Computers in Human Behavior 26 (2010) 705–715

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Effectiveness of image-based mnemonic techniques for enhancing the memorability and security of user-generated passwords

Deborah Nelson [a], Kim-Phuong L. Vu [b,*]

[a] Sage North America, 80 Technology Drive, Irvine CA 92618, USA
[b] Department of Psychology, California State University Long Beach, 1250 N Bellflower Blvd., Long Beach, CA 90840, USA

ARTICLE INFO

Article history:
Available online 9 February 2010

Keywords:
Human memory
Security
Passwords
Username-password
Mnemonics

ABSTRACT

Complex passwords are hard to remember, so people often pick simple passwords, write complex ones down, and reuse the same password across multiple accounts. Proactive password checking (PPC) restrictions and mnemonic techniques can enhance password security and memorability. Participants in this study were assigned to one of three password generation groups: PPC restrictions alone, image-based mnemonic, or text-based mnemonic. They were asked to generate and later recall passwords for five separate fictitious online accounts. The use of mnemonic techniques resulted in the generation of longer and more complex passwords. Furthermore, passwords were more accurately recalled when they were generated using the image-based mnemonic technique or PPC restrictions alone, as opposed to the text-based mnemonic technique. However, passwords generated using PPC restrictions alone were more easily forgotten and susceptible to being cracked. Thus, the image-based mnemonic technique was shown to be the most effective method for generating secure and memorable passwords.

1. Introduction

In 2009, there was an estimated 1.7 billion people using the Internet (http://www.internetworldstats.com/stats.htm). The use of the Internet to access personal accounts is convenient for the many people who use computers on a daily basis. Many users readily provide personal information to a site to perform online transactions without question, suggesting that they trust that their information is secure on the sites that they visit (Vu, Garcia et al., 2007). However, the cost of reported Internet security incidents was approximately $290 million in 2008 (Internet Crime Complaint Center, 2009), indicating that there is a need for security measures aimed at protecting a user's personal information.

The most basic security feature of many online accounts is user identification and authentication. To gain access to a specific account with this feature, a user must supply a username that identifies the account to be accessed. Authentication is achieved by having the user provide information, such as a password, to assure the system that the user is who that person claims to be (Schultz, 2005). Because passwords are oftentimes the only means of authentication for Web-based accounts, passwords must be kept secure by users if this method is to be successful. Yet, users readily admit to sharing passwords or giving out their passwords to other people (SafeNet, 2005). Password security issues can be illustrated with phishing schemes where, in one case, hackers successfully harvested over 40,000 MySpace passwords and posted them across several websites (Grimes, 2006). If the passwords collected from the MySpace accounts were used across more than one account and/or site, then each of the subsequent accounts protected by that password may in turn be at risk, because as Ives, Walsh, and Schneider (2004) pointed out, there is an obvious and probable overlap in the customer base for prominent online businesses, such as banks, e-mail services, top commerce sites, and social networking sites. The reuse of passwords across multiple accounts can lead to "a domino effect," where access to an important bank account is only as secure as the least secure site for which that same password is used, such as a library or social networking account. These studies illustrate why it is so important for each password to be unique or distinctly different from every other password.

When users are asked to create their own passwords, many pick weak ones that can be easily guessed or cracked. In one of the first studies addressing password security, Morris and Thompson (1979) compiled 3289 passwords from a UNIX time-sharing system. They classified 86% of the passwords they analyzed as "extremely weak," in that they were determined to be too short (an average four characters in length) and composed of only lower case letters, digits, or a combination of both lower case letters and digits. In addition, many of the passwords were easily found in dictionaries or lists of names. More recently, Florêncio and Herley (2007) analyzed the password habits of a half million Microsoft users and found that users managed an average of 25 password-protected

* Corresponding author. Tel.: +1 562 985 5021.
E-mail address: kvu8@csulb.edu (Kim-Phuong L. Vu).

0747-5632/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.chb.2010.01.007

accounts, but only used six to seven passwords, each of which was commonly shared across approximately four accounts. They also found that the average length of passwords was six characters, only a slight improvement over the four character length found in the Morris and Thompson study conducted 30 years earlier.

Florêncio and Herley's (2007) study also indicated that, much like the users of 30 years ago, modern users frequently used passwords consisting of letters typed all in one case, and rarely used numbers or special characters. When considering how users managed their passwords for accurate recall, people showed difficulty in password recall and often forgot their passwords. For Yahoo! accounts alone, it was estimated that each month, approximately 1.5% of users forgot their passwords, resulting in costly and time-consuming resets. Additionally, they found that password change/reset operations occurred 15% as frequently as sign-in operations. Florêncio and Herley surmised that while computer use and technology has grown by leaps and bounds during the past three decades, user behavior concerning passwords has changed little. People use memory, pieces of paper, and trial and error methods to remember the password for a specific account, and when that fails they result to requesting for their password to be reset.

Given that the human user is considered to be the weakest link in password security (Armstrong, 2003; Schneier, 2000), research has been devoted to assessing a myriad of password schemes aimed at improving security. Some of the more interesting password technologies include biometrics, keystroke rate, vocal verification, graphical, and image-based passwords (De Angeli, Coventry, Johnson, & Renaud, 2005; Ives et al., 2004; Stubblefield & Simon, 2004; Tullis & Tedesco, 2005; Wiedenbeck, Waters, Birget, Brodskiy, & Memon, 2005). Notably, each of these techniques involves the use of sophisticated, and often expensive, tools and/or software. Due to technological and cost restraints, the comparatively inexpensive alphanumeric password is likely to remain the popular choice for user authentication. Simply put, this means that user will likely remain responsible for creating and recalling passwords as they are needed.

Vu, Proctor et al. (2007) showed that there are two factors influencing the success of user name and password combination, security and memorability. However, there is a tradeoff between the two, in that more secure passwords are harder to remember. Brostoff and Sasse (2001) found that people tend to use familiar and easily accessible personal information when creating passwords, such as variations on one's own name, nickname, birthday, or address, because these are memorable (e.g., "dnelson69", which is a combination of the author's name and birth year). However, Vu, Bhargav, and Proctor (2003) found that the use of such personal information results in the creation of passwords that are not only easy for the user to remember, but are also more easily cracked. Similarly, Tullis and Tedesco (2005) found that the use of pictures, specifically personal pictures, allowed for participants to form a more personally meaningful and memorable association with a password, however, they also found that the use of personal pictures also made it easier for people who were familiar with the user to guess the password.

Klein (1990) analyzed over 15,000 user-submitted account entries from personal */etc/password* files, and also found that users choose short, simplistic passwords (e.g., name, birth date, favorite sport, etc.), that are vulnerable to guessing or brute force cracking. One way to prevent users from selecting easily guessable passwords is to assign strong ones. However, Klein noted that this approach actually may not produce the desired effect. Specifically, he noted that while pronounceable nonsense words (combinations of random syllables) might be slightly more memorable than completely meaningless system-generated passwords (characters randomly selected and organized into strings of varying lengths), neither has a meaningful personal association for the user. Thus, the user most likely would revert to writing the password down for later use. Klein called this behavior "akin to leaving the key under the door mat" (p. 8). Ausubel (1960) showed that systematic organization of "unfamiliar" elements and association of these elements to one's knowledge base can give meaning to them and improve their memorability. Thus, the memory difficulties with secure passwords may be overcome by considering cognitive factors influencing password memorability during password generation.

