View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by NSU Works

Nova Southeastern University
NSUWorks

CEC Theses and Dissertations

College of Engineering and Computing

2015

The Impact of Image Synonyms in Graphical-Based Authentication Systems

Jonathan William Sparks
Nova Southeastern University, js3063@nova.edu

This document is a product of extensive research conducted at the Nova Southeastern University College of Engineering and Computing. For more information on research and degree programs at the NSU College of Engineering and Computing, please click here.

Follow this and additional works at: http://nsuworks.nova.edu/gscis_etd


Part of the Information Security Commons

Share Feedback About This Item

NSUWorks Citation

Jonathan William Sparks. 2015. *The Impact of Image Synonyms in Graphical-Based Authentication Systems.* Doctoral dissertation. Nova Southeastern University. Retrieved from NSUWorks, Graduate School of Computer and Information Sciences. (33)
http://nsuworks.nova.edu/gscis_etd/33.

This Dissertation is brought to you by the College of Engineering and Computing at NSUWorks. It has been accepted for inclusion in CEC Theses and Dissertations by an authorized administrator of NSUWorks. For more information, please contact nsuworks@nova.edu.

The Impact of Image Synonyms in Graphical-Based Authentication Systems

by

Jonathan W. Sparks

A dissertation report submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

in

Computer Information Systems

Graduate School of Computer and Information Science

Nova Southeastern University

2015

We hereby certify that this dissertation, submitted by Jonathan Sparks, conforms to acceptable standards and is fully adequate in scope and quality to fulfill the dissertation requirements for the degree of Doctor of Philosophy.

_____ _____

Maxine S. Cohen, Ph.D. Date
Chairperson of Dissertation Committee

_____ _____

James D. Cannady, Ph.D. Date
Dissertation Committee Member

_____ _____

Timothy J. Ellis, Ph.D. Date
Dissertation Committee Member

Approved:

_____ _____

Eric S. Ackerman, Ph.D. Date
Dean, Graduate School of Computer and Information Sciences

SIGNED ORIGINAL ON FILE AT NSU

Graduate School of Computer and Information Sciences
Nova Southeastern University

2015

An abstract of a Dissertation Report Submitted to Nova Southeastern University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

The Impact of Image Synonyms in Graphical-Based Authentication Systems

by
Jonathan W. Sparks
March 2015

Traditional text-based passwords used for authentication in information systems have several known issues in the areas of usability and security. Research has shown that when users generate passwords for systems, they tend to create passwords that are subject to compromise more so than those created randomly by the computer. Research has also shown that users have difficulty remembering highly secure, randomly created, text-based passwords.

Graphical-based passwords have been shown to be highly memorable for users when applied to system authentication. However, graphical-based authentication systems require additional cognitive load to recognize and enter a password compared to traditional text-based authentication that is more muscle-memory. This increase in cognitive load causes an increased security risk of shoulder-surfing created from the longer amount of time needed to input a password.

Graphical-based authentication systems use the same images for each possible input value. This makes these authentication systems vulnerable to attackers. The attackers use their ability to remember visual information to compromise a graphical-based password.

This study conducted research into a graphical-based authentication scheme that implemented pictorial synonyms. The goal is to decrease security risk of graphical-based authentication systems while maintaining (or even increasing) the usability of these systems. To accomplish this goal, a study to evaluate the impact on the cognitive load required using an image synonym authentication system compared to traditional graphical-based authentication schemes.

The research found that there was not a significant difference in the areas of user cognitive load, shoulder-surfing threat, and user effectiveness. The research evaluated users' accuracy, cognitive load, and time to authenticate and found to have significant impact of pictorial synonyms on graphical-based authentication systems. The research

shows that the accuracy of pictorial synonyms was greater than word password. This appears to due to people's ability to recall pictorial information over text information. Future research should look at the impact of pictorial synonyms on shoulder-surfing attackers and different ages.

Acknowledgements

I would also like to thank my wife Janae for constantly reminding me of the truth from 1 Thessalonians 5:24, “Faithful *is* he that calleth you, who also will do *it*.”

I would like to thank Dr. Cohen for all of her help throughout the entire process of my degree. I could not have made it this far without her help.

I would also like to Dr. Ellis, and Dr. Canady in helping to guide and improve my dissertation and take it to the next level.

Table of Contents

List of Tables

Tables

List of Figures

Figures

Chapter 1

Introduction

Background

A primary means of security implemented on computer systems has been the use of a text string of characters to prevent unauthorized access to data or processes. The method was a simple solution to the security problems of the emerging computer information systems at its onset of use. This method of authorization had the potential to be sufficiently secure. Text-based authentication has many different character symbols that could possibly be used. However, with the many possible alpha-numeric and symbol character combinations, brute force attacks became very difficult. The one major problem to this is the human limitation of remembering passwords makes the implementation of this security vulnerable to attacks (Pilar, Jaeger, Gomes, & Stein, 2012).

To decrease the amount of information needed to access data, users often reuse authentication credentials across multiple accounts. This practice can make the security of one computer information system dependent on the security implemented on another information system (Bang, Lee, Bae, & Ahn, 2012). This security risk is shown in the case of a recent attack by two Korean hackers. The hackers compromised several small information systems and used the gained user account names and password combinations

to compromise accounts on a major portal site that was considered to have strong security (Ives, Walsh, & Schneider, 2004).

One approach that has been proposed to address the inherent security vulnerabilities of text-based authorization methods was the use of a graphical password system (Wright, Patrick, & Biddle, 2012). Graphical systems rely on the human ability to remember visual information better than their ability to remember a series of characters (Biddle, Chiasson, & Van Oorschot, 2012; Gehring, Toglia, & Kimble, 1976). Such a system of authentication could allow users to maintain multiple secure passwords with minimal memory effort, while also reducing security risks for information systems.

Gehring et al. (1976) used qualitative measures to evaluate the short and long term memory requirements of words compared to pictures. Gehring et al. found that pictorial information was significantly easier to recall in short and long term memory. Using this knowledge, a graphical-based password could reduce the memory load required in an authentication system.

Tullis, Tedesco, and McCaffrey (2011) reported that when using a graphical-based recognition password ,users were able to successfully remember the target images through a six-year separation from the last use of the password. This indicates that visual recognition authentication could improve the overall recollection performance in users, while maintaining high security protocols. The implementation of graphical recognition could hold the key to creating effective secure authentication systems in the future.

In implementing graphical-based authentication, many potential security problems arise. Wright et al. (2012) noticed a significant difference in the time needed to log in using the recognition authentication scheme. This introduced a security risk in allowing

attackers a longer window of opportunity to observe user passwords. Wright et al. (2012) suggest that this is due to the lack of familiarity with the placement of the recognition-based passwords.

In many graphical-based authentication systems, positions of possible response choices are randomized. This requires users to seek through the given selectors and identify the correct selector. This type of system requires users to rely solely on recognition of the selectors and not on the recollection of positions of the locations of those selectors on an input device. Graphical-based authentication systems still have security vulnerability because the authentication system uses the same images in a randomized position. Attackers use the same ability to remember visual cues as legitimate users to watch and compromise a password.

Reaction time is generally faster with text-based passwords because keyboard layout is generally in the QWERTY layout, which users are familiar with. After typing a password several times on a standard keyboard, a user's password entry becomes muscle memory requiring little seeking and recollection for password entry (Wright et al., 2012). The quick reaction time and accuracy of entering a password provides little time for an attacker to successfully observe and remember a password.

Problem Statement

There is an important security issue with text-based authentication that needs to be addressed in research. Authentication schemes need to fulfill security and usability concerns to adequately protect computer information systems (Bang et al., 2012; Biddle et al., 2012; Pilar et al., 2012; Wright et al., 2012).

Removing the user's ability to create his or her own passwords has been shown to increase security on information systems; however, the difficulty in the user's ability to retain randomly generated authentication has been a limiting factor in the success of these systems (Bang et al., 2012; Wright et al., 2012). Computer information systems research is needed to address the problem of user memory limitation of randomly generated passwords that are found in current authentication methods with new methods that are secure and usable.

Dissertation Goal

The goal of this study was to examine the ability of graphical-based passwords to address the security limitation of users to remember randomly-assigned passwords found in traditional authentication systems. To accomplish this goal, a graphical-based password system that randomly generates graphical-based password with pictorial synonyms was evaluated and compared to other known graphical-based authentication systems.

A pictorial synonym is an non-identical image of the same subject matter as another image (Gehring et al., 1976). The use of pictorial synonym images in the proposed authentication system was designed to address the security issue of potential observation attacks.

Research Questions and Hypothesis

Graphical-based passwords have shown promise in creating strong, memorable passwords for users of information systems. There are several research questions that this research attempted to answer to better understand the viability of graphical-based passwords as an effective solution to information system's security needs.

RQ1. What is the impact of a randomly generated graphical-based password with pictorial synonyms on cognitive load?

Georgakakis, Komninos, and Douligeris (2012) found in the evaluation of a graphical password scheme (NAVI) that having multiple authentication questions which could reduce the impact shoulder-surfing attacks have on a computer information system. This premise could be modified to implement a recognition-based password scheme where the answers are random pictorial synonyms. This should make attacks on the authorization by observation more difficult because the attacker would not know why a selector was chosen immediately.

Additionally, users could have several possible images that are synonyms for the same idea, pictorial synonyms, instead of having multiple authentication questions. Gehring et al. (1976) evaluated the memory capability of subjects for both pictures and words in both long and short intervals. Gehring et al. found that there is an equal level of comparison between word synonyms and pictorial synonyms among people. People are able to easily differentiate between images that are of the same subject matter but in different contexts. Gati and Tversky (1987) studied people's ability to find missing components in pictorial stimuli and found that people are able to find missing components better in pictorial stimuli that is of the same subject matter compared to stimuli of distinctive subject matters.

RQ2. How would the security threat of shoulder-surfing be affected by implementing a graphical-based authentication system that uses pictorial synonyms?

The human ability to differentiate between similar images shows that pictorial synonyms have the potential to be used in graphical-based passwords, but that alone would not be a good solution for shoulder-surfing because of the universal nature of pictorial synonyms. The addition of pictorial synonyms could be enough to help reduce the success rate of shoulder-surfing attackers.

One of the major security concerns is shoulder-surfing (Chang, Tsai, & Lin, 2012)**.** Research is needed on the use of pictorial synonym in a graphical-based authentication system, and the impact that addition has on shoulder-surfing. The threat of shoulder-surfing could possibly be mitigated or reduced by implementing a graphic-based password using several similar visual representations for the same authentication response input choice. The use of pictorial synonyms would increase the amount of information that an attacker would need in order to compromise a system. This system has the potential to reduce an attacker's ability to easily observe a graphical-based password and compromise the security of a system.

