like an L-shape or a V-shape* (create pattern), (b) *I memorized the names of the objects in the password* (object recognition) (c) *The objects I selected had common visual features (e.g., color, shape) or functional (e.g., cutting) features,* (similar features) (d) *I created a story* (story). The participants were allowed to make multiple choices. The participants' answers were analyzed in terms of the relation between the group type, the session, the adopted strategy and the login success. A main effect was obtained for strategy, $\chi^2(1) = 91.7$, $p = .00$, indicating that there was a significant difference between the adopted strategies. Pattern creation was the most preferred strategy (33 out of 77). Furthermore, the interaction between strategy and group type was significant, $\chi 2(3) = 8.61$, $p = .03$, indicating that adopted strategy was significantly affected by group type. On the other hand, no relationship was observed between the selected strategy and the login success, $\chi 2(3) = 4.46$, $p = .21$.

Discussion

The results of the experimental investigation showed that a high majority of the participants (57 of 63 participants) in the first session was able to login the system by clicking on a black screen. This finding indicates that the participants memorized the locations of the pass-items in the graphical password. The rest of the analyses were conducted on those 57 participants. Overall, the coherent-group participants, who were presented a coherent image as the graphical password, achieved better memory performance compared to the jumbled-group participants, who were presented a jumbled image as the graphical password. This finding was obtained in terms of a set of measures, including login success, login time, eye movement parameters and visual saliency, as well as password creating strategies. The analysis of login success showed that the coherent group exhibited higher login success compared to the jumbled group, independent of the session. This difference was obtained both in the first attempt to login and in the analysis of all attempts to login (the participants were allowed to try three times to login). The performance difference between the groups was evident, for some of the measures, in the second session of the experiment, which was conducted three days after the first session. For example, the analysis of the login duration showed that in the first session, there was no difference between the groups. The difference, however, was significant in the second session in favor of the coherent group: the coherent-group participants were able to login in shorter time. These findings have implications for end-users, as well as password system designers. The facilitating role of image coherence suggests that users should be encouraged to select coherent images for graphical passwords rather than jumbled images.

The analysis of fixation counts revealed two major findings: not in the first session but in the second session, the jumbled group fixated more frequently on the black screen compared to the coherent group. Moreover, in the second session, the coherent group spent less fixations on non-pass items, thus exhibiting a higher memory efficiency for the pass items. The analysis of fixation durations revealed that in the second session, the mean fixations of the jumbled group were shorter than the mean fixations of the coherent group. Shorter fixation durations may be indicators for visual search (compared to normal scene viewing, Rayner, 1998). Jumbled-group participants' higher effort to find the pass-items, as well as the longer Levenshtein distance exhibited by the jumbled group, provide support for our interpretation that they had more difficulty in remembering the pass items compared to the coherent-image group participants. Finally, we observed no relationship between likelihood of the selected pass items and their visual saliency. This may be due to participants' strategies in selecting the pass items. The analysis of the

strategies, however, returned no significant relationship between the selected strategy and the login success, though higher preference of certain strategies (in particular, pattern creation) by the participants over the others.

Conclusion and Future Work

Coherence has been a research topic in relevant domains to human cognition. In linguistics, discourse coherence is described as constructing the continuity in context by constructing the meaning between the parts of the written text or spoken utterance (Wolf, 2005). A coherent discourse has comprehension advantages compared to an incoherent discourse. In visual cognition, the studies reveal an improved efficiency in object identification and memory in favor of coherent images. These findings reveal the importance of coherence for cognition in different modalities. The findings in the present study show that the coherence effect is also applicable to practical settings, in this case graphical passwords. The present study also shows that the advantages of visual coherence can be observed in various measures, including login success and duration, as well as eye movement parameters. Future studies will address extending the evaluation by additional eye tracking metrics, such as scan path ratio, the investigation of the role of specific memory components, and a more extensive analysis of users' strategies for creating graphical passwords.

Acknowledgments. Thanks the Capital Markets Board of Turkey for providing the setting for the experimental investigation. Thanks Deniz Zeyrek and our reviewers for their valuable comments and suggestions.

References

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

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

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

Brady, T.F., Talia, K., & Alvarez, G.A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *J. Vision, 11*(5):4, 1-34.

Brockmole J.M., Castelhano, M.S., & Henderson, J.M. (2006). Contextual cueing in naturalistic scenes: Global and local contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(4), 699-706.

Chiasson, S. (2008). *Usable autentication and click-based graphical passwords*. Carleton University dissertation.

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

Cranor, L.F., & Garfinkel, S. (2005). *Security and usability: Designing secure systems that people can use*. L. Cranor & S. Garfinkel, (Eds.), O'Reilly.

Dhamija, R. & Perrig, A. (2000). Déjà Vu: A user study using images for authentication. *Proceedings of the 9th conference on USENIX Security Symposium*.

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

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

Herley, C., Van Oorschot, P., & Patrick, A. (2009). Passwords: If we're so smart, why are we still using them? In Dingledine &Golle (Eds.) *Financial cryptography and data security* (pp. 230-237). Springer.

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

Hollingworth, A., & Rasmussen, I.P. (2010). Binding objects to locations: The relationship between object files and visual working memory. *Journal of Experimental Psychology: Human Percept. and Perf., 36*(3), 543-564.

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

Lashkari, A.H., Saleh, R., Farmand, S., Zakaria, O. (2009). A wide range survey on recall based graphical user authentications algorithms based on ISO and attack patterns. *International Journal of Computer Science and Information Security*, 6(3), 17-25.

Mandler, J.M. & Parker, R.E. (1976). Memory for descriptive and spatial information in complex pictures. *Journal of Experimental Psychology: Human, Learning, and Memory, 2,* 38-48.

Mandler, J.M., & Ritchey, G.H. (1977). Long-term memory for pictures. *Journal of Experimental Psychology: Human Learning & Memory, 3,* 386-396.

Nelson, D.L., Reed, U.S., & Walling, J.R. (1976). Pictorial superiority effect. *Journal of Experimental Psychology: Human Learning & Memory*, 2, 523-528

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372-422.

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

Tatler, B. W., & Vincent, B. T. (2009). The prominence of behavioural biases in eye guidance. *Visual Cognition*, 17, 1029-1054.

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

Walther, D., & Koch, C. (2006), Modeling attention to salient proto-objects. *Neural Networks*, 19, 1395-1407.

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

Wolf, F. (2005). Coherence in natural language: Data structures and applications. MIT Dissertation.