# A Comparison of Perceived and Real Shoulder-surfing Risks between Alphanumeric and Graphical Passwords

Furkan Tari
Dept. of Information Systems, UMBC
1000 Hilltop Circle
Baltimore, MD 21250
+1 (410) 455 3984
furkantari@umbc.edu

A. Ant Ozok
Assistant Professor of Dept. of Information Systems, UMBC
1000 Hilltop Circle
Baltimore, MD 21250
+1 (410) 455 8627
ozok@umbc.edu

Stephen H. Holden
Assistant Professor of Dept. of Information Systems, UMBC
1000 Hilltop Circle
Baltimore, MD 21250
+1 (410) 455 3936
holden@umbc.edu

## ABSTRACT
Previous research has found graphical passwords to be more memorable than non-dictionary or "strong" alphanumeric passwords. Participants in a prior study expressed concerns that this increase in memorability could also lead to an increased susceptibility of graphical passwords to shoulder-surfing. This appears to be yet another example of the classic trade-off between usability and security for authentication systems. This paper explores whether graphical passwords' increased memorability necessarily leads to risks of shoulder-surfing. To date, there are no studies examining the vulnerability of graphical versus alphanumeric passwords to shoulder-surfing.

This paper examines the real and perceived vulnerability to shoulder-surfing of two configurations of a graphical password, Passfaces™[30], compared to non-dictionary and dictionary passwords. A laboratory experiment with 20 participants asked them to try to shoulder surf the two configurations of Passfaces™ (mouse versus keyboard data entry) and strong and weak passwords. Data gathered included the vulnerability of the four authentication system configurations to shoulder-surfing and study participants' perceptions concerning the same vulnerability. An analysis of these data compared the relative vulnerability of each of the four configurations to shoulder-surfing and also compared study participants' real and perceived success in shoulder-surfing each of the configurations. Further analysis examined the relationship between study participants' real and perceived success in shoulder-surfing and determined whether there were significant differences in the vulnerability of the four authentication configurations to shoulder-surfing.

Findings indicate that configuring data entry for Passfaces™ through a keyboard is the most effective deterrent to shoulder-surfing in a laboratory setting and the participants' perceptions were consistent with that result. While study participants believed that Passfaces™ with mouse data entry would be most vulnerable to shoulder-surfing attacks, the empirical results found that strong passwords were actually more vulnerable.



## Categories and Subject Descriptors
H.5.2 [**Interfaces and Representation**]: User Interfaces – *Graphical user interfaces;* K.6.5 **[Computing Milieu]**: Security and Protection – *Authentication.*

## General Terms
Experimentation, Security, Human Factors, Design

## Keywords
Authentication, Human Factors, Social Engineering, Shoulder Surfing, Graphical Passwords, Authentication, Password Security, Usable Security.

## 1. INTRODUCTION
Authenticating users in network- and Internet-based environments has been a challenge for network administrators and end users. Attackers, on the other hand, recognizing certain facts about humans, may easily gain access to these organizational or personal information resources. Despite their vulnerabilities, passwords are still the most commonly used authentication mechanism. Although organizations may adopt "strong" password policies [11] that encourage or require users to select passwords less susceptible to discovery, such policies typically increase the burden on the users' ability to remember those passwords [46]. Users tend to type such passwords (considered a non-dictionary word) about 40 percent slower than dictionary words [37], making the data entry process for such authentication more vulnerable to shoulder-surfing attacks (defined below). Alternative authentication solutions, such as token-based or biometric authentication, do not rely on the users' memory and introduce an increased level of security at the expense of increased hardware and software costs and usability, and are therefore not used as frequent means of user authentication [2, 13, 42].

The issue of how to design authentication systems that are both secure and usable is yet another example of what continues to be a challenge to the human computer interaction (HCI) and security communities [8]. Is it possible to have both security and usability, or do both users and organizations have to accept what has typically been considered a trade-off between these two attributes [28]? Historically, you could have one, but not the other. This paper further explores the possibility that graphical passwords, in this particular study Passfaces™, may offer both a secure and usable solution to network- and Internet-based user authentication.

The focus of this paper emerged from a study comparing user attitudes and preferences between alphanumeric passwords and Passfaces™, a commercially available graphical password system. That study reported on a variety of issues, including user preference for passfaces (the term "passface" refers to the authentication element that replaces alphanumeric passwords in the Passfaces™ authentication system) versus passwords on factors such as satisfaction, usability, memorability, trust and security. The study results indicated that users generally preferred Passfaces™ to passwords in a variety of issues including usability and reliability [27]. Study participants raised a concern, though, about the possibility that Passfaces™ might be more vulnerable to shoulder-surfing than alphanumeric passwords, which therefore may affect users' trust level in the graphical password system.

These user concerns led to the follow-on study presented in this paper that compares two kinds of alphanumeric passwords and two most common and viable configurations of Passfaces™ for vulnerability to shoulder-surfing. This study examines the perceived and real trade-off between usability and security for these two forms of passwords when at risk to shoulder-surfing. For purposes of this study, the term *shoulder-surfing* is defined as:

*"…the observation of an individual entering their password without their knowledge. Historically, this involved looking over the individual's shoulder while they were sitting at a terminal"* [4].

The paper is organized as follows. The literature review examines previous research on the usability of authentication systems with a particular emphasis on alphanumeric and graphical passwords. It also summarizes the literature of password security and offers a more detailed review of password vulnerability to shoulder-surfing attacks. The research methodology section identifies the research questions for this paper, background information on study participants and the experimental procedure for gathering the data. Data gathering focused on users' real and perceived ability to shoulder surf alphanumeric passwords and Passfaces™ in a laboratory setting. The results section presents an analysis of the success rates of study participants' shoulder-surfing the four configurations (two configurations of Passfaces™ and two types of passwords) and then assesses participant perceptions of the vulnerability of each configuration to shoulder-surfing attacks. The conclusions and discussions section examines the real and perceived security and usability trade-offs for these two kinds of password systems.

# 2. LITERATURE REVIEW

## 2.1 Usability and Authentication

Recent studies have acknowledged that secure systems in general and authentication solutions in particular can benefit from improvements in usability. Unfortunately, most studies on security and usability seem to confirm the widely-held belief that systems can be either secure or usable, but not both. More recently, though, there is a concerted effort by usability and security researchers to work together with the aim of building systems that are both secure and usable [7]. In fact, there is some literature that suggests that considering usability earlier in the development of secure systems might help ensure the proper configuration and use of secure systems so that they achieve desired levels of assurance that otherwise might not be achieved because of user misfeasance [39].

