Journal of Usability Studies
Vol. 13, Issue 4, August 2018 pp. 196–219

Usability Comparison of Over-the-Shoulder Attack Resistant Authentication Schemes

Ashley A. Cain
Old Dominion University
5115 Hampton Blvd.
Norfolk, VA
USA
acain001@odu.edu

Jeremiah D. Still
Old Dominion University
5115 Hampton Blvd.
Norfolk, VA
USA
jstill@odu.edu

Abstract

Graphical authentication schemes offer a more memorable alternative to alphanumeric passwords. However, they have been criticized for being susceptible to over-the-shoulder attacks (OSA). To solve this shortcoming, schemes have specifically been designed to be resistant to OSA. Common strategies used to decrease the ease of OSAs are grouping targets among distractors, translating them to another location, disguising the appearance of targets, and using gaze-based input. We are the first to provide a direct comparison of the common strategies regarding usability and OSA resistance. Specifically, we examined three OSA resistant graphical schemes, an eye tracker scheme, and a traditional alphanumeric password. To capture usability performance, we measured login times, learnability, memorability, satisfaction, acceptability, and error rates. OSA performance was examined to determine the relative resistance of each scheme. We found that graphical schemes are memorable after three weeks, and they were resistant to OSAs. Login time was acceptable for some schemes and not others. Learnability and satisfaction were disappointing, and error rates were high likely due to the novelty of these graphical schemes. Alphanumeric passwords offer the best learnability.

Keywords

graphical authentication, over-the-shoulder attack, cyber security, usability

Introduction

The frequency and cost of cybersecurity attacks, such as worm attacks, is increasing (Walters, 2014). The average cost of a breach in data has more than doubled since 2010 (Walters, 2014), with the cost reaching an average of $4 million per breach in 2016 (per a survey of 383 companies across 12 countries; IBM, 2016). For most attacks, a vulnerability in authentication must be exploited (Zviran & Haga, 1999). Authentication protects valuable or confidential information (e.g., company files, banking information, and health records) by requiring the user to confirm their identity. A user is granted access to a network or system if they can confirm something they know (e.g., a password), something they have (e.g., a token), or something they are (e.g., a fingerprint; Cazier & Medlin, 2006).

Alphanumeric passwords, a knowledge-based scheme, are the most commonly used authentication scheme (Grawemeyer & Johnson, 2011; Zviran & Haga, 1999). Alphanumeric passwords have widespread use because they are effective, efficient, subjectively satisfactory, and learnable. These passwords offer security against attacks, such as guessing attacks or worm attacks, when they have a large dimensional space (Zviran & Haga, 1999). They should be long and complex (Barton & Barton, 1984; Choong & Greene, 2016). They should not contain common words, and they should contain numbers and symbols (Barton & Barton, 1984; Choong & Greene, 2016). To be secure, alphanumeric passwords should not be written down, they should be different for every account, and they should be changed often (Barton & Barton, 1984). Users have difficulty applying the given rules for creating strong passwords, especially as guidance varies from system to system (Choong & Greene, 2016). In a recent study, Choong and Greene (2016) asked participants to classify passwords as whether or not they comply with a given set of rules. Although "special characters," "symbols," and "non-alphanumeric characters" have the same meaning, participants interpreted rules differently depending on how the rules are explained. Even when users apply password rules, stronger security can lead to trade-offs with usability. End users deal with limitations of usability by using "workarounds" and not using the system as it was intended to be used (Grawemeyer & Johnson, 2011). Long, complex strings of characters, symbols, and numbers are hard to remember (Zviran & Haga, 1999). Memorability is further hindered by the considerable number of passwords users have and by the need to routinely change passwords (Still, Cain, & Schuster, 2017). The security of alphanumeric passcodes is often undermined by users when they write them down or share passwords with loved-ones to solve problems with memorability (Grawemeyer & Johnson, 2011; Kaye, 2011; Paans & Herschberg, 1987). When forced to use unfamiliar alphanumeric passwords to bolster security, users are 18 times more likely to write them down (Grawemeyer & Johnson, 2011). A third of users report sharing their email password with someone else (Kaye, 2011). Users also undermine security to aid memorability by reusing passwords (Grawemeyer & Johnson, 2011). Up to 50% of alphanumeric passwords are reused (Grawemeyer & Johnson, 2011), and they are typically reused for 1.7 to 3.4 websites (Wash, Rader, Berman, & Wellmer, 2016). Although usability and security compete when selecting alphanumeric passcodes, alternate approaches to authentication may be able to provide both usability and security.

Graphical authentication schemes are knowledge-based approaches that utilize pictures as passcodes rather than complex strings of characters. Graphical passcodes offer a solution to the problem of memorability that accompanies the alphanumeric scheme (Biddle, Chiasson, & Van Oorschot, 2012). The pictures used in graphical passcodes are more easily remembered than the strings of characters used in alphanumeric passwords because pictures allow for a greater depth of cognitive processing. The picture superiority effect explains that pictures are dual encoded both visually and semantically, whereas alphanumeric passcodes are only encoded semantically (Paivio, 2013). Further, pictures typically have more features than individual letters and numbers, thereby also facilitating retrieval.

Although, graphical schemes offer memorial advantages they must also meet other usability needs for widespread adoption. First, authentication systems must allow for quick access (Still et al., 2017) comparable to login times for alphanumeric passwords (e.g., approximately five seconds; Wiedenbeck, Waters, Birget, Brodskiy, & Memon, 2005). Authentication is a secondary task that serves as a gateway to the primary goal, so authentication processes need to be quick. Login times for elaborate graphical schemes may not meet the need for quick access

(Sreelatha, Shashi, Anirudh, Ahamer, & Kumar, 2011) because it may take users longer to search for or recognize images. Second, to be usable, appropriate actions should be apparent to a wide range of users so that logins are successful with little training (Still et al., 2017). Confusion when authenticating will frustrate users when they are blocked from their primary goal. High error rates or poor learnability over time may reflect a lack of transparency for the actions needed to authenticate. The same issues could arise due to a lack of accessibility (Behl, Bhat, Ubhaykar, Godbole, & Kulkarni, 2014). Measurements of login times, error rates, and learnability can be used to index the usability of novel graphical schemes, and subjective satisfaction can reflect users' reactions to these objective dimensions.

Graphical schemes need to meet requirements of security as well as usability. One prominent susceptibility has been over-the-shoulder attacks (OSA). OSAs occur when an observer steals a passcode in a public place. Images associated with some graphical passcodes can be clearly observed on the screen. Just as users can quickly recognize and remember pictures in their passcodes, attackers may be able to casually peek at and reproduce the pictures. Vulnerability increases if an attacker has the opportunity to view a login more than once. Concern over OSA vulnerability has delayed broader deployment of these schemes. To overcome this concern, many graphical schemes have been designed to resist OSAs by allowing for non-direct selection of targets, by obscuring the appearance of the targets, or by obscuring target selection (Hayashi, Dhamija, Christin, & Perrig, 2008; Khot, Kumaraguru, & Srinathan, 2012; Wiedenbeck, Waters, Sobrado, & Birget, 2006). We have identified in the literature, four strategies that defend against OSAs: (a) grouping targets among distractors (Manjunath, Satheesh, Saranyadevi, & Nithya, 2014; Wiedenbeck et al., 2006) in which users can select a group of images rather than directly selecting targets, (b) translating targets to another location (De Luca, Hertzschuch, & Hussmann, 2010; Khot et al., 2012) in which passcodes are also obscured when users translate targets to another location rather than directly clicking them, (c) disguising targets (Cain & Still, 2016; Hayashi et al., 2008) such as by degrading images to interfere with an attacker's recognition of the passcodes, and (d) using gaze-based input (De Luca, Denzel, & Hussmann, 2009) to enable users to select targets using an eye tracking device, which is difficult for an attacker to observe.

To advance the study and eventual use of graphical schemes, it is important to consider how the OSA-resistant schemes compare to traditional passwords (e.g., De Luca et al., 2010; Sasamoto, Christin, & Hayashi, 2008), and how they compare to each other (e.g., Bulling, Alt, & Schmidt, 2012; Cain & Still, 2016; De Luca et al., 2009; Liu, Gao, Wang, & Chang, 2011). We conducted a study that directly compared the usability and security of prototypical OSA-resistant graphical schemes (grouping, translating to another location, disguising, and gaze-based input) and the alphanumeric scheme.

Prototypical Approaches

The following sections provide detail for the graphical authentication scheme prototypes used in our study.

Grouping

Previous schemes have provided resistance to OSAs by grouping targets among distractors to avoid direct selection of targets (Ankush, Dhanashre, & Husain, 2014; Behl et al., 2014; Chen, Ku, Yeh, Liao, 2013; Joshuva, Rani, & John, 2011; Kiran, Rao, & Rao, 2012; Sreelatha et al., 2011; Sun, Chen, Yeh, & Cheng, 2016; Manjunath et al., 2014; Rao & Yalamanchili, 2012; Tao, 2006; Vachaspati, Chakravarthy, & Avadhani, 2013; Zhao & Li, 2007). When targets are grouped with distractors and the user selects the group rather than the target, it is unclear to an attacker which images compose the passcode. The grouping strategy is often accomplished by allowing users to find three or more targets on a grid. Then rather than selecting each target, users select anywhere inside the region created by the targets (see Figure 1; Rajavat, Gala, & Redekar, 2015; Vachaspati et al., 2013; Wiedenbeck et al., 2006; Zhao & Li, 2007). Joshuva et al.'s (2011) scheme used the same strategy except that instead of targets and distractors being on a grid or a blank background, targets were points on an image. Other schemes present targets on a grid and users select a distractor that is at the intersection of the targets' locations (Behl et al., 2014; Sreelatha et al., 2011).