1.1. Memory and its implications for passwords

Psychological research on memory has implications for password recall in that several factors known to influence human memory may be exploited in order to enhance the memorability of user-generated passwords. The next sections will provide discussion of the multiple memory features that may be tapped to improve password recollection, including: episodic memory, the self-reference effect, the generation effect, chunking, superior memory for pictures, depth of processing, and mnemonic strategies.

1.2. Episodic memory and the self-reference effect

Studies of episodic memory, memory which receives and stores information about temporally dated episodes and events and temporal–spatial relations among such events, show that people use "landmarks" to guide their recall (Tulving, 1972). Groninger and Groninger (1988) conducted a series of experiments to look at whether autobiographical episodes were able to serve as mediators in the recall of words and found that life episodes served as very effective mediators in cueing certain types of words during the recall test. Words that had a unique time and place tag, and were therefore more distinctive; and words that were considered to be part of an easily generated set of cues involving the self-system were found to be more easily recalled. Episodic memories have been called the "gateway" to semantic memory (Squire & Zola-Morgan, 1998) because new information is always presented initially as part of some event in the person's life. In addition, people often create memory markers for personal events, such as a wedding date or the birth of one's child, as these are quite memorable landmarks in a person's life and carry a time-stamp. In order to preserve these memories, people may also take pictures to capture and "immortalize" their (episodic) experiences.

Similar to the effects of episodic memory, people remember information best when it is related to them, a phenomenon known as the self-reference effect. Rogers, Kuiper, and Kirker (1977) asked participants to determine which words on a provided list described them personally. Participants showed the highest levels of recall for the words they felt were self-descriptive. However, even the words that participants assessed as not describing themselves were recalled at high levels, simply as a result of the participant considering whether the words did or did not describe themselves. Greenwald and Banaji (1989) determined that the self-reference effect might be explained through the link between encoding and retrieval. That is, when individuals generate cues for retrieval, the cues that are seen to be associated with them are much more potent than other cues.

Familiarity and relevance of information to one's life appears to be a central feature of password selection (Riddle, Miron, & Semo, 1989). As noted earlier, it has been well-documented that users tend to generate passwords contrived from personal information or personal pictures (e.g., Bishop & Klein, 1995; Klein, 1990; Morris & Thompson, 1979; Proctor, Lien, Vu, Schultz, & Salvendy, 2002; Tullis & Tedesco, 2005). However, passwords based on personal information, whether alphanumeric or picture-based, are generally weak because they can be easily guessed by others. Perhaps by using information that is not biographical, but still personally relevant, a user can generate a password that would be memorable and more resistant to being guessed. In addition to well-known verbal descriptors, a well-known or related picture may also serve well as a cue for a password.

1.3. The generation effect

Self-generated words are better remembered than those that are externally presented, a phenomenon known as "the generation effect" (Slamecka & Graf, 1978). Slamecka and Graf conducted five experiments comparing participants' memory for self-generated words and those that were simply presented to the participants to be read. Across all five of the experiments, results indicated that participants were better able to recall the words they produced themselves as compared to those words which they simple read (see deWinstanley & Bjork, 2004).

The generation effect extends to password memorability as well (Vu, Proctor et al., 2007). Vu et al. had participants generate passwords using a proactive password checking technique that allows the system to place adjustable restrictions on acceptable passwords to improve their security. Typically, the system would require that a password contain at least one lower and one uppercase letter, at least one number, and at least one special character, but the actual restrictions are left to be decided by the system administrator. Thus, the system may be programmed to reject passwords that look like license plates, are the same as the username, or are composed of common words. The benefit of proactive password checking over assigning secure passwords to users is that through the processing of generating the password, it is more likely that the password would be accurately recalled at a later time. The increase in memory is due to the generation effect itself, and to the user being able to associate the password (or components of the password) to personally relevant information that can cue recall later.

Additionally, the simple act of generating the password, entering it, and successfully confirming it the first time forces the user process the password at several points and can reinforce associations for later retrieval. However, usually more than one recall/confirmation attempt is needed to improve memory for the password (Vu, Cook, Bhargav, & Proctor, 2006). One drawback known to occur with proactive password checking, though, is that while studies have indicated that such restrictions do not make it more difficult for users to remember a single password (Proctor et al., 2002), it has been found that users have more difficulty accurately recalling multiple passwords (Vu, Proctor et al., 2007). Thus, proactive password generation needs to be paired with other memory techniques, such as mnemonic techniques (described later), to improve the recall of multiple secure passwords.

1.4. Chunking

Memory can be enhanced by actively combining information together into bits or chunks. Because the combining process allows people to relate individual pieces of information into a global unit, the unit can be remembered more easily than the individual components. For example, the string of 24 letters, HUMANCOMPUTER-INTERACTION, can be combined into one topic of human–computer interaction. Miller (1956) concluded that people can reliably remember seven individual items, plus or minus two. This restriction is applicable across many stimulus types, including letters, numbers, and even musical notes. By meaningfully recoding information into larger and larger "chunks," a large mass of information can be worked into manageable a form. "Meaningful" recoding provides the person with understanding (Ausubel, 1960), which might literally reduce the amount of information to-be-remembered – a summary instead of an entire passage, and reduce the load on memory. However, if the information is not chunked in a meaningful manner, error may occur in the "unpacking" of the chunk (Vu, Proctor et al., 2007). Vu, Proctor et al. found that participants had difficulty recalling where they had integrated numbers and special characters in their passwords, because they were placed randomly. However, when the participants mindfully placed the numbers and special characters in meaningful chunks of information, memory for their generated passwords was improved. Thus, successful use of a chunking strategy requires a meaningful combination of the material and a good organizational pattern. Since chunking can be applied to memory of words and pictures, it can be applied to password generation techniques that rely of either alphanumeric characters or pictures.

1.5. Depth of processing

Craik and Lockhart (1972) introduced a view of the memory process that accounts for the involvement of a short-term memory system in information processing. Craik and Lockhart believed the system would allow a person to process information on many levels, ranging from simply taking note of the visual characteristics of a printed word, through rehearsing it or paying attention to its sound, up to elaborate coding in terms of its meaning. Craik and Tulving (1975) conducted a series of experiments to determine how different levels of information processing affected recall. Participants were given lists of words, with each word proceeded by a question, intended to encourage three distinct levels of processing, in progressive order of depth: physical, acoustic, and semantic. As noted earlier, processing on the physical level entailed encoding the words by focusing on the visually apparent features of the letters of the presented words (e.g., all lower case letters or all upper case letters). Processing on the acoustic level entailed focusing on the sound combinations associated with the letters of the words (e.g., rhyming). Processing on the semantic level entailed focusing on the meaning of the presented word (e.g., is the word a member of a given category?). The results indicated that that the main determinant for learning was the type of processing that occurred. Specifically, participants who processed items with a "deeper" level of meaning (e.g., determining if the word fit in a particular category) showed much better retention than those who engaged in a "shallow" or superficial processing method (acoustic or visual feature matching).

In regards to passwords, users are often asked to come up with passwords on the spur of the moment in order create an account on a password protected site to which they would like to gain immediate access. However, because depth of processing can be applied to both alphanumeric and image-based passwords, users should devote time to think about passwords at a deeper level (i.e., relating them to their existing knowledge base) rather than coming up with a quick and superficial password that would likely be forgotten after a few days.