There are two main streams of research into the usability and security of various authentication solutions. Computer security research tends to focus on the ability of attackers to "crack" password solutions for authentication with little emphasis on usability [16, 25, 29]. Usability research focuses on memorability of passwords with some emphasis on user satisfaction, but with little emphasis on security implications [46]. Another school of thought argues that poor authentication usability leads to poor security as users, as an example, write down passwords that they cannot memorize and recall. As a result, these researchers argue that it is imperative that developers design in both security and usability from the beginning of the system or product life cycle [1, 7, 23]. The following section outlines the current literature on security and usability of electronic authentication, organized by solution or technology along with a review of password attacks through shoulder-surfing.

## 2.2 Passwords / Personal Identification Numbers (PINS)

Arguably, nearly every participant of information systems uses the most prevalent form of individual authentication: passwords and PINs. These authentication solutions are used for a variety of security functions such as authorization, access control, and signatures. A challenge arises, though, for the many users who have to manage a large number of username/PIN/password combinations as they navigate all of the information systems, including e-commerce and e-government web sites they might use as part of their personal and professional life. A number of studies [1, 41, 46] have documented the problem that most users cannot remember a unique set of authenticators and identifiers for each of the systems they use. These authors typically cite basic human cognitive limitations from the psychology literature in explaining why this is so. For example, one issue is the amount of memory burden put on the users relating to the chunking principle by Miller [21]. This is especially true when organizations (typically employers) require employees to create "strong passwords" that are less susceptible to dictionary and brute force attacks.

A classic example of organizational instructions for creating strong passwords is the Department of Defense [11] guideline that suggests passwords be used only for one year, that they should be memorized and not written down, and that randomly assigned passwords are the most secure. It has become apparent, though, that users develop coping strategies over time for dealing with requirements for password creation like the Department of Defense's and the limits of human cognition. These strategies include writing down the passwords [1], reusing the same username/password combination across systems [16], or having to reset their passwords when they inevitably cannot remember them. These practices have real costs for organizations, including lost worker productivity and driving approximately 30 percent of helpdesk calls for password resets as found by Gartner [44, 45].

## 2.3 Graphical Passwords

The now well-documented weaknesses of username/PIN/password solutions for electronic authentication has led both researchers and practitioners to find and/or create hybrid solutions that might approach the familiarity (and to some extent usability) of usernames/PINs/passwords and the security of cryptographic solutions. One such alternative is sometimes called a "graphical password." Generally used in lieu of an alphanumeric password, graphical passwords rely on a user to select a predetermined image or set of images on a visual display (like a Web browser or PDA screen) by selecting those images in a particular order to authenticate the user [17]. Claims of enhanced usability from graphical passwords derives from humans' innate ability to recognize faces, which machines have been trying to emulate with mixed success for some time now [6, 38].

Early research into this idea found the potential for increased security [17] without much focus on usability. More recent research involves a design that allows user choice in graphical passwords, in particular, faces of individuals, but leads to weaknesses similar to self-selected passwords. Davis et al. [9] found that users pick faces for graphical passwords in such predictable patterns (i.e., based on ethnic background, attractiveness of members of opposite gender, and similarity to user) that user selection leads to a greater probability of cracking such graphical passwords. As a result, Passfaces™, a commercially available implementation of graphical passwords, does not allow for user selection of faces [30].

Brostoff and Sasse [5] conducted some of the first empirical research on Passfaces™ and found a significantly lower rate of password resets and higher levels of memorability compared to passwords in a comparative test spanning over five months. They also found that performance (i.e., time to complete the authentication process) was slower than for passwords, in part because users had to pass through a number of screens with faces and also because of the relatively out-of-date hardware and software platforms used for the experiments. However, this research did confirm the presumed increase in memorability of graphical passwords compared to alphanumeric passwords. Subsequent research has replicated this relatively slower performance time for graphical versus alphanumeric passwords, with mixed results on memorability and ease of use [27, 43].

## 2.4 Password Security (or Lack Thereof)

Computer security is more of a human-centered problem than a technology problem. People have been the easier target for collecting authentication information for attackers to gain unauthorized access to systems. As a result, some authors have labeled users as the 'weakest link' in a computer system [26]. Even though there are other authentication methods that arguably provide more security, the costs associated with these mechanisms make username-password combinations indispensable for most computer networks, e-commerce, and e-government applications [32].

Despite the popularity of username/password authentication solutions, the weakness of this approach to security is well documented. According to the taxonomy devised by Vasiu and Vasiu [40], password attacks can be grouped into three different categories: guessing, cracking, and harvesting. If the password can easily be guessed, then this is a clear indication of a weak password set by the user. In some cases the password is set to be the same as the username, full name or birth date of the victim. If the password can be found using special software or algorithms, then that password is cracked. Finally, if the attacker manipulates their victims physically and/or psychologically so as to retrieve their passwords, this is referred to as password harvesting.

A number of studies indicate the need for creating more 'secure' passwords against algorithmic attacks. These recommendations include, first and foremost, not selecting a password from a dictionary in any language. Furthermore, the length of the characters should be at least eight and include a combination of letters, numbers and special characters. As noted earlier, the classic example of a policy intended to "harden" passwords to make them less vulnerable to dictionary and brute force attacks is the DOD policy [11].

While these two approaches generate passwords that are hard to crack using any type of brute force attack or dictionary attack, they significantly increase the burden on the users to memorize the password than writing it down. To some extent, pass-phrase approach seems to alleviate the memorability burden [46]. Instead of selecting random characters, pass-phrase approach recommends users to use the first letters of a phrase in password generation. A phrase like "My uncle and aunt have 12 cousins" will generate a password something like "Mu&ah12c".

Hacking into computer networks can often times be less convenient than "social engineering" the people who have access to these computer networks [24]. In most instances password attacks take the form of harvesting as a form of social engineering [22]. Orgill et al. [26] define social engineering as:

"*…a technique used by hackers or other attackers to gain access to seemingly secure systems through obtaining the needed information (for example, a username and password) from a person rather than breaking into the system through electronic or algorithmic hacking techniques…*"

Based on the example given above, the password "Mu&ah12c" can be considered to be resilient against guessing and cracking but for social engineers it does not really matter how strong the password is against brute force or dictionary attacks. Social engineering uses human emotion and manipulation to trick the victim into giving out privileged information using various methods including but not limited to 'dumpster diving', 'persuasion', 'observation', etc [15, 26].