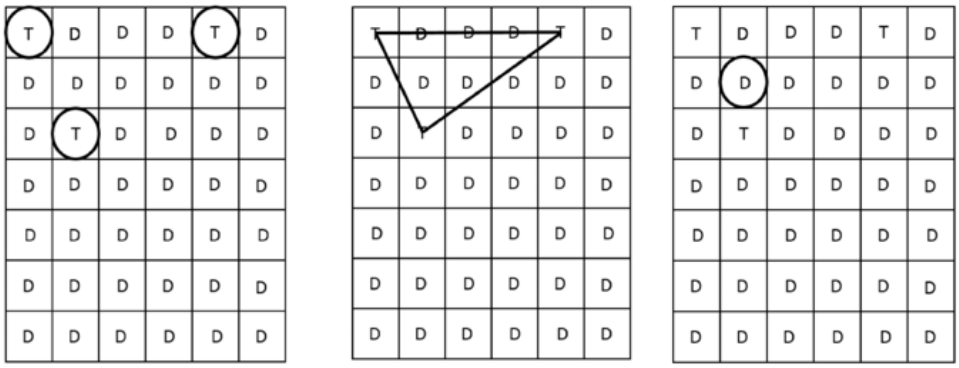

Figure 1. The interaction of grouping schemes (from left to right): find targets, mentally identify the shape, and select a distractor inside the shape.

Translating to another location

Other schemes offer resistance by allowing users to transfer targets elsewhere rather than clicking directly on them (Bianchi, Oakley, & Kim, 2016; Brostoff, Inglesant, & Sasse, 2010; De Luca et al., 2010; Gao, Liu, Dai, Wang, & Chang, 2009; Gupta, Sahni, Sabbu, Varma, & Gangashetty, 2012; Kawagoe, Sakaguchi, Sakon, & Huang, 2012; Kim et al., 2010; Lashkari, Manaf, & Masrom, 2011; Perkovic, Cagalj, & Rakic, 2009; Zangooei, Mansoori, & Welch, 2012). Similar to the grouping schemes, these schemes avoid the direct selection of targets. Some schemes number images on a grid and the numbers are selected elsewhere (Rokade, Hasan, & Mahajan, 2014; Van Oorschot & Wan, 2009). The same strategy was used by Liu, Qiu, Ma, Gao, and Ren (2011) and Sun et al. (2016) except that targets were points on a background image that was divided into numbered cells. Brostoff et al. (2010) and Zangooei et al. (2012) used the same strategy except that the numbers appeared after the grid of images. Other schemes allow users to select their targets using arrows or pressure bars on the side (Kim et al. 2010; Perkovic et al. 2009). Khot and colleagues (2012) allowed users to select the locations of their targets on a blank grid after performing a transformation (see Figure 2).

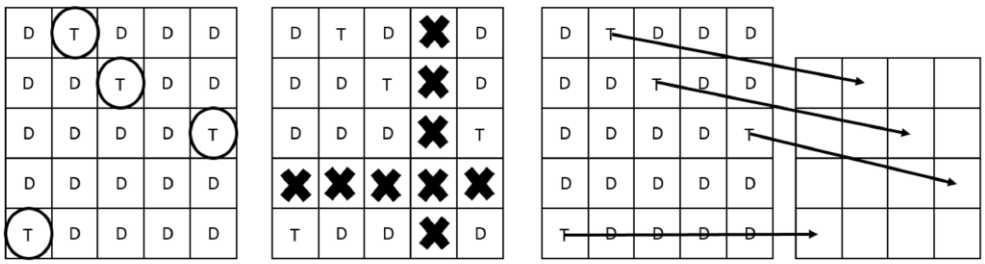

Figure 2. The interaction of a scheme that translates to another location (from left to right): find targets, mentally delete the row and column that do not contain targets, and click resulting locations of targets on the blank grid.

Disguising

Graphical schemes have been made resistant by disguising targets to interfere with attackers' recognition processes (Cain & Still, 2016; Gao, Guo, Chen, Wang, & Liu, 2008; Ghori & Abbasi, 2013; Hui, Bashier, Hoe, Kwee, & Sayeed, 2014; Jenkins, McLachlan, & Renaud, 2014; Lin, Dunphy, Olivier, & Yan, 2007; Liu, Gau et al., 2011; Meng & Li, 2013; Nicholson, 2009; Sasamoto et al., 2008; Yakovlev & Arkhipov, 2015; Zakaria, Griffiths, Brostoff, & Yan, 2011). Chakrabarti, Landon, and Singhal (2007) allowed for rotation of a free hand doodle, and Liu,

Gao, and colleagues (2011) allowed users to draw smaller free hand doodles. Zakaria et al (2011) disguised free hand doodles using line snaking, disappearing strokes, and decoy strokes. Sasamoto and colleagues (2008) and Hayashi and colleagues (2008) degraded images by removing detail and retaining general colors and shapes. Then the images were directly selected (see Figure 3). Cain and Still (2016) degraded line drawings by removing lines associated with intersections and curvatures.

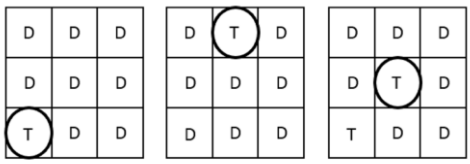

Figure 3. The interaction of a scheme that disguises targets (from left to right): find and select the first target, find and select the second target, and then find and select the third target.

Gaze-based input

OSAs can be protected against by having users authenticate using their eyes (Arianezhad, Stebila, & Mozaffari, 2013; Bulling et al., 2012; Dunphy, Fitch, & Olivier, 2008; Forget, Chiasson, & Biddle, 2010; Hoanca & Mock, 2006; Kumar, Garfinkel, Boneh, & Winograd, 2007). The spatial location of fixations on a display are harder to determine than a mouse cursor or touch locations. Schemes have allowed users to select faces on a grid (Dunphy et al., 2008), points on background images (Bulling et al., 2012), and patterns of dots using their eyes (De Luca et al., 2009; see Figure 4).

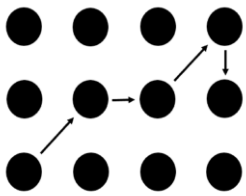

Figure 4. The interaction of a scheme that uses gaze-based input: fixate on the first, second, third, and fourth target. The arrows represent the scanpath target shape.

Needs Addressed by the Current Studies

Previous literature has offered many schemes for graphical authentication that are designed to be OSA resistant. Many schemes have been experimentally tested to determine their usability and security (Hayashi et al., 2008; Khot et al., 2012; Wiedenbeck et al., 2006). Schemes have also been compared with traditional PIN authentication (Bulling et al., 2012; De Luca et al., 2009). As different methods were used by different experimenters to assess each scheme, comparisons among schemes are difficult. For example, different amounts of training are given before experimental trials, and there are different lengths of delays before measuring memorability. OSA measures may allow for one or multiple viewings of the passcode, they may allow for one or multiple attempts to identify a passcode, and they may be motivated by reward or not. Satisfaction was measured by a variety of methods.

Limited previous studies have directly compared graphical schemes. Schaub, Walch, Könings, and Weber (2013) compared five graphical schemes that allow for authentication on small touch screen devices using strategies of recall and cued-recall. The schemes were also compared with the PIN scheme. Use Your Illusion (UYI) was the only scheme included in Schaub and colleagues' analysis that was designed to be resistant to OSAs. They found that the graphical schemes had similar usability to the PIN scheme, and they were more resistant to OSAs on small touch screens than the PIN scheme.

Johnson and Werner (2008) compared the memorability of four graphical passcodes and an alphanumeric password after 30 minutes had passed, and then again after one week. The prototypes of graphical schemes the researchers included combined an image with a background, had grids of faces, had grids of images, and had one large image with click points. All of the graphical schemes were more memorable than the alphanumeric scheme.

Our current studies provide a direct comparison of prototypical OSA resistant passcodes and an alphanumeric passcode. Convex Hull Click (CHC; Wiedenbeck et al., 2006) represented graphical passcodes that are made resistant to OSAs by allowing users to authenticate without clicking directly on the targets. What You See is What You Enter (WYSWYE; Khot et al., 2012) represented graphical passcodes that are made resistant to OSA by translating targets to another location. UYI (Hayashi et al., 2008) represented a group of graphical passcodes that disguise targets. Eye-Pass Shapes (De Luca et al., 2009) represented passcodes that are entered using gaze to obstruct OSAs. We performed a within-subjects runoff among these schemes on a variety of measures, including error rates, login times, learnability, memorability, OSA performance, satisfaction, and acceptability. We aimed to determine whether these schemes can be correctly entered and learned, whether they can have appropriate login durations, and whether they can be memorable compared with alphanumeric passwords. And, we aimed to determine whether graphical passcodes can be resistant to OSAs. We developed two separate studies—Study 1 and Study 2—to gather data for this paper. Study 2 had a two-phase design to assess the memorability of the tested schemes. The following sections provide detail about each study.

Study 1: Usability and Security Runoff

The following sections describe the particulars of Study 1, including the methods used and results. Study 1 compared four OSA-resistant graphical schemes and the alphanumeric scheme on dimensions of usability and security. Usability dimensions included error rates, login times, learnability, and acceptability. OSA performance was measured for security.