1.6. Superiority of pictures over words

Paivio (1971) proposed that there are two memory systems, one for the encoding of verbal or symbolic information and the other for images. This dual-coding hypothesis suggests distinct differences between memories for verbal information and images. For example, information about meaning may be better retrieved by verbal labels, whereas, information about size and shapes may be better retrieved through imagery. Several studies have shown that people have superior memory for images/pictures as compared to words (Shepard, 1967; Standing, Conezio, & Haber, 1970). Shepard indicated that neither words nor sentences were recognized as accurately as were images, a finding he termed the "picture superiority effect." Participants in Shepard's study were able to recognize images with extremely high accuracy, up to 98.5%, even following a 1 h delay (see also, Paivio, Rogers, & Smythe, 1968).

Graphical or image-based passwords, passwords that involve a user's recognition of selected pictures, or multiple click points on a single picture, has been posited as an alternative authentication method to alphanumeric passwords. Some researchers studying passwords believe that this type of user authentication may solve several of the security issues concerning alphanumeric passwords (De Angeli et al., 2005; Dhamija & Perrig, 2000; Stubblefield & Simon, 2004; Wiedenbeck et al., 2005). Graphical or image-based passwords are said to be much more difficult to crack because they are stored by secure means within the system and they cannot so easily be written down and insecurely stored (De Angeli et al., 2005; Stubblefield & Simon, 2004; Wiedenbeck et al., 2005).

Click-point graphical passwords were proposed by Blonder (1996). With this technique, the user clicks on selected regions or images of a picture appearing on a screen. Failure to replicate the click pattern successfully would mean the user was not authenticated and was rejected access to the site. More recently, Wiedenbeck et al. (2005) proposed a graphical password system called PassPoints. Participants created a password by clicking on five user-determined target points within a picture displayed on a computer screen. Weidenbeck et al. suggested that their scheme was more flexible than previously proposed visual-based password schemes because it allowed the user to select an image provided from a pool stored on the system, or they had the option of uploading and using an image of their own. The only specified requirement was that the image must be complex enough to provide multiple potential click points. Results indicated that while participants recognized their image and the general location of their selected click points, overall they had difficulty clicking in defined pixel areas on full-size computer screens. Therefore, it is difficult to imagine this type of graphical password system working on the fraction of screen space available on mobile devices. Due to such space constraints, it is doubtful that it will be a popular choice among users or organizations any time soon. It would be more feasible for people to use pictures or imagery to cue associated alphanumeric passwords.

1.7. Mnemonic strategies

The use of mnemonic strategies (e.g., interactive imagery, method of loci, keyword system) has been shown to greatly improve memory (see, e.g., Bellezza, 1981; Roediger, 1980; Snowman, McCown, & Biehler, 2008). Mnemonic strategies invoke deep processing because the to-be-remembered items are processed for meaning and associated with well-learned knowledge structures, such as a rhyme or well-traveled route. Baddeley (2004) noted that imaging plays an important part in mnemonic strategies, which is one reason for why mnemonics lead to improved memory in the long run.

Yan, Blackwell, Anderson, and Grant (2005) surveyed approximately 288 university students to assess password memorability and security. A third of the participants were asked to choose a password (given certain password restrictions), another third were assigned random passwords, and the remaining third were asked to choose a password using a mnemonic based on a phrase. Results indicated that the passwords generated by the survey participants were similar in length (on average 7–8 characters) and that only passwords generated by participants in the mnemonic category contained digits or special characters. Additionally, the phrase-based mnemonic passwords were found to be just as hard to crack as the assigned random passwords, but were also just as easy for the participants to recall as were the simplistic naïve-user password selections. Unfortunately, Yan et al. did not provide any examples of the passwords generated by the survey participants.

Another mnemonic strategy that has been used with passwords is the first-letter technique (Vu et al., 2003). Vu et al. asked participants to generate a sentence that had personal meaning, and then to take the first letter of each word in that sentence and combine them, in order, to make a password. To enhance the security and reduce the crackability of the passwords, they required half of the participants to embed at least one special character and at least one number in their password. Results indicated that these extra restrictions did indeed effectively reduce the crackability of the user-generated first-letter passwords. However, the improved security came at the cost of both increased generation and recall time, as well as increased errors committed during password input.

Subsequently, Vu, Proctor et al. (2007, experiment 3A) compared passwords generated using the first-letter mnemonic technique with those generated using a whole-word (passphrase) mnemonic strategy. Vu, Proctor et al. expected to find that passwords created using the whole-word method would be better recalled than those generated using the first-letter technique because the increased meaningfulness of the password would enable stronger encoding. It was also anticipated that the whole-word method would better enable participants to successfully recall where they had embedded the digit(s) and special character(s) in the generated password. However, results indicated that there was little difference between the whole-word mnemonic technique and the first-letter mnemonic technique with regards to generation time, recall time, or security (crackability). Participants had more difficulty recalling the passwords generated using the whole-word mnemonic strategy, due apparently to difficulty in recalling exactly where in the password digits and special characters had been incorporated.

Image-based password mnemonics can also be used to tap into humans' superior memory for pictures. De Angeli et al. (2005) assigned participants a series of images to memorize as their visual passcode and then later asked them to recognize those images from a wider set of distracter images. Results indicated that the participants used a variety of mnemonic techniques for remembering their passwords, but that there were differences in the techniques used by participants tasked with remembering numbers, such as a typical 4-digit ATM code, as compared to those who had to recall pictures. For the numbers; repetition, chunking, and associations with dates or math were used to facilitate memory, whereas association was the most prevalent (98%) strategy for pictures, sometimes supported by repetition. Interestingly, most of the participants indicated that they associated their pictures with words, often creating a story to support sequence retrieval.

Imagery can help associate to-be-remembered items with the knowledge structure used in many mnemonic techniques to improve the memory. Thus, perhaps the detrimental effects of embedding digits and special characters into passwords using the whole-word mnemonic technique described earlier could be alleviated if the incorporation of digits and special characters is more effectively encoded with imagery during password creation.

1.8. Summary

When applied to the password generation process, the aforementioned features of human memory may help us to understand and mitigate the memory problems associated with password recollection. For example, we can now see that several memory benefits can be tapped by having a user generate their own password by chunking personally relevant information in a meaningful manner. This process alone should provide for deeper encoding of the new material, thus improving subsequent recall. Although the picture superiority effect implies that graphical passwords may be a better alternative to alphanumeric passwords, their use is limited by the system constraints of some computing devices, especially handheld devices. The use of mnemonic strategies should provide the user with an organizational structure to their passwords, which should in turn improve both their short-term and long-term recollection of the generated passwords. As stated earlier, the main problem indicated with use of the first-letter and entire-word mnemonic technique for generating passwords is that users are not able to recall special characters and digits embedded into the password.

Memory for digits and special characters in alphanumeric passwords may be improved if uers base the special characters and digits on information derived from a well-known personal picture. That is, pairing imagery with the more widely used alphanumeric password technique may provide a methodology that results in more memorable user-generated passwords. To test this hypothesis, three different password generation techniques for will be examined in the present study. The first technique required participants to adhere to guidelines aimed at improving the security of the password. The other two techniques relied on mnemonics to provide a structure for transforming verbal descriptors and pictures into meaningful units to improve memorability, while adhering to the guidelines known to improve password security. It was hypothesized that the two mnemonic techniques would lead to better memory of passwords. Furthermore, because pictures can be remembered better than words, it was hypothesized that the image-based mnemonic technique will be better than the text-based one.