RQ3. What are the implications to users' effectiveness using a graphical-based authentication system using pictorial synonyms?

Wright et al. (2012) noticed that recognition-based password systems could be just as effective as recall-based systems; however, it was shown that recognition-based systems required users to take more time to enter the password. To consider the feasibility of implementing a new authentication system, the memorability and usability of the systems' password implementations needs to be comparable to current authentication systems.

Based on the research questions and review of literature the following null hypotheses are formed:

H1: There is no significant difference on a user's cognitive load of a randomly generated graphical-based password with pictorial synonyms compared to passwords without pictorial synonyms.

H2: There is no significant difference to the security threat of shoulder-surfing by implementing a graphical-based authentication system that uses pictorial synonyms compared to a system without pictorial synonyms.

H3: There is no significant difference to user's effectiveness using a graphical-based authentication system using pictorial synonyms compared to graphical-based systems without pictorial synonyms.

This research used user performance time and accuracy to evaluate H1, H2 and H3. In addition, the NASA-TLX was used to also evaluate H1. Chapter 3 of this study further discusses the methodology used.

Relevance and Significance

Graphical-based password authentication systems have been proposed and studied to strengthen the security of information systems (Wright et al., 2012). These graphical-based authentication systems have not adequately resolved all security issues and have introduced a potentially larger vulnerability to shoulder-surfing type attacks (Biddle et al., 2012). Current research needs to address the shoulder-surfing threat found in graphical-based authentication systems for these systems to be considered a viable solution to authentication in information systems.

De Angeli, Coventry, Johnson, and Renaud (2005) discuss the potential security risk of shoulder surfing with graphical-based passwords in their research on the feasibility of graphical authentication systems. Shoulder surfing attacks are more successful compromising a system when users require more time to enter a password. The extra time allows the attacker more time to observe and remember passwords as they are entered. The attacker success rate could be diminished with the implementation of graphical-based passwords with pictorial synonyms.

There have been several proposed and implemented graphical-based authentication systems; however, there is limited research into the use of pictorial synonyms in these systems. Hunt and Elliot (1980) noted that the distinctiveness of an image impacts the memorability of that image. Studies on graphical-based authentication have focused on the viability of these systems and are only now beginning to look at improving these systems though modification to existing systems (Biddle et al., 2012).

If the proposed graphical-based system is shown to lessen the user's cognitive load, then the password would be retained by the users, and the users would be able to input the password with a minimal amount of time. The impact level of usability using pictorial synonyms could result in the reduction of the likelihood of a successful shoulder-surfing attack on an information system through its authentication scheme.

Barriers/Issues

The measure of password retention is a barrier to this study. The amount of repetition of a task increases the retention of that task. The retention levels may vary depending on the participant's habits in using an authentication system.

In a comparative study, the demographics population of the study groups can affect the conclusions that can be made. If one group skews to one demographic, that could impact the integrity of the research being conducted. To address this problem, groups can be matched to have similar demographics. The makeup of the groups must be similar to create a representative population of users. The quantitative data can then be compared without questioning the validity of the comparisons. To address the potential issue in this study, demographic information was gathered before the start of the study to ensure the homogeny of the groups. This information was used to help balance the test groups into a similar representative user configuration before the data collection process began.

Finally, the users' satisfaction towards the environment where the research is being conducted can directly impact their interaction and effectiveness in using an authentication system. One issue that was monitored was the subject's feelings towards the system and look for balance in overall feelings about the system across the different groups. This can be mitigated by collecting the overall satisfaction about the system before and after the study period to verify a balance among the test groups.

Assumptions, Limitations, and Delimitations

This study was limited by the conclusions that could be made on password retention by using a six-week period to observe participants using this type of authentication system. The conclusions drawn by this research were able to state findings within the given time frame. The study could generate implications towards the use of the system over a longer period of time; however, further study would be required to verify those conclusions. The time spans that have been used to test memory recall varied from

minutes, weeks, months, and years. For example, Chowdhury, Poet, and Mackenzie (2014) used a 14 day span between their participants receiving a graphical-based password and using that password to test the memorability of that password. This research used a six week timeframe which provided ample time to measure retention and gather good usability data. Chapter three of this study elaborates the methodology of the time span used in this study.

The measure of cognitive load is another limitation of this study. Bang et al. (2012) discuss the measure of cognitive load as an issue in the research and evaluation of graphical-based authentication. Studies on graphical-based systems have not had consistency in the measurement of cognitive load, which makes reproducing and comparing results difficult. One studied and verified method of workload evaluation is the NASA-TLX (Hart & Staveland, 1988). NASA-TLX is a subjective evaluation of workload and as a result findings have the possibility of not accurately reflecting true cognitive load. Other non-subjective measurements like time required to log in and number of failed log in attempts was also taken to allow for a more complete picture of cognitive load. The use of these measurements helped to verify the results from the NASA-TLX and give a more complete idea of the cognitive load required for the use of the authentication systems being studied.

Definition of Terms

Authentication – the identification of a user to access computer system's resources or information (Wright et al., 2012).

Account locking – the practice of preventing access to an account after a number of unsuccessful authentication attempts (Kirushnaamoni, 2013).

Biometric authentication – user behavioral or physical characteristics are used for authentication (Jain, Ross, & Pankanti, 2006).

CAPTCHA – Completely Automated Public Turing Test to tell Computers and Humans Apart is an automatic system used to tell if human or a computer is attempting authentication (Kirushnaamoni, 2013).

Cognitive load – a measure for the effort needed by a user to observe and identify visual stimuli (Back & Oppenheim, 2001).

Cognometric Systems – an authentication system in which the user is given several images and must select the correct images from the distractor images (De Angeli et al., 2005).

Denial of Service Attack – an attack in which an attack attempts to guess passwords to lock accounts and prevent legitimate users from authenticating (Kirushnaamoni, 2013).

Graphical password – a secret involving an image or many images used to verify identity on a system (Biddle et al., 2012).

NASA-TLX – the National Aeronautics and Space Administration – Task Load Index is a tool developed to measure subjective workload (Hart & Staveland, 1988).

Recall-based password – a password based on remembering information given no additional aids (Wright et al., 2012).

Recognition-based password – a password based on remembering information given a cue of possible choices (Wright et al., 2012).

Pictorial synonym – an non identical image of the same subject matter as another image (Gehring et al., 1976).

Shoulder-surfing – an attack where the attacker is able to observe or record a user's password over their shoulder (Forget, Chiasson, & Biddle, 2010).

Text password – a secret text used to verify identity on a system (Wright et al., 2012).

Summary

An increasing number of common tasks require the use of computer information systems. Securing these systems is done through the use of authentication. It is pertinent for all stakeholders that authentication systems are secure and limit the vulnerability of the system.

Password--based authentication systems are commonly used; however, the usability of these systems decrease as the security requirements of the passwords used increase. One proposed method of increasing usability and security of password-based authentication systems is the use of graphical-based passwords. These systems are often more susceptible to the problem of shoulder-surfing. Research is needed to look at ways of combating shoulder-surfing in graphical-based passwords while maintaining both security and usability. This study suggests the use of pictorial synonyms with graphical-based authentication as a solution to combat shoulder-surfing attacks.

Chapter 2

Literature Review

Introduction

This chapter reviews the following key areas of graphical-based authentication systems: text-based passwords, cognometric systems, graphical authentication systems, methods for evaluating passwords, and cognitive load measurement. This literature review was designed to inspect the areas of text-based authentication, security issues of text-based authentication, other authentication methods, graphical authentication, and authentication evaluation to show the significance of further study into the impact of pictorial synonyms in graphical-based authentication systems.

Text-based Passwords

A recurring issue with text-based passwords is that users commonly choose weak passwords when given the opportunity. Despite the massive range of passwords that could be chosen, users consciously and subconsciously restrict the range of possible passwords (Biddle et al., 2012; Georgakakis et al., 2012).

Morris and Thompson (1979) evaluated text-based password security by looking at over 3000 passwords on a system that allowed users to generate a password with no security restrictions. Morris and Thompson found that over 86% of users had simple passwords of less than six characters or a simple dictionary name.

Several methods have attempted to mitigate the risk by requiring users to meet various password policies when they create a password for an information system. Recent research has verified that, despite increased password security requirements implemented by organizations, users still engage in high-risk behaviors such as reusing passwords across many systems, sharing passwords with others, and writing passwords down and leaving those passwords in unsecured locations (Grawemeyer & Johnson, 2011).

Shay et al. (2014) studied the usability of long passwords with many different security requirements. Given strict requirements, users would often employ tricks that met the security requirement while still being simple to remember. For example, 28% of participants fulfilled the symbol requirement by adding "!" at the end of their password string. Shay et al. also discovered that 54.4% of participants fulfilled the uppercase letter requirement by capitalizing the first letter of the password and no other letters. These common patterns show that users tend to choose passwords that are vulnerable to possible attacks.

The problem of security compliance with password selection could be completely eliminated by removing users' ability to choose a password in a system. This would allow an information system to create secure passwords without natural prejudices that the users have in generating a password. Users find password retention to be difficult when a computer generates a secure password (Wright et al., 2012). Lack of password retention causes many potential problems within an information system. Users either have to reset a password or participate in risky security behaviors to prevent disruption of service. Both of these scenarios introduce security vulnerabilities into a system. Users

form habits in their implementation and usage of text-based passwords that have created an atmosphere full of potential security risks.

Bang et al. (2012) showed that across multiple websites users tend to have similar usernames and passwords. This tendency shows user's inability to self-generate secure authorization usernames and passwords on systems. Bang et al. found that users interact with an average of over 100 websites that require authentication accounts with a minimum of 27 websites and a maximum of 199 websites. The users had on average approximately seven unique user names and five unique passwords for these websites with a minimum of two unique user names and one unique password and a maximum of 14 unique user names and 15 unique passwords. When given a choice, users will create passwords that are similar to other known passwords. This practice has less impact on the user's cognitive load in memory and entry of the password.

One way many users are attempting to have more secure passwords and are able to manage them is the introduction of password-management software to their system access workflow. Password managers allow users to store passwords for the many systems that the user interacts with. The manager allows the user to no longer have the requirement of remembering all of their passwords. Instead the users only needs to remember the password for the manager and then they have access to all their passwords. Chiasson, van Oorschot, and Biddle (2006) performed an analysis of two different password managers to evaluate the usability and user's acceptance of these solutions. This study found participants had a less than 50% success rate completing basic tasks using these password management systems. Users then were susceptible to attacks through the misuse of the password management systems. Research is needed into

finding a solution that better addresses the memory limitations of users and security concerns of information systems. Traditional text-based passwords may not hold the key to solving this problem.