Method

The following sections describe the methods used for Study 1, which include discussion about the participants, stimuli and apparatus, and the procedure used in this phase of the study.

Participants

Twenty undergraduate students participated (females = 11, males = 9). They were recruited through the SONA system and compensated with class research credit. One participant reported being left hand dominant. Ages ranged from 18 to 53 (M = 23.05, SD = 8.60). Reported computer use ranged from three to 15 hours a day (M = 7.2, SD = 3.28). All participants reported normal or corrected to normal vision.

Stimuli and Apparatus

We created five prototypes of authentication schemes for this study. Four graphical schemes were based on Eye-Pass Shapes (De Luca et al., 2009), CHC (Wiedenbeck et al., 2006), UYI (Hayashi et al., 2008), and WYSWYE (Khot et al., 2012). These four schemes were compared to an alphanumeric scheme. CHC, UYI, WYSWYE, and alphanumeric prototypes were presented on a Windows desktop computer with a 24-inch monitor. The gaze-based scheme was presented on a Windows desktop computer with a 16-inch monitor.

CHC

In the current study, the prototypes of CHC, UYI, and WYSWYE were created in Paradigm©. Paradigm recorded selection locations on the grids and login times. CHC consisted of icons on a 10 x 15 grid (see Figure 5). The icons came from an online, open source database (http://www.fatcow.com/free-icons). The grid was 4138 x 1126 pixels. Each icon was 55 x 45 pixels. The passcode consisted of three system-assigned icons. Because there were three target icons, they would always form a triangle shape on the grid (see Figure 6). Target icons were never located in a straight line. A correct login occurred when a participant selected one time anywhere inside of the triangular region created by the three icons. They were told not to click directly on target icons and not to hover the mouse cursor over their target icons. Verbal

feedback of correct or incorrect selections was provided by the researcher after each authentication attempt. After each attempt, the icons were repositioned.

Figure 5. Prototype of CHC.

Figure 6. Three target icons form a region.

WYSWYE

The interface for WYSWYE showed a 5 x 5 grid of images on the right side of the screen. The grid of images was 715 x 549 pixels. Each image was 139 x 103 pixels. A blank 4 x 4 grid was on the left side (see Figure 7). The blank grid was 578 x 459 pixels. The blank cells were 139 x 103 pixels. A passcode consisted of four system-assigned images. Participants logged in by mentally deleting a row and column that does not contain a target on the 5 x 5 grid. They would mentally shift the remaining cells together and click the resulting locations of their four targets on the blank grid. Verbal feedback of correct or incorrect selections was provided by the researcher after each authentication attempt. The images were repositioned for every attempt.

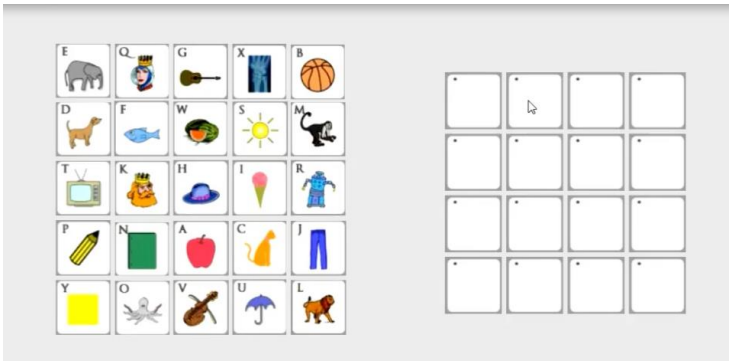

Figure 7. Prototype of WYSWYE.

UYI

UYI was presented as images in a 3 x 3 grid that were degraded by removing detail but retaining general colors and shapes (see Figure 8). The grid was 774 x 571 pixels. Each image was 213 x 175 pixels. A passcode consisted of three system-assigned images. A correct login occurred when a participant selected the degraded versions of each of their three targets on three subsequent grids. Verbal feedback of correct or incorrect selections was provided by the researcher after each authentication attempt. After each attempt, the images were repositioned.

Figure 8. Prototype of UYI.

GAZE-BASED SCHEME

The gaze-based prototype based on Eye-Pass Shapes consisted of a 3 x 4 dot configuration with an enter button on the upper right (see Figure 9). The grid was 513 x 379 pixels. Each dot had a radius of 50 pixels, and the selection area for each was a 120 x 120-pixel square. The interface was implemented using HTML. The Internet browser was Firefox. An Eye Tribe© eye-tracker was used to control the mouse cursor. Java code made the mouse cursor invisible while over the grid of dots. Java code measured the time of every selection of the enter button and determined if it was correct. The time and feedback of correct or incorrect selections were presented below the grid of dots. Dragger© made a selection every .7 seconds at the locations of the invisible mouse cursor. Minor movements of the mouse cursor were controlled by a jitter box of 22 pixels. The passcode consisted of four dots in a sequential order. Every participant used this same system-assigned passcode. A researcher recorded the time of all attempts to login.

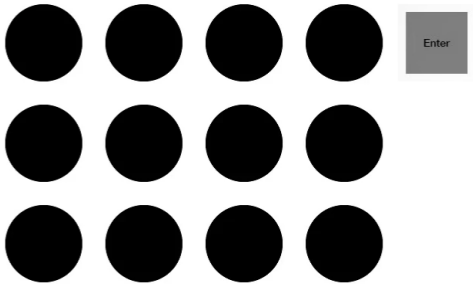

Figure 9. Prototype of a gaze-based graphical scheme.

ALPHANUMERIC

The alphanumeric interface consisted of a box for text entry and an enter button. It was implemented with HTML and run in Firefox. Java code captured login times. It displayed them below the primary interface frame, and the researcher recorded them. The passcode was system-assigned. The passcode, "col2Wlan6" was complex, containing nine characters, no dictionary words, a capitol letter, and two numbers.

Procedure

Participants were run individually. They were seated in front of a desktop computer. The researcher explained that they would be authenticating using five different schemes, and participants signed a consent form after being allowed time to ask questions. The schemes were as follows: gaze, CHC, UYI, WYSWYE, and alphanumeric. The order of the schemes was counterbalanced using a Latin Square design across participants. For each scheme, participants were given instructions, and they had one practice trial. After instructions, the experimenter would answer questions at any time. The experimenter would tell the participants whether they had correctly authenticated on every trial to provide them with feedback. Participants completed nine experimental trials of CHC, WYSWYE, and alphanumeric. Because three challenges are necessary to log in with UYI, participants logged in three times with UYI. Participants logged in four times with the gaze-based scheme.

After each set of experimental trials, participants took on the role of a casual attacker using OSAs for each of the graphical schemes and the gaze-based scheme. They viewed a video of the researcher logging in one time. The video only showed the screen, including mouse movements as it appeared while the researcher logged in. Then they circled the passcode they thought they observed on an answer sheet. They viewed the same passcodes being entered two more times, and they made another attempt to identify the passcode on the answer sheet.

Results

Dependent variables were the following: error rates, login times, learnability, OSA performance, and acceptance. Post hoc power analyses were conducted in GPower© for two-tailed repeated measures ANOVAs. For error rates, login time, learnability, and OSA performance, effect sizes were entered (.8 for large, .5 for medium, and .2 for small; Cohen, 1992), alpha (.05), and sample size (18). The likelihood of detecting a large and a medium effect for these measures exceeded 99%. The likelihood of detecting a small effect size was 57% for error rates and login times, and 51% for learnability and OSA performance. There was adequate power for a large or medium effect. Bonferroni corrections were used for all post hoc comparisons.

Error rates

Error rates were calculated as the number of incorrect trials of the total experimental trials for each scheme. A repeated measures ANOVA (authentication scheme: CHC, UYI, WYSWYE, gaze-based, and alphanumeric) was conducted to explore error rates. Sphericity was violated, so a Greenhouse-Geisser correction was applied. Differences were found among error rates (see Figure 10), $F(2.49, 37.29) = 9.02$, $p < .001$, partial $\eta^2 = .376$. Post hoc comparisons revealed that CHC ($M = .18$, $SD = .16$) was entered with fewer errors than the gaze-based scheme ($M = .42$, $SD = .20$), and alphanumeric passcodes ($M = .01$, $SD = .04$) were entered with fewer

errors than passcodes for UYI (M = .34, SD = .40), WYSWYE (M = .28, SD = .23), and gaze-based, $p < .05$ for all comparisons. No error rate differences were found between gaze-based scheme and UYI, p = 1.00, or WYSWYE, p = .55, and no differences were found between CHC and UYI, p = .32, WYSWYE, p = 1.00, or alphanumeric passcodes, p = .29. No differences were found between UYI and WYSWYE, $p > .99$. Participants made more errors using the UYI, WYSWYE, and the gaze-based schemes compared to the alphanumeric scheme. Of the graphical schemes, CHC was the only one that did not have more errors than the alphanumeric scheme.

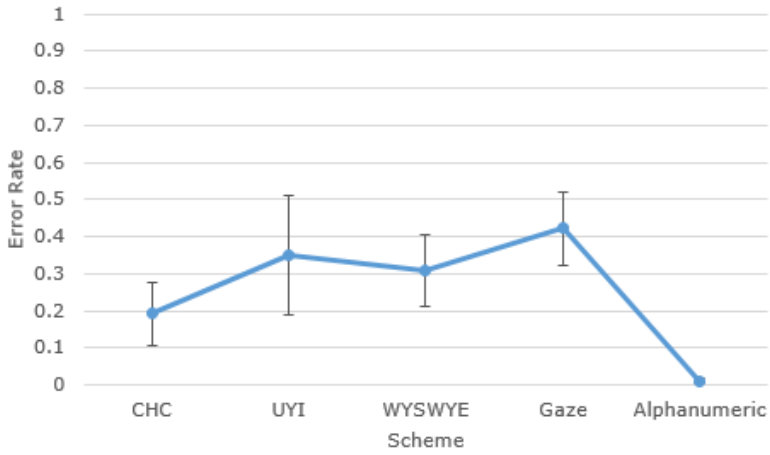

Figure 10. Error rates. Error bars represent 95% confidence intervals.