1.9. Current study

The current experiment considered the effectiveness of a password method using proactive password checking (PPC) restrictions alone as a guide for generating passwords, compared with the use of an image-based mnemonic technique plus PPC restrictions, and a text-based mnemonic technique plus PPC restrictions. Although the generation technique varied depending on the group to which the participant was assigned, all participants were asked to generate their own uniquely different passwords for each of five different fictitious online accounts. Participants were measured on the amount of time needed to generate acceptable passwords. At two specific times, 10-min and 1-week following the password generation phase, participants were tested on their ability to recall the passwords they had generated for each account. In addition, the strength of each of their passwords was assessed as well as the number of passwords forgotten over the week delay.

2. Method

2.1. Participants

Seventy-eight participants from the Southern California region were recruited from fliers posted at California State University Long Beach and the surrounding community. These fliers specified that the participants would be taking part in a research study and that they were to bring five personally-relevant pictures with them to the first session. Twenty-six participants were randomly assigned to the image-based mnemonic group, 26 to the text-based mnemonic group, and the remaining 26 to the PPC restrictions alone group.

2.2. Design, apparatus, and procedure

This study used a 3 (generation technique: image-based mnemonic, text-based mnemonic, or PPC alone) by 2 (recall delay: 10-min or 1-week delay) mixed design. Generation technique was the between-subjects variable and recall delay was the within-subjects variable. The experimental apparatus closely modeled that used in the research conducted by Vu, Proctor et al. (2007). The program for this experiment was written in Java and ran on a personal computer with a 14 in. monitor. The Java program was used to present instructions to the participants, check that the generated passwords met the set criteria (Each password must be unique and cannot be an existing password, your name, username, or ID number. The password must be at least 8 characters in length and have at least an upper case letter, a lowercase letter, a digit, and a special character.), and record participants' responses. Generation and recall times were measured in seconds by Java code (System.currentTimeSeconds).

The experiment consisted of two sessions, 1 week apart, and was broken up into three parts. The first two parts of the experiment occurred during session 1, and the third part of the experiment occurred during session 2. At the start of the first session, each participant was provided with a brief written introduction to the experiment, asked to sign a consent form, and give the researcher their five pictures. All pictures were digitized for storage and logged with the data collected from each participant.

All participants were tested individually in a quiet room. If the participant was assigned to the image-based mnemonic group, then his/her pictures were placed on the table and the participant was asked to write the answers to some basic questions about each of the pictures as best as possible (e.g., brief title of the picture, a brief description of the picture, and why it was important to them or why they brought it in). Once completed, the experimenter collected each description sheet. However, if the participant was assigned to the text-based mnemonic group, or the PPC restrictions group, then he/she skipped this step of the procedure.

All participants were verbally informed that they would be asked to generate, and later recall, a unique password for each of five different fictitious online accounts: a bank account, a bookstore account, an e-mail account, a social networking account, and a computer account. Participants were instructed not to write down any of their passwords during the course of the experiment. Prior to doing any work on the computer, the participants were provided both printed and verbal directions of the experimental task based on their experimental group.

Participants assigned to the PPC restrictions alone group were asked to generate passwords that met nine proactive password checking criteria. Participants in the image-based mnemonic group were instructed to pair each of the pictures with one of the each of the five fictitious online accounts. They were then shown how mnemonic techniques were used to generate passwords based on the image in each of two sample pictures. The first example of the image-based mnemonic technique used a picture of a young man with the phrase "My boyfriend Matt" written underneath the image. After the participant had approximately 10 s to take in the image and statement, the experimenter explained step by step how particular personal knowledge involved with the picture could be incorporated into password, such as the fact that a person "dates" their boyfriend, and so the phrase "I date Matt" could be used as a basis for a password.

The experimenter then explained how pieces of information could be transcribed in a manner that increases complexity while maintaining meaningful to the participant. Specifically, the experimenter pointed out Matt's eyes and remarked on how they were bright blue. The experimenter explained how that remarkable feature could be translated in the password through replacing "I" with "Eye" and capitalizing the "E" similar to how one would capitalize the "I," or because Matt's eyes "popped" out of the picture. Additionally, the experimenter explained how letters and numbers or symbols could be combined in meaningful ways, such as combining "D" and "8" to represent the word "Date." Through the explanation of the example, the experimenter provided the participant all the steps involved in transforming information associated with the picture of Matt into the complex password "EyeD8M@tt". A second example of the image-based mnemonic technique was illustrated using a picture of a calico cat with the name "Kali" written under the image. Again, participants were shown how bits of personal information associated with the picture could be incorporated into a password. The experimenter explained that when she was little Kali stuck her nose in a candle flame, she has had many seizures, has cost her owner a lot of money in vet bills, and was named after the goddess Kali. The experimenter pointed out how the picture of Kali serves to cue the password "G0d$$ofF!re". As with the participants in the PPC group, a sheet of paper listing the proactive password checking requirements was also provided.

Participants in text-based mnemonic plus proactive password restrictions group were provided directions for how to use text-based mnemonic techniques to create passwords based on self-generated phrases. The instruction sheet explained the mnemonic strategy through the two separate examples described above, except they were provided without the pictures. As with the previous two groups, a sheet listing the proactive password checking requirements was provided.

2.2.1. Generation of passwords

Each participant was assigned a unique username for use during the experiment. Following input of the username, the participant was presented with a prompt to enter a password for one of the five accounts. If the generated password met all of the requirements, then the participant was informed that the password had been accepted for that account, the screen scrolled down to where the generated password was out of view and was asked to retype the password. Following successful recall of the password for that account, the program then prompted the participant to move onto the next account. However, if the password did not meet one or more of the restrictions, a prompt to reenter the password was presented along with a list of the restrictions that were not met by the attempted password entry. Additionally, if the password was accepted but then not accurately recalled, the participant was shown the incorrectly entered password along with the accepted password and then prompted to retype the correct password.

The generation time and number of attempts needed to correctly generate each of the five passwords was recorded and sent directly to a log file. The participants were not informed that generation times were being measured, and thus were not provided feedback about their password generation times. Once acceptable passwords were generated for all five accounts, a list containing each account name and the associated password was printed for the participant to review for 30 s before the program was closed out and the monitor turned off. The participant was then provided with a sheet of paper containing simple math problems and was instructed to complete as many of the problems as possible within 10 min. The participant was informed that he/she could skip around and choose which problems to work on. At the end of 10 min, the participant was asked to step outside for a moment while the experimenter reset the computer for the next task.

2.2.2. Short-term recall

Once the computer was reset the participant was instructed to come back in the room and retake the seat he/she had been in during the first part of the experiment. The researcher informed the participant that he/she would now be asked to recall each of the passwords he/she had generated for the five accounts. The participant was told that one-by-one, a total of four times each, the account names would randomly appear on the screen, and that his/her task was to recall and enter the correct password for that account, with a maximum of ten tries to correctly enter the correct password before the program would automatically move him/her onto the next account. For each account occurrence, the login time and number of incorrect attempts were recorded. At the completion of this part of the experiment and prior to leaving, the participant was again instructed not to write down the passwords that he/she had generated during the study.

2.2.3. Long-term recall

For the second session and third part of the experiment, participants returned 1 week later and were once again asked to recall the passwords they had generated for each account. The procedure was identical to the second part of the experiment. At the end of the third part of the experiment the participants were asked to complete a demographic questionnaire, a post-study recall questionnaire, and were provided with a debriefing form. Participants were provided the opportunity to have questions regarding the experiment answered and thanked for their participation.

3. Results

Demographic data for the participants are presented in Table 1 and only summarized here. The study consisted of 78 participants, ranging in age from 18 to 62 years old ($M = 30.2$, $SD = 11.3$). Approximately two-thirds of the participants were female. Participants widely varied in their reported years of computer experience ($M = 14.4$, $SD = 6.3$, Range = 3–38).