Strong password policies and procedures, along with user education, may help raise an awareness against these types of attacks. However, all these policies and procedures [11, 36] may be useless in preventing shoulder-surfing as this attack relies on the users' disclosing their password visually, thereby obviating the need to guess the passwords.

Unlike traditional alphanumeric passwords, graphical password authentications rely on images or pictures to replace the traditional alphanumeric "what you know" secrets [3] that users often have to write down to remember or re-use across multiple systems. Because users cannot share or write down such passwords, graphical passwords should arguably be less

susceptible to most of the "social engineering" attacks in general.

## 2.5 Password Attacks through Shoulder-surfing

The primary benefit of graphical passwords compared to alphanumeric passwords is the improved memorability. However, the potential detriment of this advantage is the increased risk of shoulder-surfing. Graphical passwords that use graphics or pictures [32] such as PassFaces [30] , Jiminy [31], VIP [10], Déjà Vu [12], Passpoints [43] or a combination of graphics and audio such as AVAP [18] are likely all subject to this increased risk unless somehow mitigated in implementation. This section addresses the literature on shoulder-surfing attacks on passwords in more detail.

A review of the information systems and computer science literature of security and information assurance uncovered only a few papers addressing the 'shoulder-surfing' phenomenon. Shoulder-surfing is considered a form of 'social engineering' that is gaining more and more importance as devices such as video camcorders and even cellular phones with audio-visual capabilities become more affordable to consumers. In one study, researchers attempted to evaluate users' perceptions of alternative authentication mechanisms including cognitive questions (where participants' responses to open-ended questions are used as passwords, such as mother's maiden name) and ImagePINs (selecting images from a number of icons) [14]. More participants believed that cognitive questions method is more resistant to being cracked compared to ImagePIN method (77% and 45% respectively). But these results only demonstrate the perceptions of participants and no actual data were collected to support this finding empirically. In another study, researchers proposed a method, which they referred to as "cognitive trapdoor games" that offers increased security for users entering their PIN numbers even if an attacker fully observes the entry [33]. However, the method used in said study is not based on authentication literature and the sample consisted of only eight users. The findings of that study may therefore be difficult to generalize to a larger population.

There is some research that addresses the issue of shoulder-surfing vulnerabilities of graphical passwords directly. In one study, researchers developed a graphical authentication system in which icons (or pictures) move on the screen at a desired speed as the user has to locate them and select them as they move. The purpose for having the icons or pictures move was to defeat attacks from malicious code that records the mouse click-stream of the victim by recording the X and Y axis of each click. On the other hand, the authors believe the system is still vulnerable to shoulder-surfing attacks through direct observation or video recording [35]. Man et al. [19] propose an alternative form of graphical passwords, where icons presented to the user have a number of variations (creating convex hulls and thus limiting the attacker's ability to identify the correct password. A follow up study by the authors is still in progress, where they plan to mathematically prove the resistance of this system to shoulder-surfing [20]. Neither of these studies relies on a commercially available graphical password nor do they compare the vulnerability of graphical passwords to alphanumeric passwords.

The research to date on the usability and security of electronic authentication solutions has either looked at security or usability, but not both facets of these systems. Therefore, the present study seeks to compare users' perception versus the actual security of graphical passwords against the threat of shoulder-surfing. Additionally, none of the usability studies of electronic authentication assessed user satisfaction or preferences of security between two forms of authentication, namely, alphanumeric and graphical passwords. Considering that effective security systems require user trust and competence, the lack of research into user attitudes on authentication alternatives represents a serious gap in the information systems literature. As a result, this study included the development of a research methodology to assess both the perceived and real vulnerability to shoulder-surfing of a graphical password solution, Passfaces™, compared to an alphanumeric authentication solution.

# 3. RESEARCH METHODOLOGY

## 3.1 Research Questions

The current research can be considered a continuation of the previous research efforts [27], where user performance, satisfaction and preferences between two authentication systems, namely, alphanumeric passwords and Passfaces™, were assessed. For each authentication method in the current study, the experiment used two configurations, resulting in a total of four different authentication configurations. For alphanumeric passwords, the configurations consisted of dictionary and non-dictionary passwords. For the graphical authentication system Passfaces™, the configurations consisted of data entry methods using the mouse versus the keyboard. After this classification, the objectives of this study can be explained as follows:

1) Determining the actual and user-perceived vulnerabilities of two kinds of passwords and two configurations of Passfaces™ to shoulder-surfing under optimal conditions for a shoulder-surfer,

2) Determining the significant differences among the levels of shoulder-surfing vulnerabilities (real and perceived) for each of the authentication configurations,

3) Determining the interrelationships (correlations) between the perceived and real vulnerability levels of these authentication mechanisms to shoulder-surfing, and

4) Determining the perceived vulnerability of the authentication configurations to shoulder-surfing, both by ordinary onlookers and professional, malicious hackers.

This research relies on an experiment to address the above research questions, focusing on real and perceived shoulder-surfing vulnerabilities between passwords and Passfaces™. The exact procedures for the experiment are presented in the Experimental Procedures section below.

## 3.2 Independent and Dependent Variables

Since the research questions sought to discover the real and perceived vulnerabilities to shoulder-surfing of the two authentication types, the independent variable is the authentication type (passwords vs. Passfaces™). The dependent

variables are vulnerability to shoulder-surfing (as measured by the number of correct character/face responses in the shoulder-surfing procedure), and three perceived vulnerability items for both authentication systems to shoulder-surfing. These three items are also discussed in detail in the Experimental Procedures section.

## 3.3 Participants

The study was conducted as a laboratory experiment in a within-subject design, meaning all participants received the same treatment by going through all of the experimental procedures the same way. Participants were Information Systems and Computer Science graduate students at University of Maryland, Baltimore County (UMBC). A total of twenty students participated in this study. While true shoulder-surfing would occur by people with intent to steal authentication information by looking over the victims' shoulders, recruiting a participant group that would represent the true population was practically impossible. However, a participant group consisting of graduate students with authentication system experience was deemed appropriately representative of "potential shoulder surfers," as they use password-based logins on a day-to-day basis and are familiar with authentication in general. The researchers believe that professional hackers do not commonly use shoulder-surfing as a method to steal identities and other malicious activities, and choosing a group of hackers as participants, besides the obvious difficulties in recruitment, would not result in the correct methodology.

The age mean for the participant group was 29.85 with a standard deviation of 5.66. Gender distribution among the participants was exactly half and half. All participants had a master's degree in a related field and indicated their level of expertise with computers as 'expert'.