Table 1. Descriptive Statistics for Scheme on Error Rates

Variable scheme	***N***	***M***	***SD***
CHC	20	0.18	0.16
Gaze-based	20	0.42	0.20
UYI	20	0.34	0.40
WYSWYE	20	0.28	0.23
Alphanumeric	20	0.01	0.04

Login times

Login times were calculated for correct, experimental trials. Login times were cleaned for CHC, UYI, WYSWYE, and alphanumeric schemes by removing outliers that were 2.5 standard deviations above or below an individual participant's mean for that scheme. No outliers were removed for CHC or UYI. Two outliers were removed for WYSWYE and three for alphanumeric. No outliers were removed for the gaze-based scheme. A repeated measures ANOVA (authentication scheme: CHC, UYI, WYSWYE, gaze-based, and alphanumeric) was conducted to explore login times. Sphericity was violated, and we used a Greenhouse-Geisser correction. Differences were found among login times (see Figure 11), $F(2.21, 70.59) = 12.34$, $p < .001$, partial $\eta^2 = .278$. Post hoc comparisons revealed that UYI (M = 20.22, SD = 10.73) had longer login times than CHC (M = 10.97, SD = 5.74), gaze-based (M = 11.98, SD = 6.87), and alphanumeric schemes (M = 6.81, SD = 3.24), and WYSWYE (M = 15.84, SD= 11.76) had longer login times than the alphanumeric scheme, $p < .05$ for all comparisons. No differences were found between login times for the gaze-based scheme and CHC, $p > .99$, WYSWYE, p = .62, or alphanumeric passcodes, p = .13. No differences were found between login times for CHC and WYSWYE, p = .19, or the alphanumeric schemes, p = .44. No differences were found

between UYI and WYSWYE, p = .34. Alphanumeric passcodes were entered efficiently. However, the gaze-based scheme and CHC also had acceptable login times.

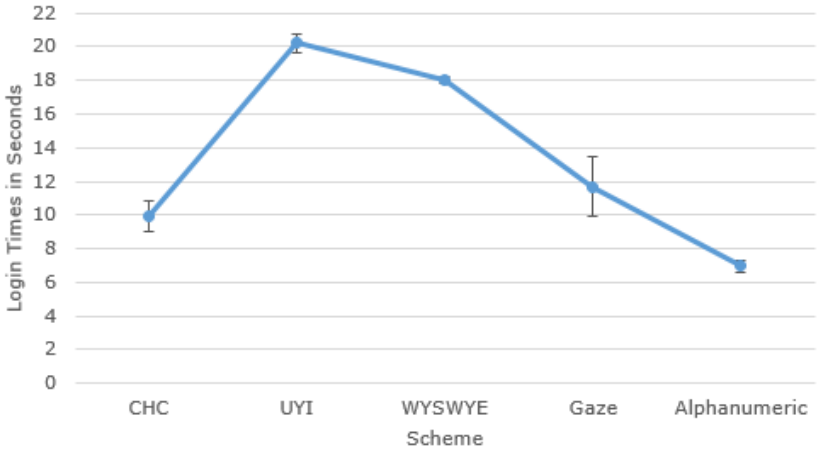

Figure 11. Login times. Error bars represent 95% confidence intervals.

Learnability

The correct attempts over time were measured to reflect learnability. The first three trials for each scheme including the practice trial composed the first rate for learnability, the second set of three interactions composed the second, and the third set of three interactions composed the third. The gaze-based scheme was not coded for learnability because it had fewer trails than the other schemes. In order to include UYI in learnability and allow for equivalent practice among schemes, the challenges were considered individually instead of in sets of three. A 3 (number of attempts: first set, second set, and third set) x 4 (authentication scheme: CHC, UYI, WYSWYE, and alphanumeric) repeated measures ANOVA was conducted to explore learnability. Sphericity was violated, so a Greenhouse-Geisser correction was employed. Main effects revealed differences among correct logins for the schemes and differences among correct logins over time, $p < .001$. The main effects were qualified by a two-way interaction between learnability and scheme, $F(3.65, 69.41) = 2.61$, $p = .021$, partial $\eta^2 = .121$. When using UYI, participant performance with UYI improved with practice, whereas it did not with the other schemes (see Figure 12 and Table 2). Because there was an interaction, the researchers tested for simple effects. Post hoc comparisons revealed differences between the first set and second set and between the first set and third set, $p < .05$ for both comparisons. No differences were found between the second and third set of attempts, $p = .09$.

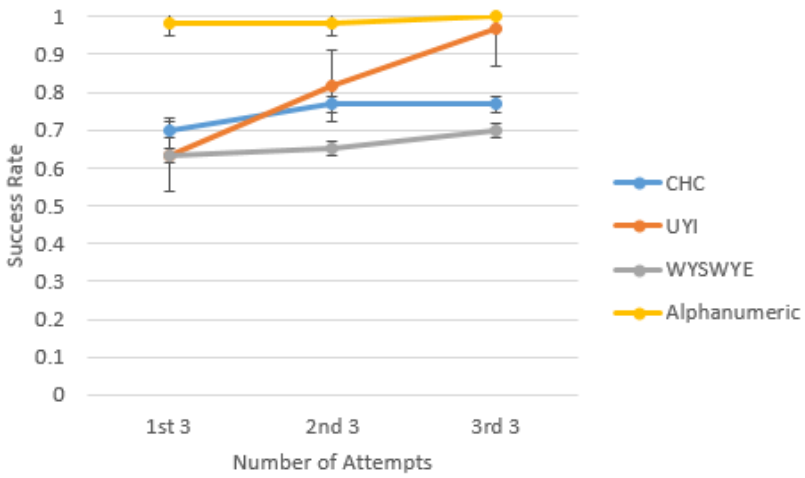

Figure 12. Learnability. Error bars represent 95% confidence intervals.

Table 2. Descriptive Statistics for Learnability and Scheme on Success Rates

Variable scheme	**Learnability**	***N***	***M***	***SD***
CHC	1st set	20	0.70	0.31
	2nd set	20	0.77	0.25
	3rd set	20	0.77	0.25
UYI	1st set	20	0.63	0.29
	2nd set	20	0.82	0.30
	3rd set	20	0.97	0.10
WYSWYE	1st set	20	0.63	0.29
	2nd set	20	0.65	0.37
	3rd set	20	0.70	0.29
Alphanumeric	1st set	20	0.98	0.07
	2nd set	20	0.98	0.07
	3rd set	20	1.00	0.00

OSA performance

OSA was calculated for the first and second attempts to identify passcodes. The researchers calculated what percent of the passcode was identified for each attempt. For example, if one out of three images was identified for UYI, the OSA performance would be .33. A 2 (OSAs: one viewing and three viewings) x 4 (authentication scheme: CHC, UYI, WYSWYE, and gaze-based) repeated measures ANOVA was conducted to explore OSA performance. The alphanumeric password was not tested for OSA resistance because the password was made clearly visible in our prototype. Main effects revealed differences among OSA performances for the schemes and differences among OSA performances for the number of viewings, $p < .001$ (see Figure 13). The main effects were qualified by a two-way interaction between OSAs and scheme, $F(3, 51) = 10.38$, $p < .001$, partial $\eta 2 = .379$. UYI became vulnerable after three viewings while the other schemes did not become more vulnerable with additional viewings. Because of our research questions, we explore simple effects. Post hoc comparisons revealed differences between partial passcodes identified for the schemes after viewing a log in one and three times, $p < .001$, such that viewing videos additional times improved attack performance. Differences were found

between attack performances on CHC (one observation: M = .06, SD = .13; three observations: M = .09, SD = .15) and UYI (one observation: M = .33, SD = .23; three observations: M =.80, SD = .20), CHC and WYSWYE (one observation: M = .32, SD = .22; three observations: M = .36, SD = .21), UYI and WYSWYE, UYI and the gaze-based scheme (one observation: M = .11, SD = .32; three observations: M = .17, SD = .38), and WYSWYE and the gaze-based scheme, $p < .05$ for all comparisons. No difference was found between attack performances for CHC and the gaze-based scheme, $p = .001$. All the schemes offered resistance to OSA. UYI was most vulnerable to attack followed by WYSWYE.

None of the participants were able to identify a full passcode after viewing a login one time. However, participants identified partial passcodes; this was more likely for UYI and WYSWYE. Seventeen participants could not identify any correct icons for CHC on the first viewing, and three participants identified one correct icon. Three participants could not identify any correct images for WYSWYE, 10 participants identified one image, five identified two, and two identified three. Four participants identified none for UYI, 12 identified one image, and four identified two. Two participants identified the gaze-based pattern by observing where the mouse entered and left the interface, which was a problem with our implementation rather than the scheme.

Given three viewings, no participant identified all of the targets for CHC or WYSWYE. Fifteen participants identified no correct icons for CHC, and five identified one. Two participants identified no correct images for WYSWYE, nine identified one, seven identified two, and two identified three. UYI was the most vulnerable after three viewings. Nine out of 20 participants identified the full passcode for UYI. One participant identified one image, and 10 identified two.