Participants' report of the number and characteristics of their password-protected accounts also varied. On average, they reported having 14.9 ($SD = 24.9$, Range = 2–208) password-protected accounts, used the same password across 7.2 ($SD = 10.7$, Range = 0–88) accounts, and had 5.85 unique passwords ($SD = 6.6$, Range = 1–38), of which an average of 2.1 ($SD = 3.0$) of the passwords could be personally associated with them. Ninety-six percent ($n = 74$) of the participants reported using numbers in their personal passwords, while 56% ($n = 44$) reported using both upper and lower case letters, and only 32% ($n = 25$) reported using special characters. Additionally, 94% ($n = 73$) of the participants reported that they believed that the requirement of a password for an online account increases security, and 82% ($n = 64$) considered password policies to be at least somewhat important.

Consideration was also given to the password generating habits of the participants during the course of the current study. Across all three generation groups, PPC alone, text-based mnemonic, or image-based mnemonic, participants generated passwords ranging in length from 8 to 26 characters ($M = 10.62$) for each of their five accounts. On average, participants incorporated 1.32 (Range = 1–5) special characters, 1.66 (Range = 1–7) digits, and 1.60 (Range = 1–6) non-dominant characters into their generated passwords.

3.1. Generation time and number of attempts

Participants who use deep processing tend to take a longer time than participants who use shallow processing (Craik & Tulving, 1975). Thus, generation time was recorded to determine how much time users spent thinking about their password prior to submitting it to the system. A univariate ANOVA was conducted on mean generation time for the five passwords as a function of generation technique. For the current and subsequent analyses, the p-value was set to a .05 significance level and partial eta squares (η^2) are reported for effect sizes. There was a significant effect of generation technique, $F(2,75) = 9.82$, $\eta^2 = 0.21$. Follow-up Tukey HSD post hoc analysis indicated that password generation times were significantly longer for participants in the text-based group ($M = 156$ s) than for those participants in either the image-based group ($M = 89$ s) or the PPC group ($M = 67$ s). No other pairwise comparisons were significant (see Table 2a).

Table 1
Demographic characteristics ($N = 78$).

Characteristics	*n*	%	*M*	*SD*	Range
Age (years)	78	100	30.205	11.287	18–62
Gender					
Female	48	61.5	–	–	–
Male	30	38.5	–	–	–
Computer experience (years)	78	100	14.436	6.297	3–38
Computer knowledge	78	100	3.423	.814	–
1 = Novice	3	3.8	–	–	–
2 = Some knowledge	3	3.8	–	–	–
3 = Average knowledge	34	43.6	–	–	–
4 = Above average knowledge	34	43.6	–	–	–
5 = Expert	4	5.1	–	–	–
Knowledge of computer security	78	100	2.603	1.061	–
1 = Novice	11	14.1	–	–	–
2 = Some knowledge	28	35.9	–	–	–
3 = Average knowledge	24	30.8	–	–	–
4 = Above average knowledge	11	14.1	–	–	–
5 = Expert	4	5.1	–	–	–
Number of managed password-protected accounts (accounts)	78	100	14.872	24.91	2–208[a]
Number of accounts where password is reused (accounts)	78	100	7.192	10.689	0–88[a]
Number of unique passwords (passwords)	78	100	5.846	6.6	1–38[a]
Complexity: uses numbers					
Yes	75	96.1	–	–	–
No	3	3.9	–	–	–
Complexity: uses special characters					
Yes	25	32.1	–	–	–
No	53	67.9	–	–	–
Complexity: uses both upper and lower case					
Yes	44	56.4	–	–	–
No	3	43.6	–	–	–
Number of personally associated passwords (passwords)	78	100	2.064	3.008	0–15
Number of passwords that are a name or word followed by a number (passwords)	78	100	3.526	4.58	0–30
Importance of password policies	78	100	3.885	.939	–
1 = Not at all important	3	3.8	–	–	–
2 = Minimally important	5	6.4	–	–	–
3 = No opinion	6	7.7	–	–	–
4 = Somewhat important	48	61.5	–	–	–
5 = Very important	16	20.5	–	–	–
Does the requirement of passwords increase security?					
Yes	73	93.6	–	–	–
No	5	6.4	–	–	–

[a] *Note:* Two of the participants were in managerial positions that required them to manage over 100 password-protected accounts.

Table 2
Mean password generation time (in s) as a function of generation technique.

Generation technique	Sample size	Mean (s)	Standard deviation
Image-based mnemonic plus PPC	26	89.069	42.288
Text-based mnemonic plus PPC	26	156.762	118.342
PPC alone	26	67.915	36.299

The number of attempts needed by a participant to generate an acceptable password was also recorded and then averaged across the five accounts. A one-way ANOVA was conducted on the mean number of attempts as a function of generation technique. There was no significant effect of generation technique, $F(2,75) = 1.01$, $\eta^2 = 0.03$. In general, participants were able to generate an acceptable password within 1–2 attempts.

3.2. Recall time

Recall time was used as a measure of ease of memory retrieval. More memorable passwords tend to be recalled faster than less memorable ones (Vu, Proctor et al., 2007). A 3 (generation technique) × 2 (recall delay) repeated measures ANOVA was performed on mean password recall times (see Table 3 for means). Recall delay was the within-subjects factor and generation technique was the between-subjects factor. There was a main effect of generation technique, $F(2,75) = 3.85$, $\eta^2 = 0.09$. Follow-up Tukey HSD post hoc analysis indicated that participants in the text-based group ($M = 32.78$ s) took significantly longer to recall their passwords than did those participants in the PPC group ($M = 15.674$ s). No other pairwise comparisons were significant.

There was no significant effect of recall delay on recall time, $F(1,75) = 1.73$, $\eta^2 = 0.02$, and the interaction between the recall delay and generation technique was not significant either, $F(2,75) < 1.0$, $\eta^2 = 0.01$.

3.3. Recall attempts

The mean number of attempts needed for successful password recall was calculated for each participant and submitted to a 3 (generation technique) × 2 (recall delay) repeated measures ANOVA (see Table 4 for means). There was only a significant main effect

Table 3
Mean time (in s; standard error in parentheses) to correctly recall and enter a generated password as a function of generation technique and recall delay.

Generation technique	Sample size	10-min delay time (s)	1-week delay time (s)
Image-based mnemonic plus PPC	26	28.507 (5.861)	24.488 (2.916)
Text-based mnemonic plus PPC	26	32.780 (5.861)	27.251 (2.916)
PPC alone	26	15.674 (5.861)	14.459 (2.916)

Table 4
Mean number of attempts needed (standard error in parentheses) to correctly recall and enter a generated password as a function of generation technique and recall delay.

Generation technique	Sample size	10-min delay attempts (s)	Week delay attempts (s)
Image-based mnemonic plus PPC	26	1.532 (.197)	1.557 (.161)
Text-based mnemonic plus PPC	26	2.145 (.197)	2.012 (.161)
PPC alone	26	1.329 (.197)	1.361 (.161)

of generation technique, $F(2,75) = 6.04$, $\eta^2 = 0.14$. Follow-up Tukey HSD post hoc analysis indicated that it took significantly more attempts for the participants in the text-based group ($M = 2.15$ attempts) to accurately recall their passwords than it did for participants in either the image-based group ($M = 1.53$ attempts) or the PPC group ($M = 1.33$ attempts), which did not differ from each other.