## 3.4 Experimental Procedures

Data collection for the study occurred in a controlled laboratory environment where participants worked with the experimenter in an isolated room to avoid any distractions. Participants were first given a questionnaire that collected demographic information including age, sex, highest degree earned and level of proficiency with computers. The participants were then given a brief introduction to Passfaces™. After that introduction, participants received a short training session on how Passfaces™ authentication system is used. Because the participant group was very familiar with alphanumeric passwords, they received no training on this kind of authentication mechanism.

Following the training session, the participants were asked to play the role of a hacker using 'shoulder-surfing' method to gain access to passwords and passfaces of authentication system users. In this experiment, the experimenter played the role of "the victim" for each shoulder-surfing condition. In order to gain access to the authentication information through shoulder-surfing, participants were given "optimal" shoulder-surfing conditions in which they had the option to sit next to the person (in this case the experimenter), entering their information or to stand behind them. The participants were free to move from one side to the other depending on how they felt the most comfortable trying to obtain the "victim's" password. They were also given a notebook and pencil, allowing them to take notes while observing the experimenter entering the passwords or passfaces. After the experimenter entering the authentication information as the "victim" in each of the four authentication system configurations, the participants were asked to enter the same passwords or passfaces once, using their notes if they wanted, and data were collected on the accuracy of the learned passwords and passfaces. The experimenter trained himself to enter each password and passface in a constant, optimal speed in order to prevent any noise stemming from different entry speeds of different experimental sessions. Therefore it is believed that the experimenter's performance was representative of an average victim's.

There are five screens in Passfaces™ from each of which participants had to choose one correct face (from a 3 by 3 face grid), and therefore, the passwords containing five characters were chosen. The data recording was conducted by recording to a database with electronic entry forms specifically developed for the purpose of this study.

Participants attempted to shoulder surf four different authentication configurations with two each for passwords and passfaces, respectively. For passwords, participants first attempted to attack a weak or dictionary-based password (a familiar five-character word such as panic) and then a stronger, non-dictionary password (a combination of random alphanumeric characters such as f9dq0 that were not case-sensitive). Following attempts to attack these two types of passwords, participants attempted to attack two configurations of Passfaces™. The first configuration is the "out of the box" setup for Passfaces™, where a user clicks on the assigned face in the 3x3 grid with a mouse, making it possible for a potential attacker to see the assigned set of passfaces when the experimenter used Passfaces™. In the second configuration, the experimenter used the numeric keypad of the computer keyboard instead of the mouse to select the assigned passfaces.


**Figure 1. Passfaces™ Login Window Screenshot**

The keyboard numerical pad entry system was configured in such a way that each face selected on the 3X3 grid corresponded to one of the numbers 1 through 9 on the keypad. For example, if the first correct passface was located on the upper right corner, the corresponding keypad number relative to its position would be '9'. Keypad entry posed some challenges to the participants, as it requires them to follow both the screen and the numeric keypad at the same time to identify a passface. Remembering the numbers entered would not help them because each time a user would log on to the Passfaces™ system, the

location of the correct faces and the decoy faces on the 3X3 grid would change randomly.

Following each of the four shoulder-surfing sessions, the participants were given a post-experiment questionnaire and asked to rate these four authentication configurations on a 7-point Likert scale to assess:

1) the perceived ease of recording the victim's password/passface ("It was very easy to record the person's password/passface");

2) the perceived vulnerability of each authentication configuration to shoulder-surfing ("I think this authentication system is very vulnerable to shoulder-surfing attacks");

3) the perceived vulnerability of each authentication configuration to hacking ("I think it is easy for hackers to hack into systems utilizing this authentication method"); and

4) the strategy they used to obtain the password/passface (i.e., "by looking at the screen", "by looking at the keyboard", etc.).

The perceived vulnerability of each authentication configuration was measured by the first question of the questionnaire. The remaining two questions in the survey sought to determine participant perceptions about how easily they could shoulder surf and how easy it would be for professional hackers to shoulder surf each authentication configuration. On the 7-point Likert scale, a high score indicated a perceived high vulnerability of the authentication system to shoulder-surfing. No quantitative analysis was conducted on the fourth question concerning the particular strategy the participants used in the shoulder-surfing procedure because this particular question inquired about the shoulder-surfing strategy they used. Only the percentage values of each strategy as part of this question are discussed in the next section.

The resulting data collected from each participant were analyzed and are presented in the next section.

# 4. RESULTS

The previous study indicated a high level of concern against shoulder-surfing among the users for the Passfaces™ authentication system [27]. It was therefore that the researchers felt the need to explore the differences in vulnerability among authentication systems. Two types of password-based and two types of passface-based authentication systems were identified to simulate the real-world authentications. The vulnerability for the said systems to shoulder-surfing was measured in terms of performance and user opinion, i.e., how this shoulder-surfing vulnerability was perceived by the users. For the performance part, the success rates of participants playing the role of "hackers" were measured under the four following sets of conditions: Dictionary Passwords, Non-Dictionary Passwords, Passfaces Using the Mouse and Passfaces Using the Keyboard. The success rates corresponded to the number of correctly guessed password characters or passface images *in the correct order*. Each password character or passface was marked individually as correct if they were identical to the corresponding character in the correct password or passface. For the user opinion part, the responses to the three questions provided by all participants after completing the shoulder-surfing tasks with each authentication system were checked for differences across the four authentication configurations. An Analysis of Variance (ANOVA) was used to detect significant differences in real (as measured by the performance) and perceived vulnerability (as measured by the scores of the three survey questions) to shoulder-surfing among the four authentication configurations. In addition to the ANOVA, a Duncan's Multiple Range Test indicated which authentication configurations had significant differences in terms of shoulder-surfing performance and opinions. The means and standard deviations as well as the ANOVA results for each of the performance and participant perception values are presented in Table 1 and Table 2.