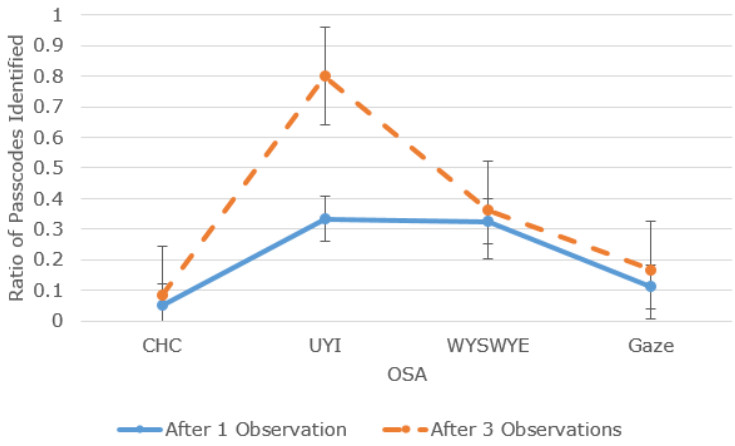

Figure 13. OSA performance. Error bars represent 95% confidence intervals.

Acceptability

Participants were asked whether they would accept the added effort the OSA prevention requires for each scheme. Eighty percent of participants accepted CHC, 45% accepted UYI, 50% accepted WYSWYE, and 68.75% accepted the gaze-based scheme.

Study 2: Memorability Runoff

Study 1 compared four OSA-resistant graphical schemes and the alphanumeric scheme on dimensions of usability and security. Study 2 adds the dimension of memorability by testing error rates and verbal memory for passcodes following a three-week delay.

Method

The following sections describe the methods used for Study 2; it was run as a two-part study.

Participants

Twenty undergraduate students participated in Part 1 (females = 13, males = 7), and 18 returned for Part 2 (females = 11, males = 7). The participants were told this was a two-part study in which they would need to return three weeks later for Part 2. Two participants were unable to return for Part 2. They were recruited through the SONA system and compensated with class research credit. Two participants reported being left hand dominant. Ages ranged for 18 to 42 (M = 21.67, SD = 5.60). Reported computer use ranged from 2.5 to 18 hours a day (M = 8.69, SD = 4.17). All participants reported normal or corrected to normal vision.

Stimuli, apparatus, and measures

Stimuli consisted of the same five authentication scheme prototypes that were used in Study 1. Study 2 included the System Usability Scale (SUS; Brooke, 1996) as a measure of satisfaction. This scale consists of 10 items that participants rate on a 5-point Likert scale.

Procedure

Participants were run individually. During Part 1 of the study, participants were seated in front of a desktop computer, and they signed a consent form. The experimenter explained that they would be logging in to five authentication interfaces and that in three weeks they would be doing the same thing. They were told that when they come in for Part 2, they would use the same passcodes as during Part 1, but they would not be reminded what the passcodes are. The order in which they logged into each interface was counterbalanced using a Latin Square design. For each interface, participants were shown their passcode and asked to memorize it. Participants were given instructions and one practice trial. Throughout the trials, the experimenter would answer any questions asked but would only volunteer feedback about whether participants had correctly authenticated. Participants logged in using CHC, WYSWYE, and alphanumeric passcodes 10 times. They logged in using UYI three times, with each login consisting of three identifications of targets. They logged in using the gaze-based scheme five times. After using each scheme, they completed the SUS.

Three weeks later, participants returned for Part 2 of the study and followed the same procedure without being shown their passcodes; however, they did not complete the SUS during this phase of the study.

Results

Error rates, the percent of each passcode remembered, and satisfaction scores were analyzed. Post hoc power analyses were conducted in GPower© for two-tailed repeated measures ANOVAs. For memorability and satisfaction, effect sizes were entered (.8 for large, .5 for medium, and .2 for small; Cohen, 1992), alpha (.05), and sample size (18). The likelihood of detecting a large and a medium effect for memorability exceeded 99%. The likelihood of detecting a large effect for satisfaction was 97%, and the likelihood of detecting a medium effect was 68%. The likelihood of detecting a small effect size was 46% for memorability and 16% for satisfaction. There was adequate power for a large or medium effect for memorability and adequate power for a large effect for satisfaction.

Memorability

Error rates were calculated for Time 1 and Time 2 as the number of incorrect trials of the experimental trials. The researchers calculated error rates at Time 1 and Time 2 for CHC, UYI, WYSWYE, and the alphanumeric scheme. The gaze-based scheme was not included in this analysis because of missing data. A 2 (elapsed time: day one or three weeks later) x 4 (authentication scheme: CHC, UYI, WYSWYE, and alphanumeric) ANOVA was conducted to explore memorability as measured by error rates (see Figure 14). Sphericity was violated, so a Greenhouse-Geisser correction was employed. Main effects revealed differences among error rates for the schemes and differences among error rates for elapsed time, $p < .001$. The main effects were qualified by a two-way interaction between elapsed time and scheme, $F(1.99, 27.88) = 35.97$, $p < .001$, partial $\eta^2 = .072$. Error rates for graphical passcodes did not differ by elapsed time. However, error rates for the alphanumeric passcode did (see Figure 10). Simple effects were investigated. A Bonferroni correction was used for post hoc comparisons. Post hoc comparisons revealed that Time 2 had more errors than Time 1, $p < .001$. The alphanumeric

scheme (Time 1: M = .05, SD = .20; Time 2: M = 1.00, SD = .00) had more errors than CHC (Time 1: M = .18, SD = .12; Time 2: M = .29, SD = .21), UYI (Time 1: M = .23, SD = .21; Time 2: M = .24, SD = .26), and WYSWYE (Time 1: M = .30, SD = .19; Time 2: M = .25, SD = .25), $p < .05$ for all comparisons. There were no differences found between CHC and UYI, $p = 1.00$, or WYSWYE, p = 1.00. There were no differences found between WYSWYE and UYI, $p = 1.00$.

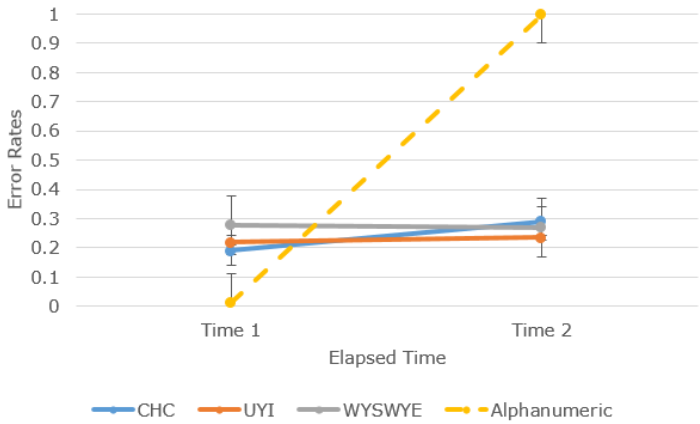

Figure 14. Memorability as reflected by error rates when logging in. Error bars represent 95% confidence intervals.

For CHC, participants had three targets to remember after a three-week delay. When asked to vocalize their targets, 7% of participants remembered just two icons and 93% remembered all three. UYI also consisted of three targets: 14% remembered one, 29% remembered two, and 40% remembered all three. WYSWYE consisted of four targets: 37% remembered none, 6% of participants remembered two, and 57% remembered all four. For the gaze-based passcode, 82% of participants remembered the passcode. No participants remembered the alphanumeric password after a three-week delay. All other passcodes had fewer errors during verbal report than entering them using the interfaces (see Figure 15 and 16).

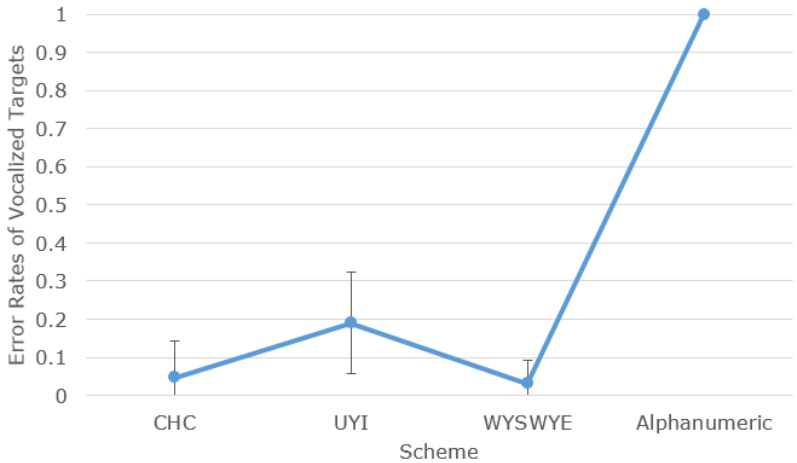

Figure 15. Errors rates stemming from interface interactions. Error bars represent 95% confidence intervals.