Security Issues of Passwords

Authentication systems introduced complicated rule sets and password expiration periods to combat these potential attacks. These complicated rule sets require users to have a minimum amount of characters of varying types with limitations on certain character ordering. The combination of rule sets and expiration dates have introduced the problem of increased cognitive load on users. Many users find remembering expiring passwords with completed rule sets unusable and tend to begin to engage in highly insecure actions. Adams and Sasse (1999) surveyed users and found that more complicated password requirements caused a reaction of users to start writing their password on post-it notes by their desk. These high-risk behaviors are not an outcome of a successful authentication system. Flor, Herley, and Coskun (2007) conducted a study on the effectiveness of strong passwords against a variety of potential attacks and found that in a bulk-attack on a system the strength of the password did not statistically protect a system as a whole. Flor et al. instead recommended focusing on the prevention of password stealing and increasing the user id space to reduce the likelihood of a successful bulk-attack on a system.

Text-based passwords are also susceptible to a guessing attack. A guessing attack is where the attacker uses a dictionary of likely passwords and attempts to access accounts with those likely passwords. These attacks are more successful in authentication systems with a smaller number of possible passwords or systems that have identifiable

patterns in the user choice of passwords (Biddle et al., 2012). A dictionary attack is a type of guessing attack that uses a dictionary of common words and attempts to use those words and combinations of those words. Morris and Thompson (1979) conducted research that attempted to attack passwords using a simple dictionary attack and found the 75% of accounts could be compromised using this attack technique.

There have been many attempts to solve the problem of guessing attacks; however, these attempts still leave certain areas open for attack. Account locking is one researched area suggested to prevent guessing attacks. A defined number of unsuccessful authentication attempts on an account will block access for a set amount of time. Kirushnaamoni (2013) attempted to solve this issue by the development of a Password Guessing Resistant Protocol. This protocol limited the number of login attempts for an unknown users to one attempt. If the user fails to authenticate the protocol challenges the user with an Automated Turing Test like a CAPTCHA (Completely Automated Public Turing Test to tell Computers and Humans Apart) security questions. The CAPTCHA is designed to tell if the user attempting authentication is human or a computer. However, the account locking schemes are still subject to potential Denial of Service attacks in which an attacker or a group of attackers will guess passwords to lock accounts and stop legitimate users from accessing their systems.

Cognometric Systems

Text-based passwords work mainly on the premise of recalling information. Most text-based authentication is purely recall-based. This requires users to remember a password with no cue or additional remembering aid. However, text-based passwords can use cognometric systems as a recognition-based password scheme.

Recognition-based passwords prompt users with a valid choice and several distracters to cue memory of a password. Wright et al. (2012) tested text-based passwords on letter recollection, word recollection, and word recognition. The study showed that use of a cognometric system could improve the usability of authentication. Wright et al. (2012) found that cognometric systems, created fewer password reset requests compared to recall-based passwords. The occurrences of users participating in high-risk behaviors were reduced when the system for authenticating is more usable.

Gehring et al. (1976) noted that people have varying success rates of retaining pictorial information in both short-term and long-term memory as compared to word information. They found that the loss of pictorial information only marginally lessened after a three month span of time without any formal form of measurement.

The design of cognometric systems can impact the memorability of visual components (De Angeli et al., 2005). In the study of several implementations of graphical-based authentication systems, De Angeli et al. found that a poorly designed system would eliminate the memorability of images used for authentication. However, De Angeli et al. also showed the feasibly of graphical-based authentication as a viable replacement for traditional text-based password schemes.

Tullis et al. (2011) were able to show a high success rate of users being able to remember a pictorial password that was selected six years prior during a previous study. Tullis et al. used handpicked distracters to try to make the image recognition more challenging for users; however, with the small sample group there was a 95.6% accuracy rate. This study gives evidence of the memory retention benefits to graphical passwords used in cognometric systems.

Other Authentication Methods

There have been several viable solutions for addressing security issues found in traditional text-based passwords. The simplest of these solutions is the implementation of a password manager. Password managers are software solutions for the generation and retrieval of authentication information for users. The password manager allows the user to need to remember only one secure password and all their other account information is locked by the one password. The problem with this solution is similar to the problem found with traditional text-based passwords. All the generated passwords have a high security level; however, attackers need only to acquire the one password and all of the user's accounts are now compromised.

Chiasson et al. (2006) studied user habits when interacting with password managers and found that users did not use these systems correctly in several instances. In addition, Chiasson et al. observed that users did not trust a password manager to be a secure solution for authentication. Password managers introduce a single point of failure to a user's secure information. An attacker only needs to access the password manager to then gain access to all of the users accounts. Chiasson et al. did not find users to trust these managers as a better security solution. Users would rather rely on their own abilities when securing the systems that they use. The usability study that Chiasson et al. conducted on password managers showed that authentication systems still need to look for new solutions to the secure problems that they face.

Another authentication method that has been proposed is the use of a physical token-based authentication. Physical token based authentication is where the user has in their possession a physical device that allows them access to their account. The physical

device will have a rotating code or a signal that will be used to authenticate the user (Bonneau, Herley, van Oorschot, & Stajano, 2012).

This authentication scheme has a few vulnerabilities that have not yet been fully addressed to make this a truly viable solution. The token-based authentication is vulnerable to an attacker by physically acquiring the token from the user. When this occurs, the user has to somehow notify the information system that they are a valid user, and their account is now compromised. In addition, token-based authentication is also vulnerable to an attacker compromising the algorithm used to generate the rotating code. This exact scenario happened in March 2011when attackers were able to compromise the database seeds of the RSA's token-based authentication. Once the attackers had the database seeds for the code generation they could predict the next code for any given token (Bright, 2011).

Jain et al. (2006) conducted a literature overview for the use of biometric authentication as a solution to password security problems. Biometric authentication is where either behavioral or physical characteristics of the user is used to authenticate the user. These types of characteristics include the use of fingerprints, iris scans, and keyboard dynamics. Biometric authentication can eliminate some of the security concerns that passwords have but have many issues associated with them.

One major hurdle that biometric authentication needs to address is the deployability of the technology. Several biometric authentication methods require additional computer hardware to implement the needed evaluation of the user's characteristics. The addition of these extra sensors to many existing systems would be cumbersome and not feasible in many cases.

Biometrics also have the additional problem of not being able to handle the leak of biometric data. A biometric leak is when a user characteristic is captured in a way that it can then be used by others to authenticate. An example of a biometric leak is a theif in Malaysia who cut off a finger to gain access to a car (Kent, 2005). Many biometric authentication system lack a way to assign a user a new means to authenticate once data has been compromised.

These other authentication systems have areas of strength; however, they have yet to fully address the security issues found in passwords. Bonneau et al. (2012) developed a comprehensive framework to evaluate authentication schemes in the areas of usability, deployability, and security. Bonneau et at. found that password managers, hardware tokens, and biometrics all have deficiencies in the areas of usability, deployability, and security. This study shows that there is still a need for research to address the usability and security to improve authentication systems.

Graphical Authentication Systems

Cognometric systems can be inputted with only text and allow users to select the correct password from incorrect distractors (Wright et al., 2012). However, these systems are often used with images creating a graphical authentication system.

In a previous study, Tetsuji, Takehito, and Hideki (2006) looked at the implementation of a graphical-based authentication where the users were allowed to upload a personal photo for use in the authentication system. Tetsuji et al. called their approach Awase-E and compared their system to other known authentication methods. In allowing the personalization of personal photos, users of Awase-E had an even stronger recollection because the photo used was tied to stronger emotional responses.

The Awase-E system was compared to a traditional password and a graphical-based password that assigned an abstract image that had no personal ties to the user. The Awase-E system had a very high memory retention rate as evidenced in the results of successful authentication after 16 weeks. Users of the Awase-E system had a 100% accuracy rate compared to 60% memory retention in those using a traditional password and 50% memory retention in those using an assigned abstract image.

Davis, Monrose, and Reiter (2004) introduced a story method to aid users in the retention of user-chosen images used in a pictorial authentication system. In the story method, users are expected to retain images and a particular sequence to recall those images. Davis et al. showed that users could easily remember their chosen images from the image distracters; however, users had trouble retaining the sequence they originally chose. Davis et al. also discovered that, like with text-based passwords, when users are given a choice to generate a password they tend to have subconscious factors that affect the images that are chosen. These factors can be exploited by attackers attempting to gain access to an information system. This study also showed a clear limitation of recognition-based authentication schemes in that the sequential ordering of image-based passwords did not appear to be tied to image recognition.

Chowdhury et al. (2014) examined the memorability of multiple graphical-based passwords over a two-week time span with the aid of a hint. When creating a password, users chose four images and give a hint for the password created. Each user created four separate passwords from four completely different categories of images. After the two-week span 85% of participants were able to remember all four passwords they generated which showed graphical-based passwords to be memorable over time.

In implementing graphical-based authentication, many potential security problems arise. Wright et al. (2012) saw a significant difference in the time needed to log in using the recognition authentication scheme. This introduced a security risk in allowing attackers a longer window of opportunity to observe user passwords. Wright et al. suggest that the increase in time is due to the lack of familiarity with the placement of the recognition-based password selectors. In many graphical-based authentication systems, positions of possible response selectors are randomized. This requires users to look through the given selectors and identify the correct selector. Reaction time is generally faster with text-based passwords. After typing a password several times on a standard keyboard, a user's password entry becomes muscle memory requiring little seeking and recollection for password entry (Wright et al.).

An additional potential problem image-based logins have to overcome is the perceived additional time it takes to authenticate. Users are resistant to systems that take too much time. Dunphy, Heiner, and Asokan (2010) noted, "User acceptance is often driven by convenience and login duration of approximately 20 seconds are unattractive to many users" (p. 11). With many users having their password committed to muscle memory, a system that cannot reach that perceived speed will feel cumbersome.

De Angeli et al. (2005) pointed out in their evaluation of feasibility of graphical-based authentication systems that many factors have a direct impact on the possible security of a graphical authentication system. These factors are the guessability, observability, and recordability of the passwords. De Angeli et al. observed that there are still many unresolved issues with security when the system's authentication interface

provides an environment that allows attackers to record, guess, or observe the code entry process.

An attack method that graphical password systems are vulnerable to is shoulder-surfing. This is an attack where an observer can watch graphical passwords as they are entered at login and repeat the pattern (Chang et al., 2012). Graphical-based systems are more susceptible to these types of attacks because it is easier for an observer to view an input device on a screen selecting graphical cues than when a user is working with keyboard strokes and the text displayed is masked by symbols. The opportunity for attacks to observe a password increase with the greater amount of time that is used to input a graphical password.