**Table 1. Real and Perceived Vulnerability Descriptive Statistics, ANOVA and Duncan's Multiple Range Test Results**

| **Number of Correct Characters Entered in Correct Order** | | | |
|---|---|---|---|
| **Authentication Type** | **Average** | **Std. Dev.** | **Duncan Group** |
| Non-Dictionary Password | 3.65 | 1.631 | A |
| PassFaces with Mouse | 3.1 | 1.119 | A |
| Dictionary Password | 1.3 | 0.923 | B |
| PassFaces with Keyboard | 0.55 | 0.510 | C |
| *ANOVA: F = 34.14, p-value < 0.001, items with the same letter are not significantly different in Duncan's Multiple Range Test.* | | | |
| **System being Vulnerable to Shoulder-surfing Attacks (1=Not Vulnerable at All, 7=Extremely Vulnerable)** | | | |
| **Authentication Type** | **Average** | **Std. Dev.** | **Duncan Group** |
| Passfaces with Mouse | 5.2 | 1.005 | A |
| Non-Dictionary Password | 5.05 | 0.945 | A |
| Dictionary Password | 4.85 | 1.309 | A |
| Passfaces with Keyboard | 2.3 | 1.129 | B |
| *ANOVA: F = 30.90, p-value < 0.001, items with the same letter are not significantly different in Duncan's Multiple Range Test.* | | | |

Table 1 presents results concerning real and perceived shoulder-surfing vulnerability (number of correct characters entered by the participants during the shoulder-surfing task and the responses concerning how vulnerable they thought the configuration was to shoulder-surfing attacks, respectively). Table 2 presents the results concerning the two remaining questions in the post-experiment questionnaire. The last columns in Tables 1 and 2 indicate the Duncan Grouping value. If two configurations in this column are marked with the same letter, this means these two configurations are not significantly different from each other in terms of the value being measured (shoulder-surfing performance or survey scores).

Table 1 indicates that for shoulder-surfing performance, there were significant differences among three of the four authentication configurations, with the exception of the difference between Passfaces with Mouse and Non-Dictionary Password authentication configurations being not significant.

Out of the five characters, on average, participants guessed only 0.55 correct with Passfaces Using Keyboard, and this finding coupled with the relatively high standard deviation of 0.510 leads to the conclusion that the participants simply failed in shoulder-surfing when the "victim" was using the keyboard data entry configuration with Passfaces™. It should be noted that unlike passwords, participants had no way of writing down any elements of passfaces and therefore relied solely on recalling them. Significantly higher than Passfaces with Keyboard, Dictionary Password participants were able to recall on average 1.3 characters out of five. With an average of 3.1, the Passfaces with Mouse trials had the second highest number of correct recalls, next to Non-Dictionary Password trials with an average of 3.65. The findings indicate that participants were highly successful in recalling non-dictionary passwords. They were also able to recall a good number of passfaces under the scenario when the "victim" used the mouse. The correct recalls for these two circumstances were not significantly different from each other. This leads to the conclusion that, to some extent, Non-Dictionary Passwords and Passfaces Using Mouse are prone to shoulder-surfing. Further implications of these findings are discussed in the Conclusions and Discussions section.

**Table 2. Participant Opinions on Recording of the Authentication Information Descriptive Statistics, ANOVA and Duncan's Multiple Range Test Results**

| **How Easy it is to Record the Passwords/Passfaces (1=Extremely Difficult, 7=Extremely Easy)** | | | |
|---|---|---|---|
| **Authentication Type** | **Average** | **Std. Dev.** | **Duncan Group** |
| Passfaces with Mouse | 5.2 | 1.196 | A |
| Non-Dictionary Password | 5.05 | 1.234 | A |
| Dictionary Password | 2.45 | 1.191 | B |
| Passfaces with Keyboard | 1.6 | 0.754 | C |
| *ANOVA: F = 53.86, p-value < 0.001, items with the same letter are not significantly different in Duncan's Multiple Range Test.* | | | |
| **How Easy it is for Hackers to Obtain Passwords/Passfaces (1=Extremely Difficult, 7=Extremely Easy)** | | | |
| **Authentication Type** | **Average** | **Std. Dev.** | **Duncan Group** |
| Dictionary Password | 6.7 | 0.47 | A |
| Non-Dictionary Password | 5.65 | 0.875 | B |
| Passfaces with Mouse | 3.95 | 1.468 | C |
| Passfaces with Keyboard | 2.75 | 1.118 | D |
| *ANOVA: F = 56.17, p-value < 0.001, items with the same letter are not significantly different in Duncan's Multiple Range Test.* | | | |

Next, the significant differences in participant attitudes concerning shoulder-surfing among the four types of authentication were investigated. Three questions were asked to the participants regarding shoulder-surfing activities: How vulnerable the system is to shoulder-surfing attacks, how easy it was to record the "victim's" authentication, and how easy it is for hackers to obtain passwords/passfaces with the current authentication. The questions were posed as statements (such as "It was easy to record the password), and a 1-7 Likert scale was used for these first three questions with a scale of 1 corresponding to the response "Extremely Difficult" and 7 corresponding to "Extremely Easy." The scores were then compared among the four types of authentication using an ANOVA model for the three questions. One final question asked where the participants looked (screen, keyboard, mouse or any combination of the three) to shoulder surf. Table 2 presents the participants' opinion that the Passfaces with Mouse, Non-Dictionary Password and Dictionary Password authentications were equally "vulnerable to shoulder-surfing attacks" according to the significant ANOVA model, with the Passfaces with numeric keypad authentication having the only score significantly lower than any of the others. The opinion of participants about the vulnerability of the systems is to some extent consistent with the actual shoulder-surfing performance. This relationship is further investigated in a correlation analysis below.

Furthermore, the ANOVA comparison model regarding participant opinions on the "easiness to record passwords" question was also statistically significant. Participants indicated it was fairly difficult to record the authentication mentally with Passfaces Using Keyboard. While a moderate score was received for Dictionary Passwords (which is significantly higher than Passfaces with Keyboard authentication), it can be concluded that shoulder-surfing was again perceived as quite difficult with Passfaces Using the Keyboard. The very high scores for non-dictionary passwords and passfaces using the mouse for this question indicate that participants felt they could copy other's authentication information if the "victim" is using a non-dictionary password or passfaces with a mouse (significantly more so than the other two authentications).

The ANOVA analysis regarding the differences in the scores for "how easy it is for hackers to steal their authentication," a question to determine their opinion concerning the vulnerability of the authentication system against professional attackers, indicated significant differences among all four authentication systems. Participants found it very easy to shoulder surf for Dictionary Passwords and Non-Dictionary Passwords. The perceived possibility of the authentication getting stolen was also high for Passfaces with Mouse, but relatively low with Passfaces Using Keyboard.

The finding indicates that participants have a lack of trust on all authentication types except for Passfaces Using Keyboard. This finding is likely a result of their failure in shoulder-surfing with this configuration.

The participant attitude findings indicate a general lack of trust and a perceived vulnerability of passfaces and passwords. In general, when keyboard is used, there is a relative perceived "safe heaven" for protection against shoulder-surfing among participants, but a high comfort level against shoulder-surfing is still lacking.