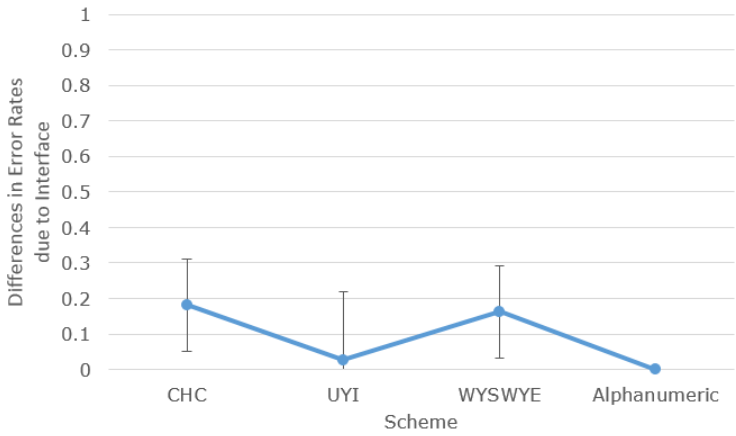

Figure 16. Increase in errors due to interacting with the interface. Error bars represent 95% confidence intervals.

Satisfaction

Satisfaction was measured by the SUS (Brooke, 1996) after the use of each scheme on Time 1 and Time 2. We did not measure SUS for the alphanumeric scheme at Time 2 because no passcodes were successfully entered. Due to technical problems with the eye-tracker and missing data, the gaze-based scheme was also not included in this analysis. A 2 (elapsed time: day one and three weeks) x 3 (authentication scheme; CHC, UYI, WYSWYE) ANOVA was conducted to explore satisfaction (see Figure 17). Sphericity was violated, and we used a Greenhouse-Geisser correction. There was no interaction between elapsed time and scheme, $F(1.25, 18.80) = 0.09$, $p = .820$, partial $\eta 2 = .006$, which indicated that satisfaction at time two did not depend on the scheme. Main effects revealed differences in satisfaction among schemes, $F(1.85, 18.80) = 10.88$, $p < .001$, partial $\eta 2 = .420$. There was no main effect for elapsed time, $p = .217$. Participants were most satisfied with CHC (Time 1: $M = 73.25$, $SD = 22.58$; Time 2: $M = 81.88$, $SD = 11.35$), followed by UYI (Time 1: $M = 64.53$, $SD = 16.34$; Time 2: $M = 68.59$, $SD = 18.37$), and then WYSWYE (Time 1: $M = 54.84$, $SD = 25.07$; Time 2: $M = 61.78$, $SD = 20.35$).

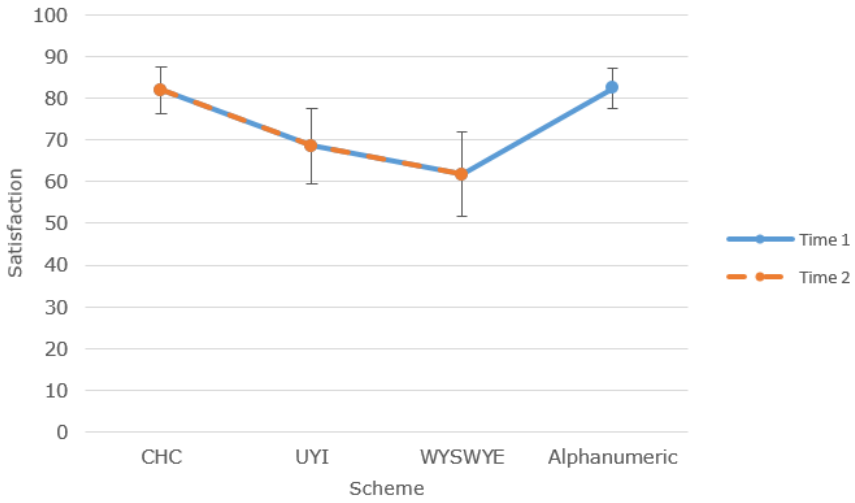

Figure 17. Satisfaction. Error bars represent 95% confidence intervals.

Discussion

The current research provided a direct comparison of prototypical examples of graphical passcodes and a gaze-based scheme that were designed to thwart OSAs. These schemes were classified as providing resistance by grouping targets among distractors, translating targets to another location, disguising targets, and using gaze-input, and they were compared to the traditional alphanumeric scheme. We determined the relative strengths of the authentication schemes regarding memorability, quick access, learnability and successful entry, security, and satisfaction. As a caveat, our sample consisted of college students who were not motivated by a reward. Many previous studies have not provided rewards when assessing graphical passwords (Gao et al., 2009; Jenkins et al., 2014; Kim et al., 2010; Lin et al., 2007; Liu, Gao, et al., 2011; van Eekelen, van den Elst, & Khan, 2013; Zangooei et al., 2012), and motivation is especially not needed in a within-subject study, in which we were comparing and contrasting schemes. Motivation would impact each scheme similarly (Cain, Werner, & Still, 2017). This study provides a high degree of internal validity but low external validity due to the college sample and laboratory controls (Still, 2011). Future studies could use a sample of actual users in the wild.

Evidence was provided for the memorability of all four prototypes of graphical and gaze-based passcodes. Participants had similar rates of error on day one and three weeks later for these schemes. For CHC, WYSWYE, and the gaze-based scheme, most participants verbally remembered the whole passcode. Forty percent remembered the whole passcode for UYI. The memorability of the graphical schemes was impressive when compared with the alphanumeric scheme, for which no participant could enter the passcode correctly or verbally remember it three weeks later. Remembering the alphanumeric passcode may have been challenging because it was long and complex to be secure and because it was system-assigned. However, the graphical passcodes and gazed-based scheme were also system-assigned. Impressively, the graphical passcodes were easily remembered, despite being system-assigned, likely due to the picture superiority effect. Muscle memory could have also aided the memorability of the gaze-based scheme. Memorability for UYI likely benefited from cued-recall. Being able to view the targets leverages cognitive abilities for memory (Al Ameen, 2016). Cued-recall would have aided memory for CHC and WYSWYE to a lesser extent than UYI because the targets were among many distractors. Findings that the grouping scheme and the scheme for translating to another location were memorable was consistent with previous literature (Brostoff et al., 2010; Wiedenbeck et al., 2006), and memorability for disguising targets and the gaze-based scheme was higher than in previous studies in which a drop in success rates were observed after three weeks (Hayashi et al., 2008) and after 10 days (De Luca et al., 2009).

Alphanumeric passcodes had fast login times, which was expected. This scheme was familiar to participants and could be entered using only keystrokes. Login times were also appropriate for CHC and the gaze-based scheme. These schemes meet the usability requirement of providing quick access. When users are focused on their primary task of interacting with data on a device, these schemes with appropriate login times will not preoccupy users with the secondary task of authenticating. CHC may allow for different numbers of challenges (e.g., participants may click to authenticate once or on multiple subsequent grids of icons). Login times were low for CHC because we had participants complete one challenge to authenticate. The previous assessment of CHC had shown much slower login times of 71.66 seconds for multiple challenges (Wiedenbeck et al., 2006). Fast login times for CHC were consistent with Behl and colleagues' (2014) five second login times for their grouping scheme and were faster than Sreelatha and colleagues (2011) login times of 29.95 seconds. The gaze-based scheme also had appropriate login times because the technology enforces selections at a certain pace. Appropriate login times for the gaze-based scheme were consistent with previous literature (De Luca et al., 2009). WYSWYE and UYI had longer login times. UYI required three image selections on subsequent grids to authenticate, and WHSWYE required some mental transformations. Findings of long login times for WYSWYE and UYI were consistent with previous research (Hayashi et al., 2008; Khot et al., 2012). Shorter login times have been found for other schemes that use the same general strategies (Cain & Still, 2016; De Luca et al., 2010; Zangooei et al., 2012).

There was only a slight, non-significant improvement in errors for CHC and WYSWYE during the course of Study 1. However, learning was demonstrated for UYI. Once participants figured out what the degraded versions looked like through trial and error and feedback from the

experimenter, participants improved for UYI. Learnability led to the low overall error rates for UYI. The quality of learnability that was present for UYI but not the other novel schemes is a shortcoming that needs to be addressed by changes in design rather than requiring training.

The graphical approaches were resistant to OSAs. No graphical passcode was stolen after one viewing. CHC, WHSWYE, and the gaze-based scheme continued to offer resistance after three viewings, after which no full passcode was stolen. However, UYI became vulnerable after three viewings. Just as participants could learn to identify degraded versions of targets and demonstrated learnability throughout the trials for UYI, the attackers were also able to learn the identity of degraded targets during additional viewings. UYI's vulnerability to OSAs likely applies to other graphical schemes that involve the direct selection of static images. Findings that schemes that translate to another location are resistant to OSAs were consistent with previous literature (Liu, Qiu, et al., 2011; Sun et al., 2016), and previous literature has shown inconsistencies in the security toward OSAs provided by disguising targets (Zakaria et al., 2011). The alphanumeric password was not assessed for OSA resistance, but the previous literature suggests it would be less secure to casual onlookers than graphical approaches. When participants attempted to steal an alphanumeric password and a graphical password, they identified an average of 3.65 characters for the alphanumeric and an average of .55 characters for the graphical (Tari, Ozok, & Holden, 2006).

There were high rates of acceptance for CHC and the gaze-based schemes, and there was high satisfaction for CHC. There was lower acceptance and satisfaction for WYSWYE and UYI. Lower rates of acceptance for WYSWYE aligned with previous literature (Khot et al., 2012). Participants may have found it difficult to transform the images in WYSWYE and may have been dissatisfied with high error rates.

Although the graphical schemes offered memorability and resistance to OSAs, error rates were high for these schemes compared to the familiar alphanumeric scheme. The high error rates found in this within-subject runoff were consistent with Behl and colleagues' (2014) finding of 20% error rates for their grouping scheme, but they were higher than some previous assessments of graphical schemes using this strategy (Wiedenbeck et al., 2006). Error rates were also higher for the scheme that translated targets to another location compared to previous studies that showed successful logins (Gupta et al., 2012; Khot et al., 2012). High error rates likely came from a lack of familiarity and the additional cognitive effort often required with graphical approaches. The alphanumeric scheme had almost no errors because this is the conventional way of authenticating. It did not have a learning curve compared to the novel approaches.