One method proposed to help reduce the threat of shoulder-surfing attacks is the use of eye-gaze entry. Forget et al. (2010) tested using an eye-gaze input entry method to allow users to select a graphical password target without any visual cues. The eye-gaze systems work by using the eye movement and position that determine the target that is being selected. The tolerance level required to implement this method was larger which reduced the overall possible password space possible with this system of input. While this scheme was an improvement to other eye-gaze methods, it still lacked a solid solution to the security problem of shoulder surfing.

DeLuca et al. (2014) proposed a method to reduce the shoulder-surfing threat by adding an additional input device that is unseen by the user. DeLuca et al. evaluated smartphones and allowed users to input shape passwords using a touchscreen input on either the front or back of the smartphone. Users could then hide some or all the shapes

being drawn as the password. DeLuca et al. found the additional input type only decreased password input speeds by less than four seconds per authentication attempt.

This input system implemented by DeLuca et al. (2014) would work as a possible solution to shoulder-surfing attacks for handheld devices. However, this input system and other graphical-based authentication systems for mobile devices are not as easily translated to traditional information system terminals. Handheld devices are smaller and held closer to the user's body, making it difficult for an attacker to successfully compromise the authentication system. Traditional information system terminals are often stationary systems with their monitors positioned so others can see what the worker is doing as an accountability measure. The implementation of a mobile device graphical-based authentication system in a traditional information system environment would not be a one-to-one match. The differences in the environments would be substantial enough to not protect the traditional system against shoulder-surfing attacks.

Graphical authentication systems have real advantages in security and usability in the mobile space. Chang et al. (2012) conducted a study on the implementation of a graphical authentication system on mobile devices. Chang et al. found that graphical authentication systems did not add any burden to the users and had a minimal risk of shoulder-surfing breaking the authentication scheme. This is caused by the more personal and smaller nature of mobile devices. These devices are designed to be held closer to users making it difficult for attackers to clearly view graphical password input.

Georgakakis et al. (2012) introduced a graphical password scheme (NAVI) that uses map routing information to create a user-friendly method of creating strong graphical passwords. Through their research, many possible risk areas were mitigated;

however, the shoulder-surfing attack was still an ever present danger to the system. The NAVI system was very secure in brute force and dictionary attacks; however, traditional text-based passwords were still shown to provide the greater protection to shoulder surfing attacks. Georgakakis et al. suggested the possibility of using additional techniques to limit the effectiveness of shoulder-surfing attacks. Georgakakis et al. described the use of a transparent image layered on top of the password interface. This would decrease the range of visibility for an attacker to clearly see the solution entered by the user. However, this solution did not eliminate the shoulder-surfing threat.

Another possible solution proposed by Georgakakis et al. (2012) would add several possible challenge questions for the user to answer to authenticate. Multiple challenges would still provide information to a shoulder-surfer, but the attacker would not be guaranteed to succeed in an attack unless the challenge asked by the system was the challenge observed by the attacker. Security can be improved by changing what the user is asked because it causes the attacker to acquire more information to successfully compromise a system.

PassPoint-style graphical passwords ask users to select multiple points on an image in sequence (Van Oorschot, Salehi-Abari, & Thorpe, 2010). These types of passwords have been shown to meet usability requirements and have been successful with users. However, graphical-based passwords that use this method of points are susceptible to hotspot area security issues. Within any given image there are common points that are naturally given to being points of interest to users. Van Oorschot et al. developed an efficient method of automated attacks on PassPoint-style passwords exploiting the hotspot pattern nature of these passwords. This security flaw is also present

in text-based passwords when users are given the choice to select a password. This can be mitigated by not allowing password creation by the users, but instead assigning randomly created passwords that are not predisposed to natural human tendencies.

Stobert and Biddle (2013) compared PassPoint-style graphical passwords to evaluate the memorability of the different types of graphical passwords. The first type is a PassPoint-style with a blank screen as the background. This represented a purely recall-based password with no hint for the user on where to click for authentication. The second type is a PassPoint-style graphical password with an image as the background. This represented a cued-recall password where there are reference points for the user to select, but there are not clear choices for authentication. The third type was an object-based PassPoint where each point is a separate object. This represented a recognition password where the user is given several clear choices for authentication. Sorert and Biddle tested users with randomly assigned passwords in these different modes and tracked the results. Users found object-based PassPoint graphical passwords that were randomly assigned were the most memorable of the graphical password systems tested. Sorert and Biddle also noted that login times for this system averaged 20 seconds, which was longer than traditional text-based password.

Schaub, Walch, Könings, and Weber (2013) studied the implementation of graphical-based authentication systems on smartphones. They noted that many smartphones do not have a traditional keyboard built into the hardware and instead they rely on touch interface. A touch interface is a very suitable platform to use graphical-based authentication because the user interaction can be dynamically shown and inputted using this technology. Schaub et al. implemented five different graphical-based

authentication systems on smartphones and evaluated the usability and the shoulder-surfing vulnerability of those systems. The results of this study showed that drawn patterns were more susceptible to shoulder-surfing attacks than the recognition-based systems. Schaub et al. also noted the need for more research into innovation in the graphical-based passwords that use recognition to help make them more secure.

There are some security weaknesses in the implementation of graphical passwords through the use of Web-based environments. In a Web-based environment, the server could be completely secured; however, client machines could have any number of possible security vulnerabilities that would compromise a graphical authentication system. Environment issues are compounded as systems come further into the website domain of system implementation. The way a graphical system displays a challenge to a user could potentially introduce more potential problems as Web-based attackers can use page refreshes, website history, and client-side source code to start deciphering methods of attack (Renaud & Olsen, 2007).

Methods for Evaluating Passwords

Researchers have created widely accepted methods of evaluating the effectiveness of traditional text-based passwords. These methods are used to help calculate the security strength of a text-based authentication system. In looking towards the implementation of a graphical-based authentication system, there is no established evaluation process. This makes the comparison process between the security of the two methods difficult to quantify (Biddle et al., 2012; Esch, 2003). As research in graphical-based authentication progresses, there is a need to establish how best to evaluate security of graphical-based passwords and how to compare that quantification with traditional text-based methods.

Gati and Tversky (1987) looked at the memory recollection of both unique and similar features from verbal and pictorial items. Participants were able to recall similarities between two items more often than the differences. This study supports the idea that two images of the same subject matter could be used interchangeably in an authentication system.

Ritov, Gati, and Tversky (1990) studied pictorial and verbal stimuli and the ability of people to determine similarities and differences. Through their research, Ritov et. al. found that pictorial and verbal representations of the same stimuli resulted in no difference in the relative weight of the commonality between the two.

Cognitive Load Measurement

Cognitive load measurement is a measure of cognitive processing capacity of individuals to apply acquired knowledge and skills to situations (Paas, Tuovinen, Tabbers, & Van Gerven, 2003). Paas and Van Merriënboer (1994) showed that cognitive load can be measured by determining a user's mental load, mental effort, and performance. The ability to measure an individual's cognitive load while using a computer information system allows for better evaluation of the effectiveness of that system.

Csinger (1992) looked at the theories in psychology as they relate to cognitive load. More time is required for tasks that are more complicated and require more attentive processing. The brain ranks tasks that are easier as requiring less time to accomplish the task successfully.

Mental workload has been evaluated in prior research through the use of a subjective workload instrument (Rubio, Diaz, Martin, & Puente, 2004). The evaluation of

subjective workload will often involve users answering several questions regarding the act of completing a task in a given context.

Reid and Nygren (1988) developed an instrument known as Subjective Workload Assessment Technique (SWAT) that measures these three areas of mental load: time, mental effort, and physiological stress. Through the use of a simple high, medium, and low scale users evaluate several aspects of the three areas of mental load, and the results are then scored through a defined procedure.

Another instrument is the NASA-TLX developed by Hart and Staveland (1988) to evaluate workload based on multiple dimensions. The six dimensions factored into the evaluation of workload are: Effort, Frustration, Own Performance, Mental Demands, Physical Demands, and Temporal Demands. NASA-TLX has been found to be a robust tool in the measurement of subjective workload (Moroney, Biers, & Eggemeier, 1995). Hart (2006) looked at the use of NASA-TLX in a span of 15 years of peer-reviewed literature and verified the acceptance of NASA-TLX as a viable subjective scale of user workload.

NASA-TLX is a tool that has been used in many studies to evaluate workload by users after interacting with a system. Rubio et al. (2004) compared NASA-TLX to work profile methods, and the SWAT and found all methods to be valid and produced similar results in the areas of intrusiveness sensitivity, convergent validity, concurrent validity, diagnosticity, implementation requirements, and acceptability. Yost and North (2006) used NASA-TLX in evaluating cognitive workload in a study looking at the scalability of information visualizations. Hart (2006) showed how NASA-TLX is used as a benchmark

for the efficacy of other measures, theories, and models in several different areas of current literature.

This research chose to use the NASA-TLX because it has been validated through many studies and has been shown to be a reliable instrument for subjective workload measurement. The NASA-TLX is public domain and can be used without additional cost. The NASA-TLX provided this study a measure on the subjective workload in using graphical-based authentication systems.

Summary

This chapter reviewed the key areas of graphical-based authentication systems including text-based passwords, cognometric systems, security issues of text-based authentication, other authentication methods such as token-based authentication and biometric authentication, graphical authentication systems, methods for evaluating passwords, and cognitive load measurement. This literature review showed the significance of study into the impact of pictorial synonyms in graphical-based authentication systems.

There is a significant amount of research showing that further research is needed on authentication system solutions to address the usability and security issues of current authentication systems. Traditional text-based passwords, though being prevalent in computer information systems, continue to have security and usability problems. Graphical-based authentication systems are an area of interest to potentially solve the usability and security issues that are found in the text-based authentication systems.

This research contributed to the understanding of graphical-based authentication systems by addressing the impact on usability with the implementation of pictorial

synonyms. This helps to guide the understanding of graphical-based authentication systems to help increase both the security and usability of these systems.

Chapter 3

Methodology

Overview of Research Methodology

This research determined the cognitive impact of implementing a graphical-based authentication system that uses pictorial synonyms. Quantitative measurements were used to evaluate if randomly generated graphical passwords with pictorial synonyms had a significant impact on cognitive load while maintaining security integrity. A comparative analysis was conducted on authorization systems that use randomly generated text words, static graphical passwords, or graphical passwords with pictorial synonyms.