As a next step, the analysis explored the relationship between how vulnerable each configuration was in reality versus participant perception. This analysis sought to determine the correct and incorrect perceptions concerning the authentication configurations about their vulnerability to shoulder-surfing. For this purpose, a correlation analysis between the score of actual success in correctly recalling information for password characters and passfaces (real vulnerability), and the survey score on participant perception of the easiness of recording the authentication information (perceived vulnerability) was conducted. In other words, the correlation analysis was conducted between the performance results and survey scores in Table 1, and the results of this correlation analysis are presented in Table 3.

**Table 3. Correlations between Real and Perceived Vulnerabilities for each Authentication Configuration.**

| Authentication Configuration | Correlation between Real and Perceived Vulnerability (Numbers in bold are significant at 0.05 alpha level) |
|---|---|
| Dictionary Password | **0.302** |
| Non-dictionary Password | 0.114 |
| Passfaces with Mouse | **0.535** |
| Passfaces with Keyboard | 0.055 |

Table 3 shows that significant correlations were detected for Dictionary Passwords and Passfaces with Mouse, indicating that the participants' perceptions of the vulnerabilities of these types of passwords and passfaces relate to the actual vulnerability of the configuration. It should be noted that dictionary passwords are more common than non-dictionary passwords, [34] and the significant correlations indicate accuracy in the participants' perception capabilities concerning the vulnerabilities of each configuration. This finding therefore indicates a well-placed concern on the participants' parts concerning the respective configurations' vulnerability. It also validates the concern expressed in previous research that Passfaces™ with mouse data input is more vulnerable to shoulder-surfing.

Interestingly, the lack of a correlation between the participants' perceptions and the reality of the vulnerability of Non-dictionary Password and Passfaces with Keyboard configurations indicate that participants did not accurately appreciate the vulnerabilities of these configurations to shoulder-surfing. The lack of correlation in the Non-dictionary Passwords might be explained by the fact that while in general non-dictionary passwords are believed to be "more secure" because of their resistance to dictionary attacks, the participants incorrectly assumed this level of assurance applied to shoulder-surfing risks too. In the case of Passfaces with Keyboard, a possible "ceiling effect" resulting from a total failure in shoulder-surfing by the participants (a very little fraction of passfaces could be successfully copied by the participants) is the likely reason for the lack of correlation due to the very low average of success scores in the particular configuration. When a data set has such low scores, it becomes virtually impossible to analyze the data. In this case, most responses for success rates using Passfaces with Keyboard configuration were zero. As a result, the ceiling effect is present.

In response to the fourth survey question, participants indicated that to shoulder surf for Dictionary Passwords and Non-Dictionary Passwords, they looked at the keyboard (100%), to shoulder surf for Passfaces Using Mouse they looked at the screen (100%), and to shoulder surf for Passfaces Using Keyboard, 15% indicated that they looked at the screen and 85% indicated they looked at both the screen and the keyboard. The responses to this question only indicate simple, common-sense shoulder-surfing strategies identified by the participants. All participants in a certain configuration used the same strategy (either looking at the screen or looking at the keyboard), except for the Passfaces Using Keyboard configuration (where participants failed to shoulder surf). Therefore, a data analysis to determine relationships between shoulder-surfing strategy and shoulder-surfing success for each authentication configuration could not be conducted.

These findings indicate that one configuration of the graphical password system, Passfaces™ with Mouse, is vulnerable to shoulder-surfing as expected and correctly perceived by study participants. It is somewhat surprising to find the vulnerability of Non-Dictionary Passwords to shoulder-surfing, which was inconsistent with study participant perceptions. In this context, partial success means being able to capture some, but not all, of the characters of the passwords or pictures from passfaces. The possible implications of these findings are discussed in the next section.

## 5. CONCLUSIONS AND DISCUSSIONS

The seminal question still remains: Can we have both usable and secure authentication systems? In particular, are graphical passwords the leading candidates to address this long-standing challenge, or do the very characteristics that make graphical passwords more memorable and usable lead to increased security vulnerabilities like shoulder-surfing? As in many cases, the answer is "maybe."

This paper presents some answers to this question based on a laboratory experiment with 20 graduate students at UMBC, which asked them to explore the real and perceived vulnerability of four configurations of authentication systems (two with alphanumeric passwords and two configurations of Passfaces™, a commercial graphical password). An examination of this real and perceived vulnerability of four configurations of authentication systems yields some predictable and some surprising results.

As expected, study participants both perceived and experienced a higher level of vulnerability of Passfaces™ with mouse to shoulder-surfing. The very characteristics that allowed users to recognize faces with higher rates of memorability in previous studies led to an increased effectiveness in shoulder-surfing. Participants were able to memorize at least some of the faces during the data entry. The Passfaces™ configuration with a mouse was relatively vulnerable to shoulder-surfing attack in the "ideal conditions" for an attack in the lab with four participants able to guess 100% of the five faces. A correlation between real and perceived vulnerability of Passfaces™ with

mouse to shoulder-surfing was found to be significant. Study participants expected this configuration to be vulnerable to shoulder-surfing and they were right, seemingly confirming previous literature that usability and security did trade-off.

This is not to say that Passfaces™ and graphical passwords are inherently vulnerable to shoulder-surfing because of configuration options. Switching the configuration from mouse input to keyboard input decreased the vulnerability to shoulder-surfing significantly. Of the four configurations, Passfaces™ keyboard entry was the least vulnerable to shoulder-surfing by far. Possibly due to the speed entry with the keyboard and need to look in two places at once, Passfaces™ with keyboard were virtually invincible against shoulder-surfing. Curiously, study participants did not perceive the Passfaces™ with keyboard to be less vulnerable and there was no statistically significant correlation between real and perceived vulnerability for this configuration. For computers and computing devices that have a numeric keypad, this offers the possibility for a secure and usable authentication solution, but for devices with more complicated numerical entry capabilities, such as laptop computers that do not have a separate numeric keypad, it may be necessary to devise other means of data entry that does not disclose which face is the assigned passface on each screen.

Somewhat surprisingly, non-dictionary passwords proved to be the most vulnerable to shoulder-surfing. The difference in real vulnerability was significantly higher than both dictionary passwords and Passfaces™ with keyboard. Study participants successfully "stole" a large percentage of password characters in non-dictionary passwords. While the participants perceived non-dictionary passwords to be vulnerable to shoulder-surfing, the differences in perceptions were not different to a statistically significant degree other than for Passfaces™ with keyboard. While non-dictionary passwords may be difficult to remember or record, shoulder-surfing is easier with this type of authentication. The successful shoulder-surfing for non-dictionary passwords may result from the ability of shoulder surfers to record the characters one by one without paying attention to the meaning of the password if it were a dictionary word. The survey results indicate that users find non-dictionary passwords strong authenticators. What they apparently misperceive is that a password that is more resistant to a dictionary attack is likely more vulnerable to shoulder-surfing. This represents a novel finding that has not yet been documented in the literature of usability and security of authentication systems.