3.4. Forgetting

The number of passwords forgotten by each participant for one or more accounts was also examined. A univariate ANOVA was conducted with generation technique as a between-subjects factor (see Table 5 for means). There was a significant effect of generation technique, $F(2,75) = 5.58$, $\eta^2 = 0.13$. Follow-up Tukey HSD post hoc analysis indicated that participants assigned to the text-based group ($M = 1.08$ passwords) forgot significantly more of their generated passwords than did those participants in the image-based group ($M = .12$ passwords). No other pairwise comparisons were significant.

3.5. Password length

All passwords were compiled and the mean password length was determined for each participant. These means were submitted to a univariate ANOVA, with password length as the dependent measure and generation technique as a between-subjects factor (see Table 6 for means). There was a significant effect of generation technique, $F(2,75) = 6.80$, $\eta^2 = 0.15$. Follow-up Tukey HSD post hoc analysis indicated that participants in the PPC group ($M = 9.40$ characters) generated significantly shorter passwords than did participants in either the image-based group ($M = 11.22$ characters) or the text-based group ($M = 11.23$ characters), which did not differ from each other.

Table 5
Mean number of forgotten passwords as a function of generation technique.

Generation technique	Sample size	Mean (passwords)	Standard deviation
Image-based mnemonic plus PPC	26	.12	.431
Text-based mnemonic plus PPC	26	1.08	1.440
PPC alone	26	.46	1.029

Table 6
Mean password length as a function of generation technique.

Generation technique	Sample size	Mean length (characters)	Standard deviation
Image-based mnemonic plus PPC	26	11.215	2.673
Text-based mnemonic plus PPC	26	11.231	3.691
PPC alone	26	9.400	1.828

3.6. Password complexity

Because differences in the length and complexity between the passwords generated in each of the three conditions may influence crackability rates, it was determined that it would be informative to examine password length and complexity. Password length was easily obtained, but there were no complexity measures that considered all of the factors contributing to crack-resistant passwords used in this study. As a result, a complexity index was created for exploratory purposes that used the following criteria: generated passwords were assigned one point of complexity for each inclusion of a digit, special character, or non-dominant case letter (use of an upper or lower case character when use of the opposite case was predominant in the password structure). An additional point of complexity was granted if the base "word" of the constructed password could not be found in the dictionary (e.g., was not an actual word in a dictionary). For example, the password "Venice#1" uses "Venice," a word that can easily be found in a dictionary, as the base "word." In contrast, the password "Clownluv#1" concatenates the words "clown" and "luv," forming a base "word" for the password that cannot be found in a dictionary. The latter password received an additional point for complexity.

Scores based on this password complexity index were computed for each password, and mean complexity scores for all participants were submitted to a univariate ANOVA with generation technique as a between-subjects factor (see Tables 7 and 8). Again, there was a significant effect of generation technique on the complexity of generated passwords, $F(2,75) = 14.85$, $\eta^2 = 0.28$. Follow-up Tukey HSD post hoc analysis indicated that participants in the PPC group ($M = 4.32$ units of complexity) generated significantly less-complex passwords than did participants in either the image-based group ($M = 5.87$ units of complexity) or the text-based group ($M = 5.73$ units of complexity), which did not differ from each other.

3.7. Password crackability

The crackability of the generated passwords was assessed by the Password Recovery Tool Kit (PRTK). Each password was used

Table 7
Mean password complexity as a function of generation technique.

Generation technique	Sample size	Mean complexity (complexity units)	Standard deviation
Image-based mnemonic plus PPC	26	5.869	1.567
Text-based mnemonic plus PPC	26	5.731	1.603
PPC alone	26	4.323	1.043

Table 8
Range of complexity differences for passwords in terms of generation technique.

Generation technique	Length	Digits	Special characters	Letters from non-dominant case
Image-based mnemonic plus PPC	8–21	1–6	1–3	1–3
Text-based mnemonic plus PPC	8–26	1–6	1–5	1–6
PPC alone	8–19	1–4	1–2	1–2

to encrypt a word document. The document was then submitted to the PRTK in an attempt to recover the password. All passwords generated in this study were randomly selected and submitted to the PRTK for a 12 h period. The passwords in the PPC group were more susceptible to being cracked ($n = 8$) than were passwords generated using the image-based ($n = 0$) or text-based ($n = 0$) mnemonic methods. 6.15% of the PPC group's total number of passwords ($n = 130$) were cracked.

4. Discussion

The results from the current study show that the use of a mnemonic technique during password generation improves users' memory for their generated passwords. The increase in memorability is a result of the participants providing meaning and structure to the characters in the password. The image-based mnemonic technique was shown to be more effective than the text-based one. Because memory for images is better than memory for words, the process of relating the password elements to the image improved its memorability. Thus, as hypothesized, the use of images helped users generate passwords more effectively than just text-based descriptions alone. In the following sections, we discuss how the generation technique influenced password memorability and security.

4.1. Influence of password generation technique on resulting passwords

Passwords generated by participants in the image-based and text-based groups were longer and more complex than passwords generated by participants in the PPC group. Passwords generated by the PPC group were typically single words close to eight characters in length that could be found in a dictionary, with relatively predictable substitutions made to meet the digit and special character inclusion specifications (e.g., "Fri3nds!", "Soccer2*", or "H0tm@il"). However, passwords in both of the mnemonic conditions were typically concatenations of phrases, with multiple inclusions of digits, special characters, and/or use of the non-dominant case (e.g., "M3y3*Space*", "L3ft<3inSF", "Be@rddr@g0nFruit"). The increased complexity of passwords generated with the mnemonic techniques are likely due to participants using deeper processing to relate the password elements to the sentence they generated to form the password.

Because of the differences in the password characteristics, there was a wide variation within each group as to the time taken by each participant to generate their passwords. The amount of time needed for a participant to generate their passwords was positively correlated to both the length ($r = .26$, $p = .02$) and complexity ($r = .36$, $p < .01$) of the generated passwords. Thus, it took longer for participants to generate passwords that were long and complex than passwords that were short and simple. However, the amount of time needed for participants to accurately generate passwords for the PPC restrictions alone group was not significantly different for participants who used the image-based mnemonic technique (21 s different). This is surprising because the use of the image-based technique should require deeper processing, which should take more time than shallow processing (Craik & Tulving, 1975). The fact that, using the image-based mnemonic technique, passwords were generated more quickly speaks to the strength of associating passwords with well-known personal pictures. One reason why the image-based technique allowed for the information to be processed at a deep, but efficient level is that the contents of the pictures were familiar to the participants, providing them a foundation from which to generate new passwords. On the other hand, participants in the text-based condition had to create the base sentences used for their passwords from scratch, requiring more time to generate and process the information.

4.2. Influence of password generation technique on passwords recall

It was expected that the participants provided only with the PPC restrictions would take the most time to accurately recall their generated passwords because this generation technique does not provide participants with an organized structure to give meaning to the password elements (see, e.g., Ausubel, 1960). However, it was found that participants who used the text-based mnemonic technique took nearly 15 s longer to recall their passwords than did participants who used the PPC restrictions alone. As noted earlier, the difference in the length and complexity of the passwords across the three generation groups is an important factor for understanding group differences in recall times. That is, longer passwords take more time to input than shorter passwords, and more complex passwords will take longer to input than less-complex passwords. This may partially explain why, overall, it took participants in the mnemonic conditions longer to recall ($M = 28$ s) their passwords compared to participants in the PPC group ($M = 15$ s). However, this does not provide an explanation for why the text-based group took slightly longer to recall and input their passwords ($M = 30.02$ s) as compared to participants in the image-based group ($M = 26.50$ s) since the passwords generated by the two groups did not differ in terms of length or complexity. Moreover, participants in the image-based group were able to enter their passwords approximately 3.5 s faster than participants in the text-based group. A likely reason why the participants in the image-based group were able to recall their passwords slightly faster than participants in the text-based group is that the pictures effectively served as a cue and lessened the amount of time needed by those participants to accurately recall their passwords.