The following research questions were addressed through this study:

1. What is the impact of a randomly generated graphical-based password with pictorial synonyms on cognitive load?
2. How would the security threat of shoulder-surfing be affected by implementing a graphical-based authentication system that uses pictorial synonyms?
3. What are the implications to user's effectiveness using a graphical-based authentication system using pictorial synonyms?

To answer these questions user performance time, accuracy, and subjective workload evaluation was used to measure the cognitive load implications of the different authentication schemes. The measures and their relationship is depicted in figure 1.

	Performance Time	Accuracy	Cognitive Workload
RQ1	X	X	X
RQ2	X	X	
RQ3	X	X	

Figure 1. Research question measures.

The performance time was measured to evaluate if the proposed authentication system maintained, increased, or decreased the time required to input a password. This measurement aided in the evaluation of RQ1, RQ2, and RQ3.

Accuracy was recorded to measure the successfulness of users with the proposed system and it indicated the memorability of the proposed authentication system. This measurement aided in the evaluation of RQ1, RQ2, and RQ3.

Subjective workload was used to better evaluate the user's cognitive load requirements in using the proposed system. This measurement aided in the evaluation of RQ1.

Performance time and accuracy was automatically logged by the authentication system. When the user began to enter their password the system started a timer behind the scenes and stopper the timer when the user submitted a password attempt. Every authentication attempt and the result of that attempt was logged automatically by the system. The user's subjective cognitive load was self measured using the NASA-TLX survey at the end of their individual participation.

Specific Research Methods Employed

There were three groups consisting of approximately 28 participants per group. These groups were comprised of volunteers from five classes in the Business Administration and Computer Science Department of Pensacola Christian College. To eliminate bias, the classes used were not taught by the researcher. The participants were given 10% extra credit on one quiz grade after successful completion of the study as an incentive to volunteer.

Participants were given an identification number to match their usage records and their survey information. All the login information was coded, and only the researcher had the list of names and codes.

The participants were informed of the requirement to log into a website to obtain course work and an authentication system would be in place. The participants were advised not to discuss the study with others and to contact their instructor with any questions on the system. The course instructor was advised to forward all questions to the researcher to address any issues that arose. The participants were randomly assigned into groups as they volunteered and they could quit or ask to be removed from the study at any time if they chose end their participation. A total of 91 students agreed to participate in the study at the beginning. No users asked to quit the study however, a total of 84 users completed the final survey and were given extra credit in their class. It is unknown why the 7 users did not complete the final survey.

Hwang and Salvendy (2010) performed an analysis of usability study papers and determined that the optimal number to find the majority of errors in a system was 10+/-2.

With a sample size of 10+/-2, 85% of issues with a system could be identified in a investigative study.

Schmettow (2012) argued against the use of a magic number in sample sizes for usability studies. In reviewing Hwang and Salvendy (2010), Schmettow found that there was consensus in finding a majority of issues and not the entirety of issues in a system. Schmettow does acknowledge the usefulness of smaller sample sizes for iterative processes but encourages larger sample sizes for more accurate results. A sample size of 28 participants per group far exceeds the accepted sample size of 10+/-2 to ensure more robust results.

All groups were given a randomly generated password of several words. The method of entry of the password varied for each of the group types. The first group used text-based passwords. The random password given to the users was entered using a recognition based password screen where each of the possible word choices were displayed on the screen with the user choosing their password. This is demonstrated in figure 2. Each time the users came to the authentication screen the words were placed in a random location to ensure the users are using recognition for their password instead of muscle memory of word locations.

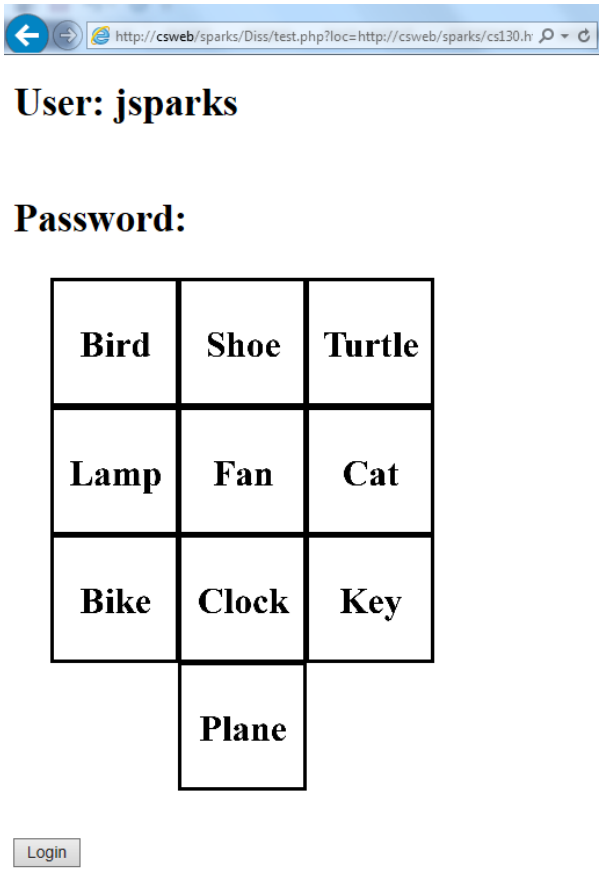


Figure 2. Text-based password screen.

The second group was given a static visual palette of images to use to enter the randomly assigned password. This is shown in figure 3. Like group 1, each time the users come to the authentication screen the images were placed in a random location to ensure the users are using recognition for their password.

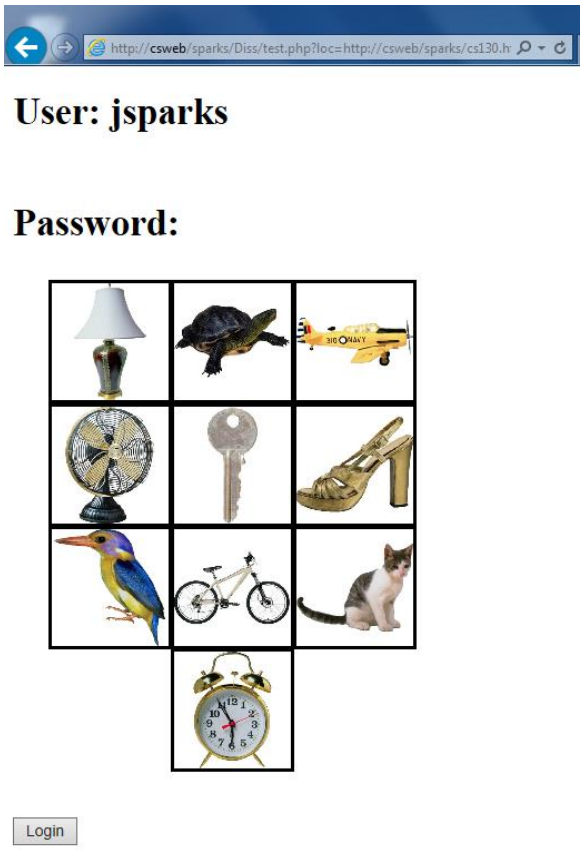


Figure 3 Graphical-based Password Screen

The final and third group used the entry method that incorporates random synonym images to enter the randomly assigned password. The input method was the same as group 2; however, the images were randomly chosen from 17 possible choices for each image category. These images were used by Konkle, Brady, Alvarez, and Oliva (2010) to study the limitations of visual memory of real world objects. Konkle et al. used many different objects which had a varying number of images to test the ability of participants to remember many images. Konkle et al. provided all the images used in their experiment for further study use. The synonym password system used in this research implemented ten of the categories used in Konkle et al. that had a total of 17 images and best fit the needs of this research. The ten selected categories and images are shown in Appendix A.

The subjects used the system for a six week time span. The system was designed to grant the users access to course content upon successful authentication. If the user failed to authenticate two consecutive times they had the option to reset their password and were then granted access to their course content. The users needed to access this content many times throughout the six-week time span provided a good sample of normal password usage. During this time, usage data was gathered automatically by the authentication system. This data consisted of time to login, number of successful and unsuccessful login attempts, and number of password reset requests. At the end of the study period, the subjective workload of the participants was evaluated using NASA-TLX (Task Load Index). The NASA-TLX (see Appendix B) is a short survey to measure subjective cognitive workload and took approximately five minutes to complete.

The time spans that have been used to test memory recall in various research studies have varied from almost immediately to longer periods of weeks and months and even into years. Schaub et al. (2013) used one 30 minute session with participants to gather usability information on the design of graphical-based passwords on smart phones measuring entry times and success rates. Chowdhury et al. (2014) used a 14 day span between their participants receiving a graphical-based password and using that password to test the memorability of that password. Hub, Capek, Myskovs, and Roudnu (2010) used a two month period in the evaluation of the retention of passwords to measure entry times, success rates, and memorability. Tullis et al. (2011) used a six-year span to test the memorability of graphical-based passwords over a long period. This research used a six-week timeframe. This allowed participants a sufficient amount of time to measure retention and gather good usability data.

Instrument Development and Validation

To verify the developed system, a pilot study was conducted with four participants. These participants were not allowed to participate in the six-week study. This was conducted over the span of a week to verify system stability. The pilot study was to verify the system would not crash on the subjects, and the system was correctly measuring the data. The data from the pilot system was not tracked or reported on. Once the system was verified, the actual study was conducted.

At the end of the six week period, the participants were given the NASA-TLX survey to measure cognitive load. The NASA-TLX is available in the public domain, and there is no cost associated with using this instrument.

Formats for Presenting Results

Once the six week research period ended and the data was gathered (time to login, number of successful/unsuccessful login attempts, number of password reset request and subjective workload), the data from each group was coded, evaluated, and analyzed. Analysis of variance was used to complete the analysis of the research data.

RQ1 was addressed by looking at the following areas: subjective work load scores where lower scores indicate less required workload; evaluating the amount of time users needed to authenticate where less time shows easier usage of the system; and by evaluating the proportion of successful login attempts where the higher the proportion the more robust the system.

RQ2 was addressed by evaluating the amount of time users needed to authenticate where less time shows less opportunity for a shoulder-surfing attack and by evaluating

the proportion of successful login attempts where the higher the proportion the more robust the system decreasing the likelihood of a prolonged login session.

RQ3 was addressed by evaluating the amount of time users needed to authenticate and by evaluating the proportion of successful login attempts where the higher the proportion the more robust the system.

Resource requirements

The developed authentication systems ran on a standard windows server running PHP and MySQL. Pensacola Christian College (the author's workplace) authorized the use of one of the college's servers for the study.

PSPP (a free alternative to IBM's Statistical Package for the Social Sciences) was used for calculating and reporting the analysis of variance of the data collected. PSPP is an open source project for use in research.