The study also indicates that dictionary passwords are less vulnerable to shoulder-surfing than non-dictionary passwords and Passfaces™ with mouse to a statistically significant degree. While generally understood to be more susceptible to dictionary attacks, possibly due to the entry speed, dictionary passwords hold up quite well against shoulder-surfing. Despite the participants having likely been warned about the lack of security in dictionary passwords, there was a significant correlation between the perceived and real vulnerability of dictionary passwords being relatively low.

The results indicate the fact that both alphanumeric and graphical password-based authentication mechanisms may have significant vulnerability to shoulder-surfing unless certain precautions are taken. Despite the common belief that non-dictionary passwords are the most secure type of password-based authentication, our results demonstrate that it is in fact the most vulnerable configuration to shoulder-surfing. This result is unexpected, but possibly explainable. Although we tried to keep the entry speed as constant as possible for each entry method, we did not control entry speed in our experiment. Future studies may focus on typing speed and possible training effects from long-term use of passwords (both dictionary and non-dictionary) to better establish the impact of long-term use of passwords on their shoulder-surfing vulnerability.

Similarly, when passfaces are entered using the mouse, this is almost no different than giving away your secret code to the shoulder surfer, allowing them to observe the entire authentication information. But when this authentication is performed using the numeric keypad, the results show that it is the most secure authentication configuration compared to other configurations in this experiment. We should also note that the participants of this study were not actual hackers, and therefore it would be reasonable to conclude that actual hackers having more experience and probably more trained eyes and various recording devices would perform better than our participants.

While we believe that our study accurately addressed real and perceived shoulder-surfing vulnerability issues concerning the four configurations of authentication systems, there are still some issues that require further exploration. Future studies may investigate shoulder-surfing methods used by real hackers (for example multiple cameras or other equipment) as well as investigation of circumstances for most popular shoulder-surfing environments (work, public access points, etc.). These issues remain to be explored. It should be noted that this study only explored authentication systems' vulnerabilities against real-time shoulder-surfing with a person physically peeking over the victim's shoulder without the aid of a recording device or technology. Hopefully such ideal shoulder-surfing conditions are not found in typical workplaces or home computing environments.

The non-dictionary passwords being highly vulnerable to shoulder-surfing attacks is a finding that calls for further investigation. The experimenter playing the "victim" in this study was due to the focus of this study mainly being the comparison of shoulder-surfing vulnerabilities between graphical and alphanumeric authentication systems. The surprising finding of non-dictionary passwords being highly vulnerable to shoulder-surfing motivates the researchers to understand the exact underlying factors behind this vulnerability, one of which may be typing speed, among others. Therefore, a future study can include experimental scenarios where actual participants would play the "victim" to further explore the mechanics behind this vulnerability, taking into consideration other user-related factors that may be resulting in this vulnerability.

This research makes several major contributions to the authentication and security communities. To our knowledge, this is one of the very few papers that addresses 'shoulder-surfing' in detail and maybe the only paper to assess the vulnerability of two authentication mechanisms to shoulder-surfing. A major finding from the study is that secure and usable authentication might be possible when considering shoulder-surfing risks, but that configuration for data entry (i.e., mouse

versus numeric keypad) is an important consideration for graphical passwords like Passfaces™. Even though there were significant differences in real vulnerability of the four configurations to shoulder-surfing, study participants perceived the vulnerability to be roughly the same except for Passfaces™ with keyboard being perceived as significantly less vulnerable than the other three options. The real and perceived vulnerabilities of these four configurations of authentication systems were correlated to a lesser degree than one might expect for experienced computer users, raising some interesting possibilities for future study.

Finally, these findings call into question the notion that non-dictionary passwords are universally “better” than dictionary passwords. The risk mitigation from password choice clearly depends on the nature of the attack. While this research could benefit from larger scale experiments with a more diverse set of study participants, it nonetheless raises some issues about user training for password choice as well as configuration choices for graphical password systems. This study may not answer the question of whether a usable and secure authentication solution will soon be universally available, but it offers some preliminary answers to questions concerning alphanumeric and graphical authentication systems’ vulnerability to shoulder-surfing.

## 6. REFERENCES


[1] A. Adams and M. A. Sasse, "Users are not the Enemy: Why Users Compromise Computer Security Mechanisms and how to Take Remedial Measures," *Communications of the ACM*, vol. 42, pp. 41-46, 1999.

[2] R. J. Anderson, "Why Cryptosystems Fail," *Communications of the ACM*, vol. 37, pp. 32-40, 1994.

[3] C. T. Beardsley, "Is Your Computer Insecure?," *IEEE Spectrum*, vol. 9, pp. 67-78, 1972.

[4] V. A. Brennen, "Cryptography Dictionary," vol. 2005, 1.0.0 ed, 2004.

[5] S. Brostoff and A. Sasse, "Are Passfaces More Usable Than Passwords? A Field Trial Investigation," presented at People and Computers XIV - Usability or Else! Proceedings of HCI 2000, Sunderland University, 2000.

[6] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," *Proceedings of the IEEE*, vol. 83, pp. 705-741, 1995.

[7] L. F. Cranor and S. Garfinkel, "Secure or Usable?," *IEEE Privacy & Security*, vol. 2, pp. 16-18, 2004.

[8] L. F. Cranor and S. Garfinkel, "Security and Usability: Designing Secure Systems that People Can Use." Sebastopol, CA: O' Reilly Media, Inc.. 2005.

[9] D. Davis, F. Monrose, and M. Reiter, "On User Choice in Graphical Password Schemes," presented at 13th Usenix Security Symposium, San Diego, CA, 2004.

[10] A. De Angeli, M. Coutts, L. Coventry, D. Cameron, G. I. Johnson, and M. Fischer, "VIP: A Visual Approach to User Authentication," presented at Working Conference on Advanced Visual Interfaces: AVI2002, Trento, Italy, 2002.

[11] Department of Defense Computer Security Center, "Department of Defense Password Management Guideline," Department of Defense, Washington, DC CSC-STD-002-85, April 12 1985.