Although mnemonic techniques were hypothesized to help participants recall their passwords, participants in the text-based group had more difficulty recalling their passwords than participants in the image-based or PPC groups. One reason why participants in the text-based group performed worse than the other two groups is that they were unable to accurately recall the placement of case changes, digits, or special characters within their generated passwords. Errors of these types occurred 57.4% of the time in the inaccurately entered passwords that were generated by the text-based mnemonic group, but only 24.1% and 31.4% of the time in the inaccurately passwords that were generated by participants in the image-based mnemonic and PPC restrictions alone groups, respectively.

The length of recall delay did not have a significant effect on the amount of time or number of attempts needed for the participants to recall their passwords. The lack of a time delay effect may be due to the 1 week interval being too short or to the fact that the short-term recall phase actually helped to strengthen the password associations for the longer, week-delay retention interval (see Vu, Garcia et al., 2007 for a discussion).

4.3. Security of generated passwords

Participants in the PPC restrictions alone group generated passwords that were more susceptible to being cracked by a commercial cracking program as compared to passwords that were generated by participants in either the image-based mnemonic group or the text-based mnemonic group. However, even with the brute force commercial password recovery product (PRTK) that is designed to find fairly complex passwords, only 2.05% ($n = 8$) of all the generated passwords were recovered (cracked). All of the cracked passwords consisted of a similar structure (e.g., a single word that had the first letter capitalized and the rest lower case, followed by a digit and then a special character, such as "Venice1!" or "Computer#1"), and all were found among those passwords generated by participants in the PPC group.

The PRTK did not crack some of the passwords that were thought to be more simplistic in nature, such as "Yahoo1!!", "Fri3nds!", or "Myspace123!". Thus, a simple test was run to measure the viability of the PRTK as a password cracking tool. Eight variations of the password "password" were submitted to the PRTK to determine cracking rates (see Table 9). All variations were cracked within 913 s (Range = 3–913 s). Thus, the PRTK appears to be a viable tool to use for password recovery. Therefore, the lack of passwords cracked in the present study is likely due to the fact that they were both complex and long, averaging over 10 characters in length.

Proctor et al. (2002) found that passwords greater than eight characters are less likely to be cracked by computer programs within the 12-h cracking window. Thus, results of this study provide further support to previous research indicating that proactive password checking restrictions alone suffice to provide participants with the tools necessary to generate relatively secure passwords (Vu, Proctor et al., 2007). It was estimated that it would have taken approximately 84 h to run a single password through all 161 levels of the PRTK. Due to the time constraints involved with the current study, the PRTK was limited to a 12 h cracking window. Therefore, the possibility exists that more of the submitted passwords would have been cracked if the program had run its full length.

Another reason why these passwords were more crack-resistant is that participants in this study may be more sophisticated in their computer security practices than those in previous studies. Evidence of this sophistication comes from information provided by participants in a post-recall questionnaire that asked questions pertaining to their password generation and management practices. The results of the questionnaire were compared with findings from two password studies described in the Introduction by Morris and Thompson (1979) and Florêncio and Herley (2007), who analyzed actual passwords from users. Participants in the current study are reportedly much more sophisticated in their password creation practices than were their counterparts in either of the previous studies. Unlike their counterparts, nearly all of the participants in this study indicated that they use numbers in their passwords (95%), slightly over half reported using both upper and lower case letters (56%), and roughly a third of the participants reported that they use special characters in their passwords (32%).

4.4. Forgetting of passwords

Participants using the text-based mnemonic technique were nearly twice as likely to forget the passwords they generated as compared to the participants who used the image-based mnemonic technique; however, they did not forget significantly more passwords than did the participants who used the PPC restrictions alone. Again, the superiority of memory for pictures over words (Paivio et al., 1968) could account for these findings. Interestingly, 62% ($n = 81$) of the passwords generated by participants using the PPC restrictions alone were clearly associated with the accounts for which they were generated (e.g., "4Friend$", "!1Bank1!", "#1Potterfan"), whereas only 30% ($n = 39$) of the passwords generated by participants in the text-based group (e.g., "Eye<32Read", "Gr1ng0tts#B@nk", "Nb0x!sfull") and 19% ($n = 25$) of the passwords generated by participants in the image-based group (e.g., "ABC&W3RDz", "=]M00Lah", "3EEEmale???") could be said to be associated with the accounts for which they were generated. Because the PPC group was not given an organized structure to relate their password elements, they may have decided to relate the password contents to the account to aid memory. However, this strategy may not be wise since these types of passwords can be more easily guessed.

Table 9
Cracking rates for test passwords submitted to the PRTK.

Password	Cracking rate (s)
password	3
Password	21
password1	24
Password1	20
p@ssword	563
P@ssword	592
p@ssword1	3
P@ssword1	913

Moreover, a few participants within the PPC restrictions alone group were naturally inclined to employ mnemonic techniques when generating their passwords. Specifically, three participants reported using mnemonic strategies when creating all five of their passwords (e.g., using the parts of a deer's antlers as a visual cue to each password, or using parts of a created phrase that was associated with each account), and one used a mnemonic strategy (parts of an associated phrase) when creating four out of five of their passwords. Additionally, six other participants appear to have used a mnemonic strategy when creating one out of their five passwords (e.g., "Love2read@2", "Pigbank1$", "R33l&w1d"). None of these 25 generated passwords were forgotten during the course of the experiment, the participants who generated them recalled them on the first try each time they were presented. Indeed, results indicate that forgetting for all groups was likely due to interference from prior passwords. Participants across all three generation techniques misattributed a correct password to the wrong account 20% of the time when passwords were inaccurately entered, meaning that they correctly entered a password they had generated, but entered it for the wrong account.

4.5. Implications for password practices and research

The results of the current study provide further support for previous research on proactive password checking restrictions (Bishop & Klein, 1995; Proctor et al., 2002; Vu, Proctor et al., 2007; Vu et al., 2003) and point to a number of considerations for password practices. Previous studies concluded that there was a significant increase in password "strength" (lower cracking rates) gained through the implementation of proactive password checking restrictions during password generation. However, the increase in password security achieved through the use of such restrictions comes at a steep cost to memory. The required use of both upper and lower case characters, digits, and special characters makes it harder for the user to accurately remember their password. Research has shown the use of mnemonic techniques, such as the first-letter or phrase-based mnemonic techniques improve users' memory for the passwords they generate (Vu, Proctor et al., 2007; Vu et al., 2003; Yan et al., 2005). Results from the current study indicated that the use of an image-based mnemonic technique improved participants' memory for their generated passwords beyond that provided by the use of a text-based mnemonic technique. In terms of theory, the present study showed that providing meaning to unfamiliar elements, such as special characters and digits used in passwords, through imposing structure and organizational schemes may not always be sufficient for improving memory (but see Ausubel, 1960). Thus, mnemonic techniques are most effective when the to-be-remembered items are associated with existing knowledge through deep processing, the use of an organizational structure, and imagery (Snowman et al., 2008; Roediger, 1980). Future research should further explore the benefits of imagery in promoting the generation of secure passwords.