Nova Southeastern University requires an Institutional Review Board approval before any experiment involving human test subjects is conducted. This approval was acquired through the Institutional Review Board by completing the required forms and following the appropriate procedures and is included in Appendix C.

The participants were students of Pensacola Christian College (all over the age of 18) and permission was granted by the college to permit their participation. The President of Pensacola Christian College was contacted to obtain the permission to work with the participants and approval was given and is included in Appendix D.

Summary

An investigative study was conducted to determine the impact of pictorial synonyms on graphical-based authentication systems by using various graphical-based

authentication systems. During this study user performance time, accuracy, and subjective workload were measured to determine the cognitive load implications of the different graphical based authentication schemes.

This research was designed to expand the knowledgebase of authentication systems by evaluating the introduction of pictorial synonyms to graphical-based authentication systems. The methodology described ensures the validity of this research through its approach and instruments.

Chapter 4

Results

Introduction

This research evaluated the usability of a graphical-based authentication system that implemented pictorial synonyms. The following research questions and hypotheses were established to determine the impact of implemented pictorial synonyms in graphical-based authentication systems:

RQ1: What is the impact of a randomly generated graphical-based password with pictorial synonyms on cognitive load?

RQ2: How would the security threat of shoulder-surfing be affected by implementing a graphical-based authentication system that uses pictorial synonyms?

RQ3: What are the implications to user's effectiveness using a graphical-based authentication system using pictorial synonyms?

H1: There is no significant difference on a user's cognitive load of a randomly generated graphical-based password with pictorial synonyms compared to passwords without pictorial synonyms.

H2: There is no significant difference to security threat of shoulder-surfing by implementing a graphical-based authentication system that uses pictorial synonyms compared to a system without pictorial synonyms.

H3: There is no significant difference to user's effectiveness using a graphical-based authentication system using pictorial synonyms compared to graphical-based systems without pictorial synonyms.

Findings

This research methodology described in Chapter 3 generated quantitative results from the subject's usage of the graphical-based authentication systems. Table 1 shows a summary of the usage of the authentication systems.

	Count of Users	**Mean User Age**	**Mean Password Resets per User**	**Mean Total Login Attempts per User**
Word Passwords	28	19.29	1.11	32.39
Static Picture Passwords	29	19.52	0.24	29.48
Pictorial Synonym Password	27	19.19	0.26	27.37
Means	28	19.33	0.54	29.75

Table 1 Authentication Usage Summary

The groups had a similar sample size (M = 28) with an age of approximately 19.33 years old (SD = 1.6). In the study time span of six weeks, the users of the word password systems made 32.39 login attempts and had a mean of 4.97 seconds to enter their password. The static picture group made 29.48 attempts and had a mean of 4.27

seconds to enter their attempt. The pictorial synonym group had a mean of 27.37 login attempts and had a mean of 4.84 seconds to enter their attempt (see table 1 and table 2).

Hypothesis 1

Chapter 3 described how this study used performance time, accuracy, and NASA-TLX score to determine the validity of null hypothesis H1. This section will discuss the findings for the measures related to H1.

Table 2 shows the mean login attempt time per user which is used for the evaluation of performance time. The mean login attempt time per user was calculated by collating the mean time of each user by the test group. The mean login attempt time per user in the word password group was 4.97 seconds per attempt. . The mean login attempt time per user in the static picture group was 4.27 seconds per attempt. The pictorial synonym group had a mean login attempt time per user of 4.84 seconds. An analysis of variance was performed on the mean login attempt time per user and there was no significant difference found, $F(2, 81) = 2.19$, $p = 0.12$ (Appendix E1).

	Mean of Login Attempt Time per User (Secs)
Word Passwords	4.97
Static Picture Passwords	4.27
Pictorial Synonym Password	4.84
Means	4.69

Table 2 Authentication Attempt Summary

Accuracy was evaluated through the measures of number of password resets and the proportion of successful login attempts. Table 1 shows the word password group had a mean of 1.11 (SD = 2.5) resets requests indicating a forgotten password. The static picture group had a lower number of reset requests (M = 0.24, SD = 0.95). The pictorial synonym group had only a slightly larger mean reset rate compared to the static picture group (M = 0.26, SD = 0.59). An analysis of variance was performed on the password reset count of each of the test groups (Appendix E2). There was no significant difference in the number of password resets, $F(2, 81) = 2.73$, $p = 0.07$.

The accuracy of the users was also measured using the proportion of successful login attempts. If the user entered the correct password, it was considered successful, and an incorrect attempt was considered unsuccessful. The results of user accuracy can be seen in Table 3. A null hypothesis was assumed that the proportions of the three groups were equal. Upon a chi square analysis of the proportions it was found that the proportion of successful login attempts by participants did differ by authentication method, $X^2 (2, N = 3) = 6.62$, $p = .04$. The data provides evidence that the proportions are not equal as claimed. There is little difference between static picture passwords and pictorial synonyms. The major difference in proportions appear to be with the word password group.

	Proportion of successful login attempts
Word Passwords	.689
Static Picture Passwords	.785
Pictorial Synonym Password	.759

Table 3 Attempt Accuracy Summary

The users completed a NASA-TLX survey after the six-week time of the study to evaluate cognitive load implications of a pictorial synonyms in graphical-based authentication system (Table 4).

	Mean NASA-TLX Score (0 – 100)	**SD NASA-TLX Score**
Word Passwords	32.30	23.01
Static Picture Passwords	32.92	17.50
Pictorial Synonym Password	31.47	14.85

Table 4 NASA-TLX Score Summary

The NASA-TLX score provides a zero to 100 score of the perceived cognitive load of a user. The lower the number the less perceived cognitive load. Pictorial synonym group had the lowest mean score of the groups (M = 31.47, SD = 14.85). The word password group had mean NASA-TLX score of 32.30 (SD = 23.01). The static picture

group had a slightly higher mean NASA-TLX score of 32.92 (SD = 17.50). Analysis of variance was performed and no significant difference was found between the three groups, $F(2, 81) = .04$, $p = 0.96$ (Appendix E3).

H1 states, "there is no significant difference on a user's cognitive load of a randomly generated graphical-based password with pictorial synonyms compared to passwords without pictorial synonyms." The analysis of these data points revealed in an aspect of accuracy there is not a significant difference between groups, and thus the data indicates a failure to reject H1.

Hypothesis 2

H2 states, "there is no significant difference to the security threat of shoulder-surfing by implementing a graphical-based authentication system that uses pictorial synonyms compared to a system without pictorial synonyms." Chapter 3 described how this study used performance time and accuracy to determine the validity of the null hypothesis. As discussed in the findings section of H1, the performance time and password resets showed no significant difference between groups. However, proportion of successful login attempts did differ by authentication method, X^2 (2, N = 3) = 6.62, p = .04 and provides evidence that the proportions are not equal as claimed. The analysis of these data points revealed there is not a significant difference between the graphical password groups and the word password groups. The data does not show a difference with the pictorial synonym group and the static picture group. The data indicates a failure to reject H2.

Hypothesis 3

H3 states, "there is no significant difference to user's effectiveness using a graphical-based authentication system using pictorial synonyms compared to graphical-based systems without pictorial synonyms." Chapter 3 described how this study used performance time and accuracy to determine the validity of the null hypothesis. As discussed in the findings section of H1, the performance time and password resets showed no significant difference between groups. The analysis of these data points revealed there is not a significant difference between groups, and thus the data indicates a failure to reject H3.

Summary of Results

The results of this study suggest that the implementation of pictorial synonyms in graphical-based authentication systems has a significant impact on cognitive load, shoulder-surfing threat, and user effectiveness. By using analysis of variance testing of user performance time, password resets, and user cognitive load there was no support found for rejecting H1, H2 and H3 based on these measures. However, the proportion of successful login attempts provide evidence that there is a significant difference between the test groups in the area of accuracy. This gives evidence that there is a difference in the graphical- based authentication compared to word-based authentication; however, this provides no indication on the impact of pictorial synonyms. The results of this study indicates H1, H2 and H3 failed to be rejected. These reported results were used to form the conclusions, implications, and recommendations in Chapter 5.

Chapter 5

Conclusions, Implications, Recommendations, and Summary

Conclusions

The following research questions and hypotheses were established to determine the impact of implemented pictorial synonyms in graphical-based authentication systems:

RQ1: What is the impact of a randomly generated graphical-based password with pictorial synonyms on cognitive load?

RQ2: How would the security threat of shoulder-surfing be affected by implementing a graphical-based authentication system that uses pictorial synonyms?

RQ3: What are the implications to users' effectiveness using a graphical-based authentication system using pictorial synonyms?

H1: There is no significant difference on a user's cognitive load of a randomly generated graphical-based password with pictorial synonyms compared to passwords without pictorial synonyms.

H2: There is no significant difference to the security threat of shoulder-surfing by implementing a graphical-based authentication system that uses pictorial synonyms compared to a system without pictorial synonyms.

H3: There is no significant difference to user's effectiveness using a graphical-based authentication system using pictorial synonyms compared to graphical-based systems without pictorial synonyms.

The analysis on the factors of performance time, NASA-TLX cognitive load measure, and proportion of successful login attempts found authentication systems implementing pictorial synonyms had no significant impact. This indicates that the authentication system that implements pictorial synonyms is similar to authentication systems implementing static pictures in those areas. This provides support that graphic-based authentication systems with pictorial synonyms could be a viable option for securing computer information systems.

The analysis of the data found a failure to reject H1, H2, and H3. The accuracy of the user was impacted by the implementation of graphical-based authentication systems compared to word-based authentication systems. The accuracy of the static picture password group was the most robust of all the groups at .785 of successful attempts. However, the pictorial synonym password group was close in value to the static picture group at .759. The word password group seemed to be the outlier with an accuracy of .689 of successful attempts. The word password group had the worst accuracy which supports the research indicating that graphical-based authentication systems are more usable than word-based authentication systems (Biddle et al., 2012).

This research shows evidence that word-based authentication systems have an impact on users in the area of cognitive load compared to graphical-based authentication systems by impacting the accuracy of login attempts. This research also provides evidence that the security risk of shoulder surfing attacks is impacted through the use of

word-based authentication systems in the area of accuracy compared to graphical-based authentication systems. Finally, this research provides evidence that user's effectiveness is impacted in the area of accuracy by the implementation of word-based authentication systems compared to graphical-based authentication systems.