Tips for Practitioners

The following tips for practitioners include general guidance that comes from previous literature and our authentication runoff findings:

- To increase memorability, use images as passcodes instead of alphanumeric characters.
- Use system assigned passcodes to avoid user biases.
- Choose graphical schemas that use the strategy of grouping, such as CHC, to promote better success rates.
- To promote learnability, choose a graphical scheme that uses the strategy of disguising the passcode, such as UYI.
- For the best learnability, choose the familiar alphanumeric scheme.
- For faster login times, choose a graphical scheme that uses the strategy of disguising, such as UYI.
- For better security against OSAs, choose a graphical scheme that uses the strategy of grouping, such as CHC, or translating to another location, such as WYSWYE.

Future Directions

The next step is to make graphical schemes acceptable to a large population of users. The first major hurdle is making them more intuitive (Still, Still, & Grgic, 2015). Participants may have limited experience with these novel graphical authentication schemes. So, initial mistakes and

learning pains are expected. However, we must design-out as many errors as possible to facilitate successful user experiences. Following some informal reflection, we suggest some likely cognitive error sources for each scheme. Errors with CHC might be classified as skill-based lapses. Participants are familiar with the direct selection of icons, and this may automatically lead them to take this approach rather than selecting within the mentally projected shape. Errors using WYSWYE may have been rule-based mistakes. Participants probably lack a clear understanding of how the authentication scheme functions. Errors with UYI might be classified as encoding-based errors. Participants simply seemed to have difficulty identifying their target images.

We are currently exploring whether the graphical scheme advantage of memorability still holds when users have to remember multiple passcodes and longer passcodes. Guessability of passcodes decrease when passcodes are longer, but the memorability tradeoffs of the improved security need to be measured. It is not enough to show that one graphical passcode is memorable because users have to remember an average of 7.95 passcodes for different accounts (Grawemeyer & Johnson, 2011). And, this number is rapidly growing with the development of new digital services. Creating authentication schemes that are both usable and secure is a challenge. However, the literature (e.g., Still et al., 2017) shows taking a human-centered approach to designing schemes promises to both mitigate risk and facilitate better user experiences.

Acknowledgements

The data on OSA performance was presented at CHI 2017 within an extended abstract.

References

Al Ameen, M. N. (2016). The impact of cues and user interaction on the memorability of system-assigned random passwords (Doctoral dissertation). Retrieved from UTA Libraries.

Ankush, D. A., Dhanashre, W., & Husain, S. S. (2014). Authentication scheme for shoulder surfing using graphical and pair based scheme. *International Journal of Advance Research in Computer Science and Management Studies, 2*(10), 161–166.

Arianezhad, M., Stebila, D., & Mozaffari, B. (2013). Usability and security of gaze-based graphical grid passwords. In A. A. Adams, M. Brenner, and M. Smith (Eds.), *Financial Cryptography and Data Security. FC 2013. Lecture Notes in Computer Science (Vol. 7862;* pp. 17-33). Springer, Berlin, Heidelberg.

Barton, B. F., & Barton, M. S. (1984). User-friendly password methods for computer-mediated information systems. *Computers & Security, 3*(3), 186–195.

Behl, U., Bhat, D., Ubhaykar, N., Godbole, V., & Kulkarni, S. (2014). Multi-level scalable textual-graphical password authentication scheme for web based applications. *REV Journal on Electronics and Communications, 3*(3-4), 166–124. doi: http://dx.doi.org/10.21553/rev-jec.64

Bianchi, A., Oakley, I., & Kim, H. (2016). PassBYOP: Bring your own picture for securing graphical passwords. *IEEE Transactions on Human-Machine Systems, 46*(3), 380–389.

Biddle, R., Chiasson, S., & Van Oorschot, P. C. (2012). Graphical passwords: Learning from the first twelve years. *ACM Computing Surveys (CSUR), 44*(4), 19. doi: 10.1145/2333112.2333114

Brooke, J. (1996). SUS: A "quick and dirty" usability scale. In P. Jordan, B. Thomas, & B. Weerdmeester (Eds.), *Usability evaluation in industry* (pp. 189–194). London, UK: Taylor & Francis.

Brostoff, S., Inglesant, P., & Sasse, M. A. (2010). Evaluating the usability and security of a graphical one-time PIN system. In *Proceedings of the 24th BCS Interaction Specialist Group Conference* (pp. 88-97). Swindon, UK: BCS Learning & Development Ltd.

Bulling, A., Alt, F., & Schmidt, A. (2012). Increasing the security of gaze-based cued-recall graphical passwords using saliency masks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (3011–3020). doi: 10.1145/2207676.2208712

Cain, A. A., & Still, J. D. (2016). A rapid serial visual presentation method for graphical authentication. In D. Nicholson (Ed.), *Advances in Human Factors in Cybersecurity. Advances in Intelligent Systems and Computing: Vol 501* (pp. 3-11). Springer, Cham. doi: 10.1007/978-3-319-41932-9_1

Cain, A. A., Werner, S., & Still, J. D. (2017). Graphical authentication resistance to over-the-shoulder-attacks. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 2416–2422). New York, NY: ACM. doi: 10.1145/3027063.3053236

Cazier, J. A., & Medlin, B. D. (2006). Password security: An empirical investigation into e-commerce passwords and their crack times. *Information Systems Security, 15*(6), 45–55.

Chakrabarti, S., Landon, G. V., & Singhal, M. (2007). Graphical passwords: Drawing a secret with rotation as a new degree of freedom. In *Proceedings of the Fourth IASTED Asian Conference on Communication Systems and Networks* (pp. 114–120). Anaheim, CA, USA: ACTA Press.

Chen, Y. L., Ku, W. C., Yeh, Y. C., & Liao, D. M. (2013). A simple text-based shoulder surfing resistant graphical password scheme. In *2013 IEEE International Symposium on Next-Generation Electronics* (pp. 161–164). IEEE.

Choong, Y. Y., & Greene, K. K. (2016). What's a special character anyway? Effects of ambiguous terminology in password rules. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 60*(1), 760–764.

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science, 1*(3), 98–101.

De Luca, A., Denzel, M., & Hussmann, H. (2009). Look into my eyes! Can you guess my password? In *Proceedings of the 5th Symposium on Usable Privacy and Security* (Article No. 7). New York, NY: ACM. doi: 10.1145/1572532.1572542

De Luca, A., Hertzschuch, K., & Hussmann, H. (2010). ColorPIN: Securing PIN entry through indirect input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1103–1106). New York, NY: ACM doi: 10.1145/1753326.1753490

Dunphy, P., Fitch, A., & Olivier, P. (2008). Gaze-contingent passwords at the ATM. In *4th Conference on Communication by Gaze Interaction (COGAIN;* pp. 59–62).

Forget, A., Chiasson, S., & Biddle, R. (2010). Shoulder-surfing resistance with eye-gaze entry in cued-recall graphical passwords. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1107–1110). ACM.

Gao, H., Guo, X., Chen, X., Wang, L., & Liu, X. (2008). YAGP: Yet another graphical password strategy. In *Proceedings of the 2008 Annual Computer Security Applications Conference* (pp. 121–129). Washington DC, USA: IEEE Computer Society.

Gao, H., Liu, X., Dai, R., Wang, S., & Chang, X. (2009). Analysis and evaluation of the ColorLogin graphical password scheme. In *Fifth International Conference on Image and Graphics* (pp. 722–727). IEEE.

Ghori, F., & Abbasi, K. (2013). Secure user authentication using graphical passwords. *Journal of Independent Studies and Research, 11*(2), 34.

Grawemeyer, B. & Johnson, H. (2011). Using and managing multiple passwords: A week to a view. *Interacting with Computers, 23*(3), 256–267.

Gupta, S., Sahni, S., Sabbu, P., Varma, S., & Gangashetty, S. V. (2012). Passblot: A highly scalable graphical one time password system. *International Journal of Network Security & Its Applications, 2*(4), 201–216. doi: 10.5121/ijnsa.2012.4215

Hayashi, E., Dhamija, R., Christin, N., & Perrig, A. (2008). Use your illusion: Secure authentication usable anywhere. In *Proceedings of the 4th symposium on Usable privacy and security* (pp. 35–45). New York, NY: ACM.

Hoanca, B., & Mock, K. (2006). Secure graphical password system for high traffic public areas. In *Proceedings of the 2006 symposium on Eye tracking research & applications* (pp. 35–35). New York, NY: ACM.

Hui, L. T., Bashier, H. K., Hoe, L. S., Kwee, W. K., & Sayeed, M. S. (2014). A hybrid graphical password scheme for high-end system. *Australian Journal of Basic and Applied Sciences, 8*(2), 23–29.

IBM (2016). 2016 cost of data breach study: Global analysis. Ponemon Institute LLC.

Jenkins, R., McLachlan, J. L., & Renaud, K. (2014). Facelock: Familiarity-based graphical authentication. *PeerJ, 2, e444*.

Johnson, K., & Werner, S. (2008). Graphical user authentication: A comparative evaluation of composite scene authentication vs. three competing graphical passcode systems. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 52*(6) 542–546.