5. Conclusion

When provided with a standard set of governing rules for password generation, such as the proactive password checking restrictions, participants were able to quickly and easily generate and then later recall passwords that were mostly resistant to being cracked by the PRTK. Furthermore, when either of the two mnemonic strategies were added to the PPC restrictions, participants generated longer, more complex, and crack-resistant passwords as compared to those participants who only used the PPC restrictions. The image-based mnemonic password methodology, though, was shown to be better than the text-based one. The benefit of the image-based technique can be attributed to the remarkable ability people have for recalling images (Shepard, 1967; Standing et al., 1970). In the current study, nearly 96% of the participants reported that they do not change their password if they are not required to do so. As a result, generation of a crack-resistant password through the image-based mnemonic technique may be beneficial to many online users.

Acknowledgements

This manuscript was supported in part by a CSULB Scholarly and Creative Activities Award to K. Vu. Portions of this paper was presented at the *2009 Human Factors and Ergonomics Society Annual Meeting* and appear in its proceedings. We thank Thomas Strybel, Dan Chiappe, and an anonymous reviewer for helpful comments on prior versions of this article.

References

Armstrong, I. (2003). Passwords exposed: Users are the weakest link. *SC magazine*. Available from: http://www.scmagazine.com/asia/news/article/419622/passwords-exposed-users-weakest-link/ Retrieved April 22, 2008.

Ausubel, D. P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology, 51*, 267–272.

Baddeley, A. D. (2004). *Your memory: A user's guide*. New York: Firefly Books.

Bellezza, F. S. (1981). Mnemonic devices: Classification, characteristics, and criteria. *Review of Educational Research, 51*, 247–275.

Bishop, M., & Klein, D. V. (1995). Improving system security via proactive password checking. *Computers & Security, 14*, 233–249.

Blonder, G. E. (1996). Graphical passwords. United States Patent 5559961.

Brostoff, S., & Sasse, M. A. (2001). Safe and sound: A safety-critical approach to security. In *Proceedings of the 2001 workshop on new security paradigms* (pp. 41–50), September 10–13, 2001. Cloudcroft, New Mexico.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671–684.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*, 268–294.

De Angeli, A., Coventry, L., Johnson, G., & Renaud, K. (2005). Is a picture really worth a thousand words? Exploring the feasibility of graphical authentication systems. *International Journal of Human–Computer Studies, 63*, 128–152.

deWinstanley, P. A., & Bjork, E. L. (2004). Processing strategies and the generation effect: Implications for making a better reader. *Memory and Cognition, 32*, 945–955.

Dhamija, R., & Perrig, A. (2000). Déjà vu: A user study using images for authentication. In *Proceedings of the ninth USENIX security symposium* (p. 4), 2000. Denver, Colorado.

Florêncio, D., & Herley, C. (2007). A large scale study of web password habits. *WWW 2007*, May 8–12, 2007. Banff, BC.

Greenwald, A., & Banaji, M. R. (1989). The self as a memory system: Powerful, but ordinary. *Journal of Personality and Social Psychology, 57*, 41–54.

Grimes, R. A. (2006). *MySpace password exploit: Crunching the numbers (and letters)*. Available from: http://www.InfoWorld.com/article/06/11/17/47OPsecadvise_1.html Retrieved November 30, 2007.

Groninger, L. D., & Groninger, L. K. (1988). Autobiographical episodes as mediators in the recall of words. *The American Journal of Psychology, 101*, 515–538.

Internet Crime Complaint Center (2009). 2008 Internet Crime Report. Available from: http://www.ic3.gov/media/annualreport/2008_IC3Report.pdf Retrieved January 14, 2010.

Ives, B., Walsh, K. R., & Schneider, H. (2004). The domino effect of password reuse. *Communications of the ACM, 47*(4), 75–78.

Klein, D. V. (1990). Foiling the cracker: A survey of, and improvements to, password security. In *Proceedings of the second usenix workshop on security* (pp. 5–14), August, 1990. Portland, OR, USA.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

Morris, R., & Thompson, K. (1979). Password security: A case history. *Communications of the ACM, 22*, 594–597.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart & Winston.

Paivio, A., Rogers, T. B., & Smythe, P. C. (1968). Why are pictures easier to recall than words? *Psychonomic Science, 11*(4), 137–138.

Proctor, R. W., Lien, M.-C., Vu, K.-P. L., Schultz, E. E., & Salvendy, G. (2002). Improving computer security for authentication of users: Influence of proactive password restrictions. *Journal of Behavior Research Methods, Instruments, and Computers, 34*, 163–169.

Riddle, B. L., Miron, M. S., & Semo, J. A. (1989). Passwords in use in a university timesharing environment. *Computers and Security, 8*, 569–578.

Roediger, H. L. (1980). The effectiveness of four mnemonics in ordering recall. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 558–567.

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology, 35*, 677–688.

SafeNet (2005). 2004 annual password survey results. Available from: www.safenet-inc.com. Downloaded on May 21, 2006.

Schneier, B. (2000). *Secrets and lies*. New York: John Wiley.

Schultz, E. E. (2005). Web security and privacy. In R. W. Proctor & K.-P. L. Vu (Eds.), *Handbook of human factors in web design* (pp. 613–628). Mahwah, NJ: Erlbaum.

Shepard, R. N. (1967). Recognition memory for words, sentences and pictures. *Journal of Verbal Learning and Verbal Behavior, 6*, 156–163.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 592–604.

Snowman, J., McCown, R., & Biehler, R. (2008). *Psychology applied to teaching* (12th ed.). Belmont, CA: Wadsworth Publishing.

Squire, L. R., & Zola-Morgan, S. (1998). Episodic memory, semantic memory, and amnesia. *Hippocampus, 8*, 205–211.

Standing, L., Conezio, J., & Haber, R. N. (1970). Perception and memory for pictures: Single trial learning of 2500 stimuli. *Psychonomic Science, 19*, 73–74.

Stubblefield, A., & Simon, D. R. (2004). *Inkblot Authentication* (Technical Report MSR-TR-2004-85). Microsoft Corporation, Redmond, WA, USA.

Tullis, T. S., & Tedesco, D. P. (2005). Using personal photos as pictorial passwords. *CHI 2005*, April 2–7, 2005. Portland, OR, USA.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). New York: Academic Press.

Vu, K.-P. L., Bhargav, A., & Proctor, R. W. (2003). Imposing password restrictions for multiple accounts: Impact on generation and recall of passwords. In *Proceedings of the 47th annual meeting of the human factors and ergonomics society* (pp. 1331–1335). Santa Monica, CA: HFES.

Vu, K.-P. L., Cook, J., Bhargav, A., & Proctor, R. W. (2006, April). Short-term and long-term retention of passwords generated by first-letter and entire-word mnemonic methods. In *Proceedings of the fifth annual security conference*. Published on CD-ROM.

Vu, K.-P. L., Garcia, F., Nelson, D., Sulatis, J., Creekmur, B., & Chambers, V. (2007). Examining user privacy policies while shopping online: What are users looking for? In *Proceedings of the 12th international conference on human–computer interaction* (pp. 792–801), July 22–27, 2007. Beijing, China.

Vu, K.-P. L., Proctor, R. W., Bhargav-Spantzel, A., Tai, B.-L., Cook, J., & Schultz, E. E. (2007). Improving password security and memorability to protect personal and organizational information. *International Journal of Human–Computer Studies, 65*, 744–757.

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*, 102–127.

Yan, J., Blackwell, A., Anderson, R., & Grant, A. (2005). Password memorability and security: Empirical results. *IEEE Security & Privacy, 2*(5), 25–31..