Implications

This study's findings provides the basis to state that the implementation of pictorial synonyms in authentication systems do not have a significant impact on user's cognitive load and effectiveness in the areas of time to login, accuracy, and cognitive load. This means that there is now evidence that changing static picture authentication systems to use pictorial synonyms would have similar accuracy outcomes. Randomly assigned passwords that use static images are similarly memorable as passwords that use pictorial synonyms. Although the accuracy of the static picture password group was the most robust of all the groups, the pictorial synonym password group was close in value to the static picture group at .759 (see Table 3). These results are similar to Dunphy et al. (2010) who found in a two week study of a similar static picture authentication system login success rate varied from .67 to .89. Biddle et al. (2012) preformed an analysis of state of graphical-based authentication by reporting on the many studies conducted in the area. This research closely matches the summary of findings in Biddle et al. and further validated the findings of this study.

The sample group used by this research was relatively young (M = 19.3, SD = 1.6). The age of the users may have made them more adaptable to a newer authentication system and allowed them to be successful remembering their passwords. An older sample group may provide different results and would be an area of interest for

future research. Shneiderman and Plaisant (2010) discussed that alternative interfaces from traditional interfaces, like the interface presented in this research, can allow seniors to better use authentication systems.

In the area of security, the accuracy was not shown to be significantly different; and, the mean of login attempt time per user was not shown to be significantly different. The attack opportunity is the same between graphical-cased authentication systems that use static images and those that implement pictorial synonyms. The natural question to ask is why would it be beneficial to implement this new authentication method that uses pictorial synonyms? This research shows that there is impact on normal users working with an authentication system that implements pictorial synonyms. If the impact of authentication systems that implement pictorial synonym on attackers is studied in future research and is found to decrease attacker's success rate, then it could make this type of system worth implementing in computer information systems.

This study focused on the impact to the normal user that is working with an authentication system that implements pictorial synonyms; it did not focus on the attackers. Future research should look at the impact of pictorial synonyms on attackers. The research question would be how successful can an attacker be observing passwords that implement pictorial synonyms and accessing restricted systems.

Recommendations

The research presented has shown that in the areas measured there is not a significant difference in using pictorial synonyms in place of static image passwords in graphical-based authentication and word-based passwords in the areas of time to login, accuracy, and cognitive load. Because of these findings, it is recommended that there is a

need for further research to expand the knowledge of the impact on shoulder-surfing attacks when facing an authentication system that implements pictorial synonyms.

Shoulder-surfing attacks occur when a user is able to successfully observe a password and compromise an authentication system. The naïve measure of login attempt time in this study should be bolstered by further investigation of the impact on attackers when pictorial synonyms are implemented in authentication systems. Valuable insight into the implementation of pictorial synonyms in authentication systems could be discovered by conducting a study similar to Dunphy et al. (2010) shoulder-surfing replay attack study. The shoulder-surfing replay attack is where participants observe other participants as they enter a graphical-based password and attempt to compromise an authentication system by replaying the password they observed. If attacking an authentication system that implements pictorial synonyms is significantly more difficult, it would be of benefit for authentication systems to implement pictorial synonyms because the impact is similar to static image systems in terms of login success rate, time to login, and cognitive load on normal users.

Further research into other security threats such as brute force guessing attacks, social engineering, and phishing attacks are of interest to the area of computer information systems. Implementing pictorial synonyms in authentication could possibly change the impact of these other types of security vulnerabilities and warrant further investigation on the implications pictorial synonyms may have on these attacks.

Another recommendation is that this research be expanded in several areas. This study should be expanded to see if the memorability of pictorial passwords are impacted by the user's age. The mean age of users in this study was 19.33. A large segment of user

populations were not represented in this research. Future research could investigate the impact of age on authentication systems that implement pictorial synonyms. This study can be expanded by looking at a different user populations like gender or computer literacy. There was not a large enough sample in this research to draw any conclusions based on gender or computer literacy in evaluation of the impact of authentication systems that implement pictorial synonyms. Expansion of this research to compare different genders or computer literacy rate could further the understanding of the impact of pictorial synonyms on authentication systems.

Another area of interest for future research is to perform a comparative study of the cognitive load implications of recall based authentication compared to recognition authentication. This could provide insight into how the two systems impact user's ability to authenticate.

Future research could also look into the impact of stress to the authentication process of graphical-based authentication. Cognitive load implications of an authentication system could be impacted as stress of work environments and other outside pressures increase.

Finally, the individual psychology traits of users could be evaluated and used to develop passwords that work best with those traits. An investigative study looking at picture passwords versus word password being used with certain psychological traits could provide a better authentication scheme.

Summary

Many common security tasks are now performed on computer information systems and authentication systems are the primary method of securing these systems.

Computer information systems stakeholders rely on the fact that authentication systems are secure and limit the vulnerability of the system.

Password-based authentication systems are commonly used to authenticate. The usability of these password systems decrease as more security requirements for the passwords are placed on the users (Biddle et al., 2012). A widely suggested method of increasing usability and security of password-based authentication systems is the use of graphical-based passwords. These systems are often more susceptible to the problem of shoulder-surfing. Further research is needed to develop ways of combating shoulder-surfing in graphical-based passwords while maintaining both security and usability. This research suggests the use of pictorial synonyms with graphical-based authentication as a solution to combat shoulder-surfing attacks.

A review was performed on the main areas of graphical-based authentication systems including text-based passwords, cognometric systems, security issues of text-based authentication, other authentication methods such as token-based authentication and biometric authentication, graphical authentication systems, methods for evaluating passwords, and cognitive load measurement. The review of the literature provided a foundation to the significance of research into the impact of pictorial synonyms in graphical-based authentication systems.

Biddle et al. (2012) investigated the first 12 years of research of graphical-based authentication systems and found that there is ample research showing a need for further investigation of graphical-based authentication system solutions. Research is needed in graphical-based authentication systems to further address the usability and security issues of current authentication systems. Biddle et al. discussed that graphical-based

authentication systems are a promising area of research because users find graphical-based password more memorable then text-based passwords. Graphical-based authentication has the potential of addressing usability and security issues that are currently found in text-based authentication systems.

An investigative study was completed using 84 participants in three groups over a six week time span to determine the impact of pictorial synonyms on graphical-based authentication systems by using various graphical-based authentication systems. During this research user performance time (the amount of time it took to enter a password) and accuracy (whether the password was correct) were measured as they used the authentication system. Subjective workload was also measured using the NASA-TLX to determine the cognitive load implications of the different graphical based authentication schemes.

This study contributed to the understanding of graphical-based authentication systems by investigating the impact on usability with the implementation of pictorial synonyms. This research helps to guide the understanding of graphical-based authentication systems to help increase both the security and usability of these systems

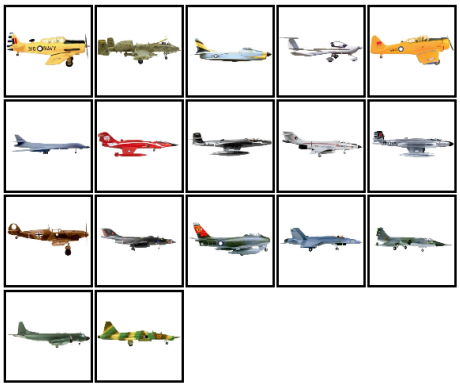
This research expanded the knowledgebase of authentication systems by evaluating the introduction of pictorial synonyms to graphical-based authentication systems. The methodology developed ensured the validity of this research through its approach and instruments

The results of this study suggest the implementation of pictorial synonyms in graphical-based authentication systems does not have a significant impact on cognitive load, shoulder-surfing threat, and user effectiveness in the areas of time to login,

accuracy, and cognitive load. This research concludes that there is evidence that pictorial synonym password showed no significant difference on user's cognitive load and effectiveness in using graphical-based authentication systems. Since there was not a significant difference in implementing pictorial synonyms in authentication systems, further research is now needed to focus on the impact of pictorial synonyms on attackers. If the impact hampers attackers, then the use of pictorial synonyms would be recommended as a solution to the problems found in text-based authentication systems. Further research is also needed to determine if age, gender, or computer literacy have an impact on these results.

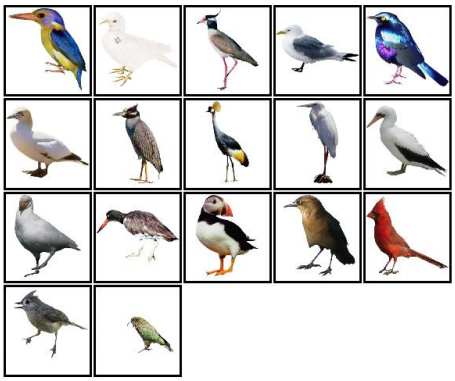
Appendix A. Image Synonyms

Plane

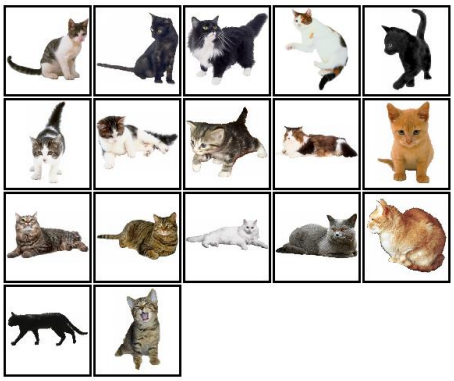


Bike

Bird



Cat



Clock

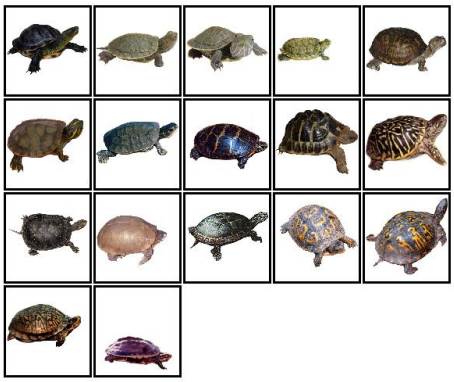
Fan

Key

Lamp

Shoe

Turtle

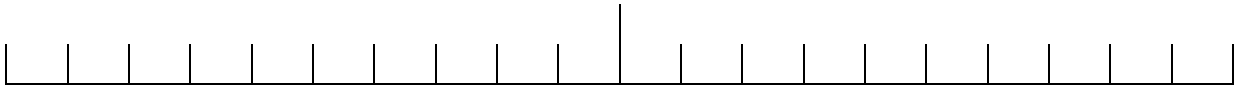


Appendix B. NASA-TLX

ID Number: Date:

Please circle a line for each of these six questions based on the scale provided.

1. **Mental Demand** How mentally demanding was it using the authentication system?



Very Low Very High

2. **Physical Demand** How physically demanding was it using the authentication system?

Very Low Very High

3. **Temporal Demand** How hurried or rushed was the pace of using the authentication system?