[12] R. Dhamija and A. Perrig, "Deja Vu: A User Study. Using Images for Authentication," presented at 9th USENIX Security Symposium, 2000.

[13] P. Doyle and S. Hanna, "Analysis of June 2003 Survey on Obstacles to PKI Deployment and Usage," Organization for the Advancement of Structured Information Standards, Billerica, MA August 8 2003.

[14] S. M. Furnell, I. Papadopoulos, and P. S. Dowland, "A long-term trial of alternative user authentication technologies," *Information Management and Computer Security*, vol. 12, pp. 178--190, 2004.

[15] S. Granger, "Social Engineering Fundamentals, Part I: Hacker Tactics," vol. 2006: SecurityFocus, 2001.

[16] B. Ives, K. R. Walsh, and H. Schneider, "The Domino Effect of Password Reuse," *Communications of the ACM*, vol. 47, pp. 75-78, 2004.

[17] I. Jermyn, A. Mayer, F. Monrose, M. Reiter, and A. Rubin, "The Design and Analysis of Graphical Passwords," presented at 8th USENIX Security Symposium, Washington, DC, 1999.

[18] J. Liddell, K. Renaud, and A. De Angeli, "Using a Combination of Sound and Images to Authenticate Web Users," presented at 17th Annual Human Computer Interaction Conference: Designing for Society, Bath England, 2003.

[19] S. Man, D. Hong, B. Hayes, and M. Matthews, "A password scheme strongly resistant to spyware," presented at Int. Conf. on Security and Management, Las Vegas, NV, 2004.

[20] S. Man, D. Hong, M. Matthews, and J. C. Birget, "A shoulder-surfing resistant graphical password scheme," 2006.

[21] G. A. Miller, "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," *The Psychological Review*, vol. 63, pp. 81-97, 1956.

[22] K. D. Mitnick and W. L. Simon, *The Art of Deception: Controlling the Human Element of Security*. New York: John Wiley & Sons, 2002.

[23] National Research Council, *Who Goes There? Authentication Through the Lens of Privacy*. Washington, DC: National Academy Press, 2003.

[24] J. Nolan and M. Levesque, "Hacking human: data-archaeology and surveillance in social networks," *ACM SIGGROUP Bulliten*, vol. 25, pp. 33-37, ??

[25] L. O'Gorman, "Comparing Passwords, Tokens, and Biometrics for User Authentication," *Proceedings of the IEEE*, vol. 91, pp. 2021-2039, 2003.

[26] G. Orgill, G. W. Romney, and P. M. Orgill, "The Urgency for Effective User Privacy Education to Counter Social Engineering Attacks on Secure Computer Systems," presented at 5th Coneference on Information Technology Education (SIGITE '04), Salt Lake City, Utah, 2004.

[27] A. A. Ozok and S. H. Holden, "Alphanumeric and Graphical Authentication Solutions: A Comparative

Evaluation," presented at HCI International 2005, Las Vegas, NV, 2005.

[28] A. S. Patrick, A. C. Long, and S. Flinn, "HCI and Security Systems," presented at CHI 2003: New Horizons, Ft. Lauderdale, FL, 2003.

[29] R. W. Proctor, M.-C. Lien, K.-P. L. Vu, and G. Salvendy, "Improving Computer Security for Authentication of Users: Influence of Proactive Password Restrictions," *Behavior Reearch Methods, Instruments & Computers*, vol. 34, pp. 163-169, 2002.

[30] Real User Corporation, "How the Passface™ System Works," vol. 2005, 2005.

[31] K. Renaud and E. Smith, "Helping Users to Remember Their Passwords," presented at Annual Conference of the South African Institute of Computer Scientists and Information Technologists, Pretoria, South Africa, 2001.

[32] K. Renaud and A. D. Angeli, "My Password is here! An investigation into visio-spatial authentication mechanisms," *Interacting with Computers*, vol. 16, pp. 1017--1041, 2004.

[33] V. Roth, K. Richter, and R. Freidinger, "A PIN-entry method resilient against shoulder surfing," presented at Proceedings of the 11th ACM conference on Computer and communications security, Washington DC, USA, 2004.

[34] M. A. Sasse, S. Brostoff, and D. Weirich, "Transforming the 'Weakest Link'--a Human/Computer Interaction Approach to Usable and Effective Security," *BT Technology Journal*, vol. 19, pp. 122-131, 2001.

[35] L. Sobrado and J. C. Birget, "Shoulder-surfing resistant graphical passwords," Draft.

[36] W. C. Summers and E. Bosworth, "Password policy: the good, the bad, and the ugly," presented at Proceedings of the winter international symnposium on Information and communication technologies, Cancun, Mexico, 2004.

[37] R. C. Thomas, A. Karahasanovic, and G. E. Kennedy, "An Investigation into Keystroke Latency Metrics as an Indicator of Programming Performance," presented at Australasian Computing Education Conference 2005, Newcastle, Australia 2005.

[38] M. Turk, "A Random Walk Through Eigenspace," *IEICE Transactions of Information and Systems*, vol. E84-D, pp. 1586-1595, 2001.

[39] J. J. Turnage, "The Challenge of New Workplace Technology for Psychology," *American Psychologist*, vol. 45, pp. 171-178, 1990.

[40] L. Vasiu and I. Vasiu, "Dissecting Computer Fraud: From Definitional Issues to a Taxonomy," presented at 37th Hawaii International Confernces on System Sciences, Hawaii, 2004.

[41] D. Weinshall and S. Kirkpatrick, "Passwords you'll never forget, but can't recall," presented at ACM Conference on Computer Human Interaction (CHI) 2004, Vienna, Austria, 2004.

[42] A. Whitten and J. D. Tygar, "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0," presented at 8th Usenix Security Symposium, Washington, DC, 1999.

[43] S. Wiedenbeck, J. Waters, J.-C. Birget, A. Brodskiy, and N. Memon, "PassPoints: design and longitudinal evaluation of a graphical password system," *International Journal of Human Computer Studies*, vol. 63, pp. 102-127, 2005.

[44] R. J. Witty and K. Brittain, "Automated Password Reset Can Cut IT Service Desk Costs," Gartner, Inc., Stamford, CT G00123531, December 13 2004.

[45] R. J. Witty, "Bank of America Implements Simplified Single Sign-On," Gartner, Inc., Stamford, CT G00123465, January 25 2005.

[46] J. Yan, A. Blackwell, R. Anderson, and A. Grant, "Password Memorability and Security: Empirical Results," *IEEE Privacy & Security*, vol. 2, pp. 25-31, 2004.