Joshuva, M., Rani, T. S., & John, M. S. (2011). Implementing CHC to counter shoulder surfing attack in PassPoint–style graphical passwords. *International Journal of Advanced Networking and Applications, 2*(6), 906–910.

Kawagoe, K., Sakaguchi, S., Sakon, Y., & Huang, H. H. (2012). Tag association based graphical password using image feature matching. In *International Conference on Database Systems for Advanced Applications* (pp. 282–286). Springer, Berlin, Heidelberg.

Kaye, J. J. (2011). Self-reported password sharing strategies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2619–2622). New York, NY: ACM. doi: 10.1145/1978942.1979324

Khot, R. A., Kumaraguru, P., & Srinathan, K. (2012). WYSWYE: Shoulder surfing defense for recognition based graphical passwords. In *Proceedings of the 24th Australian Computer-Human Interaction Conference*, 285-294. doi: 10.1145/2414536.2414584

Kim, D., Dunphy, P., Briggs, P., Hook, J., Nicholson, J. W., Nicholson, J., & Olivier, P. (2010). Multi-touch authentication on tabletops. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1093–1102). New York, NY: ACM.

Kiran, T. S. R., Rao, K. S., & Rao, M. K. (2012). A novel graphical password scheme resistant to peeping attack. *International Journal of Computer Science and Information Technologies, 3*(5), 5051–5054.

Kumar, M., Garfinkel, T., Boneh, D., & Winograd, T. (2007). Reducing shoulder-surfing by using gaze-based password entry. In *Proceedings of the 3rd symposium on Usable privacy and security* (pp. 13–19). New York, NY: ACM.

Lashkari, A. H., Manaf, A. A., & Masrom, M. (2011). A secure recognition based graphical password by watermarking. In *Proceedings of the 11th International Conference on Computer and Information Technology* (pp. 164–170). Washington DC, USA: IEEE Computer Society.

Lin, D., Dunphy, P., Olivier, P., & Yan, J. (2007). Graphical passwords & qualitative spatial relations. In *Proceedings of the 3rd symposium on Usable privacy and security* (pp. 161–162). New York, NY: ACM.

Liu, X. Y., Gao, H. C., Wang, L. M., & Chang, X. L. (2011). An enhanced drawing reproduction graphical password strategy. *Journal of Computer Science and Technology, 26*(6), 988–999. doi: 10.1007/s11390-011-1195-7

Liu, X., Qiu, J., Ma, L., Gao, H., & Ren, Z. (2011). A novel cued-recall graphical password scheme. In *Proceedings of the 2011 Sixth International Conference on Image and Graphics*, (pp. 949–956). IEEE. doi: 10.1109/ICIG.2011.16

Manjunath, G., Satheesh, K., Saranyadevi, C., & Nithya, M. (2014). Text-based shoulder surfing resistant graphical password scheme. *International Journal of Computer Science & Information Technologies, 5*(2).

Meng, Y., & Li, W. (2013). Enhancing click-draw based graphical passwords using multi-touch on mobile phones. In L. J. Janczewski, H.B. Wolfe, S. Shenoi (Eds.), *Proceedings of the IFIP International Information Security Conference* (pp. 55–68). Springer, Berlin, Heidelberg.

Nicholson, J. (2009). *Design of a multi-touch shoulder surfing resilient graphical password*. (Doctoral dissertation). Newcastle University: Newcastle, UK.

Paans, R., & Herschberg, I. S. (1987). Computer security: The long road ahead. *Computers & Security, 6*(5), 403–416.

Paivio, A. (2013). *Imagery and verbal processes*. London, Ontario: Psychology Press.

Perkovic, T., Cagalj, M., & Rakic, N. (2009). SSSL: Shoulder surfing safe login. In *Proceedings of the 17th International Conference on Software, Telecommunications & Computer Networks* (pp. 270–275). IEEE.

Rajavat, R., Gala, B., & Redekar, A. (2015). Textual and graphical password authentication scheme resistant to shoulder surfing. *International Journal of Computer Applications, 114*(19), 26–30.

Rao, K., & Yalamanchili, S. (2012). Novel shoulder-surfing resistant authentication schemes using text-graphical passwords. *International Journal of Information and Network Security, 1*(3), 163.

Rokade, A. H., Hasan, Z. U., & Mahajan, S. A. (2014). User authentication by secured graphical password implementation. In *Proceedings of the International Journal of Innovative Research in Science & Engineering* (pp. 1–8). IEEE.

Sasamoto, H., Christin, N., & Hayashi, E. (2008). Undercover: Authentication usable in front of prying eyes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 183–192). New York, NY: ACM. doi: 10.1145/1357054.1357085

Schaub, F., Walch, M., Könings, B., & Weber, M. (2013). Exploring the design space of graphical passwords on smartphones. In *Proceedings of the Ninth Symposium on Usable Privacy and security* (pp. 11–31). New York, NY: ACM. doi: 10.1145/2501604.2501615

Sreelatha, M., Shashi, M., Anirudh, M., Ahamer, M. S., & Kumar, V. M. (2011). Authentication schemes for session passwords using color and images. *International Journal of Network Security & Its Applications, 3*(3), 111–119. doi: 10.5121/ijnsa.2011.3308

Still, J. D. (2011). Experimental design: Does external validity trump internal validity? *ACM Interactions, 18*(3), 66–68.

Still, J. D., Cain, A. A., & Schuster, D. (2017). Human-centered authentication guidelines. *Journal of Information and Computer Security, 25*, 437–456. doi: 10.1108/ICS-04-2016-0034

Still, J. D., Still, M. L., & Grgic, J. (2015). Designing intuitive interactions: Exploring performance and reflection measures. *Interacting with Computers, 27*, 271–286.

Sun, H. M., Chen, S. T., Yeh, J. H., & Cheng, C. Y. (2016). A shoulder surfing resistant graphical authentication system. *IEEE Transactions on Dependable and Secure Computing, 15*(2), 180–193. doi:10.1109/TDSC.2016.2539942

Tao, H. (2006). *Pass-Go, a new graphical password scheme* (Doctoral dissertation). University of Ottawa: Canada. Retrieved from https://ruor.uottawa.ca/bitstream/10393/27297/1/MR18470.PDF

Tari, F., Ozok, A., & Holden, S. H. (2006). A comparison of perceived and real shoulder-surfing risks between alphanumeric and graphical passwords. In *Proceedings of the second symposium on Usable privacy and security* (pp. 56–66). New York, NY: ACM.

Vachaspati, P. S. V., Chakravarthy, A. S. N., & Avadhani, P. S. (2013). A novel soft computing authentication scheme for textual and graphical passwords. *International Journal of Computer Applications, 71*(10), 42–54.

Van Oorschot, P. C., & Wan, T. (2009). TwoStep: An authentication method combining text and graphical passwords. In *Proceedings of the International Conference on E-Technologies* (pp. 233–239). Springer, Berlin, Heidleberg. doi: 10.1007/978-3-642-01187-0_19

van Eekelen, W. A., van den Elst, J., & Khan, V. J. (2013). Picassopass: A password scheme using a dynamically layered combination of graphical elements. In *Proceedings of the CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 1857–1862) New York, NY: ACM.

Wash, R., Rader, E., Berman, R., & Wellmer, Z. (2016). Understanding password choices: How frequently entered passwords are re-used across websites. In *Proceeding of the 12th Symposium on Usable Privacy and Security* (pp. 175–188).

Walters, R. (2014). *Cyber attacks on U.S. companies since November 2014*. The Heritage Foundation. Retrieved from https://www.heritage.org/cybersecurity/report/cyber-attacks-us-companies-november-2014

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), 102–127.

Wiedenbeck, S., Waters, J., Sobrado, L., & Birget, J. C. (2006). Design and evaluation of a shoulder-surfing resistant graphical password scheme. In *Proceedings of the working conference on Advanced visual interfaces* (pp. 177–184). New York, NY: ACM. doi: 10.1145/1133265.1133303

Yakovlev, V. A., & Arkhipov, V. V. (2015). User authentication based on the chess graphical password scheme resistant to shoulder surfing. *Automatic Control and Computer Sciences, 49*(8), 803–812.

Zakaria, N. H., Griffiths, D., Brostoff, S., & Yan, J. (2011). Shoulder surfing defense for recall-based graphical passwords. In *Proceedings of the Seventh Symposium on Usable Privacy and Security* (Article No. 6). New York, NY: ACM. doi: 10.1145/2078827.2078835

Zangooei, T., Mansoori, M., & Welch, I. (2012). A hybrid recognition and recall based approach in graphical passwords. In *Proceedings of the 24th Australian Computer-Human Interaction Conference* (pp. 665–673). New York, NY: ACM. doi: 10.1145/2414536.2414637

Zhao, H., & Li, X. (2007). S3PAS: A scalable shoulder-surfing resistant textual-graphical password authentication scheme. In *Proceedings of the 21st International Conference on Advanced Information Networking and Applications Workshops* (Vol. 2, pp. 467–472). IEEE.

Zviran, M., & Haga, W. J. (1999). Password security: An empirical study. *Journal of Management Information Systems,15*(4), 161–185.

About the Authors

Ashley Cain
Ms. Cain is a PhD student in human factors psychology at Old Dominion University where she focuses on the human side of cyber security. She received her master's degree from San Jose State University in experimental and research psychology.

Jeremiah Still
Dr. Still is an Assistant Professor at Old Dominion University. His Psychology of Design laboratory explores the relationship between human cognition and technology; specifically, he is focusing on usable cybersecurity, visual attention, and intuitive design. He earned a PhD in Human-Computer Interaction from Iowa State University.