Very Low Very High

4. **Performance** How successful were you in authenticating?

Perfect Failure

5. **Effort** How hard did you have to work in accomplishing your level of performance?

Very Low Very High

6. **Frustration** How insecure, discourage, irritated, stressed, and annoyed were you authenticating?

Very Low Very High

For each of the following please circle the more important factor in you evaluation of this system.

1.	Effort	or	Performance
2.	Temporal Demand	or	Frustration
3.	Temporal Demand	or	Effort
4.	Physical Demand	or	Frustration
5.	Performance	or	Frustration
6.	Physical Demand	or	Temporal Demand
7.	Physical Demand	or	Performance
8.	Temporal Demand	or	Mental Demand
9.	Frustration	or	Effort
10.	Performance	or	Mental Demand
11.	Performance	or	Temporal Demand
12.	Mental Demand	or	Effort
13.	Mental Demand	or	Physical Demand
14.	Effort	or	Physical Demand
	Frustration	or	Mental Demand

Appendix C. IRB Approval Memorandum

NOVA SOUTHEASTERN UNIVERSITY
Office of Grants and Contracts
Institutional Review Board

MEMORANDUM

To: Jonathan Sparks

From: Ling Wang, Ph.D.
Institutional Review Board

Date: Sep. 5, 2014

Re: *The Impact of Image Synonyms in Graphical-Based Authentication Systems*

IRB Approval Number: wang08151406

I have reviewed the above-referenced research protocol at the center level. Based on the information provided, I have determined that this study is exempt from further IRB review. You may proceed with your study as described to the IRB. As principal investigator, you must adhere to the following requirements:

1) CONSENT: If recruitment procedures include consent forms these must be obtained in such a manner that they are clearly understood by the subjects and the process affords subjects the opportunity to ask questions, obtain detailed answers from those directly involved in the research, and have sufficient time to consider their participation after they have been provided this information. The subjects must be given a copy of the signed consent document, and a copy must be placed in a secure file separate from de-identified participant information. Record of informed consent must be retained for a minimum of three years from the conclusion of the study.

2) ADVERSE REACTIONS: The principal investigator is required to notify the IRB chair and me (954-262-5369 and 954-262-2020 respectively) of any adverse reactions or unanticipated events that may develop as a result of this study. Reactions or events may include, but are not limited to, injury, depression as a result of participation in the study, life-threatening situation, death, or loss of confidentiality/anonymity of subject. Approval may be withdrawn if the problem is serious.

3) AMENDMENTS: Any changes in the study (e.g., procedures, number or types of subjects, consent forms, investigators, etc.) must be approved by the IRB prior to implementation. Please be advised that changes in a study may require further review depending on the nature of the change. Please contact me with any questions regarding amendments or changes to your study.

The NSU IRB is in compliance with the requirements for the protection of human subjects prescribed in Part 46 of Title 45 of the Code of Federal Regulations (45 CFR 46) revised June 18, 1991.

Cc: Protocol File

03301 College Avenue • Fort Lauderdale, FL 33314-7796 • (954) 262-5369
Fax: (954) 262-3977 • Email: *inga@nsu.nova.edu* • Web site: www.nova.edu/cwis/ogc

Appendix D. Pensacola Christian Collage Approval Letter

Executive Offices
PENSACOLA CHRISTIAN COLLEGE®
P.O. BOX 18000 • PENSACOLA, FL 32523-9160 • U.S.A.
(850) 478-8496 • PCCinfo.com

August 29, 2014

Dr. Ling Wang
Graduate School of Computer
and Information Sciences
3301 College Avenue
Fort Lauderdale, FL 33314

Dear Institutional Review Board Members:

Pensacola Christian College gives permission to Jonathan Sparks to conduct the proposed study "The Impact of Image Synonyms In Graphical-Based Authentication Systems." This permission is for the fall 2014 school term. The permission grants Jonathan Sparks the use of students, over the age of 18, currently attending Pensacola Christian College and the use of computer systems at Pensacola Christian College. The methodology of the proposal approved by the Nova Southeastern University dissertation committee and Nova Southeastern University's IRB is also approved by Pensacola Christian College.

Sincerely,

Troy A. Shoemaker, Ed.D.
President

TAS:kls

Appendix E. Descriptive Analysis of Variance Tables

E1. ONEWAY Analysis of Variance on Time to Login

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Time To Login	Word Passwords	28	4.97	1.48	.28	4.39	5.55	2.7031667	8.8469333
	Static Picture Passwords	29	4.27	1.09	.20	3.85	4.68	2.3128462	7.0383750
	Pictorial Synonym Password	27	4.84	1.45	.28	4.27	5.42	2.9502500	8.9014231
	Total	84	4.69	1.37	.15	4.39	4.98	2.3128462	8.9014231

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
Time To Login	1.00	2	81	.37

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Time To Login	Between Groups	7.98	2	3.99	2.19	.12
	Within Groups	147.84	81	1.83		
	Total	155.82	83			

E2. ONEWAY Analysis of Variance on Mean User Password Reset

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Reset Password Count	*Word Passwords*	28	1.11	2.50	.47	.14	2.08	0	12
	Static Picture Passwords	29	.24	.95	.18	-.12	.60	0	5
	Pictorial Synonym Password	27	.26	.59	.11	.02	.49	0	2
	Total	84	.54	1.62	.18	.18	.89	0	12

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
Reset Password Count	5.05	2	81	.009

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Reset Password Count	*Between Groups*	13.72	2	6.86	2.73	.071
	Within Groups	203.17	81	2.51		
	Total	216.89	83			

E3. ONEWAY Analysis of Variance on Mean User NASA-TLX Score

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
NASA-TLX Score	Word Passwords	28	32.30	23.01	4.35	23.38	41.22	.0000	82.6667
	Static Picture Passwords	29	32.92	17.50	3.25	26.26	39.58	.0000	65.0000
	Pictorial Synonym Password	27	31.47	14.85	2.86	25.59	37.34	8.6667	65.0000
	Total	84	32.25	18.57	2.03	28.22	36.28	.0000	82.6667

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
NASA-TLX Score	3.55	2	81	.03

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
NASA-TLX Score	Between Groups	29.53	2	14.76	.04	.96
	Within Groups	28601.61	81	353.11		
	Total	28631.14	83			

References

Adams, A., & Sasse, M. A. (1999). Users are not the enemy. *Commun. ACM, 42*(12), 40-46. doi: 10.1145/322796.322806

Back, J., & Oppenheim, C. (2001). A model of cognitive load for {IR}: implications for user relevance feedback interaction. *Information Research 2001*. doi: citeulike-article-id:572028

Bang, Y., Lee, D. J., Bae, Y. S., & Ahn, J. H. (2012). Improving Information Security Management: An Analysis of ID-Password Usage and a New Login Vulnerability Measure. *International Journal of Information Management, 32*(5), 409-418. doi: 10.1016/j.ijinfomgt.2012.01.001

Biddle, R., Chiasson, S., & Van Oorschot, P. C. (2012). Graphical Passwords: Learning from the First Twelve Years. *ACM Computing Surveys, 44*(4). doi: 10.1145/2333112.2333114

Bonneau, J., Herley, C., van Oorschot, P. C., & Stajano, F. (2012). *The Quest to Replace Passwords: A Framework for Comparative Evaluation of Web Authentication Schemes*. Paper presented at the Proceedings of the 2012 IEEE Symposium on Security and Privacy.

Bright, P. (2011). RSA finally comes clean: SecurID is compromised. from http://arstechnica.com/security/2011/06/rsa-finally-comes-clean-securid-is-compromised/

Chang, T. Y., Tsai, C. J., & Lin, J. H. (2012). A Graphical-Based Password Keystroke Dynamic Authentication System for Touch Screen Handheld Mobile Devices. *Journal of Systems and Software, 85*(5), 1157-1165. doi: 10.1016/j.jss.2011.12.044

Chiasson, S., van Oorschot, P. C., & Biddle, R. (2006). *A usability study and critique of two password managers*. Paper presented at the Proceedings of the 15th conference on USENIX Security Symposium - Volume 15.

Chowdhury, S., Poet, R., & Mackenzie, L. (2014). *Passhint: memorable and secure authentication*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Toronto, Ontario, Canada.

Csinger, A. (1992). *The psychology of visualization*: University of British Columbia, Department of Computer Science.

Davis, D., Monrose, F., & Reiter, M. K. (2004). *On User Choice in Graphical Password Schemes*. Paper presented at the Proceedings of the 13th conference on USENIX Security Symposium - Volume 13, San Diego, CA.

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*(1-2), 128-152. doi: 10.1016/j.ijhcs.2005.04.020

DeLuca, A., Harbach, M., von Zezschwitz, E., Maurer, M. E., Slawik, B. E., Hussmann, H., & Smith, M. (2014). *Now you see me, now you don't: protecting smartphone authentication from shoulder surfers*. Paper presented at the Proceedings of the 32nd annual ACM conference on Human factors in computing systems, Toronto, Ontario, Canada.

Dunphy, P., Heiner, A. P., & Asokan, N. (2010). *A Closer Look at Recognition-Based Graphical Passwords on Mobile Devices*. Paper presented at the Proceedings of the Sixth Symposium on Usable Privacy and Security, Redmond, Washington.

Esch, J. (2003). Comparing Passwords, Tokens, and Biometrics for User Authentication. *Proceedings of the IEEE, 91*(12), 2019-2020. doi: 10.1109/jproc.2003.819605

Flor, D., Herley, C., & Coskun, B. (2007). *Do strong web passwords accomplish anything?* Paper presented at the Proceedings of the 2nd USENIX workshop on Hot topics in security, Boston, MA.

Forget, A., Chiasson, S., & Biddle, R. (2010). *Shoulder-Surfing Resistance with Eye-Gaze Entry in Cued-Recall Graphical Passwords.* Paper presented at the Proceedings of the 28th Annual Chi Conference on Human Factors in Computing Systems.

Gati, I., & Tversky, A. (1987). Recall of common and distinctive features of verbal and pictorial stimuli. *Memory & Cognition, 15*(2), 97-100. doi: 10.3758/BF03197020

Gehring, R. E., Toglia, M. P., & Kimble, G. A. (1976). Recognition Memory for Words and Pictures at Short and Long Retention Intervals. *Memory & Cognition, 4*(3), 256-260.

Georgakakis, E., Komninos, N., & Douligeris, C. (2012). *NAVI: Novel Authentication with Visual Information*.

Grawemeyer, B., & Johnson, H. (2011). Using and Managing Multiple Passwords: A Week to a View. *Interacting with Computers, 23*(3), 256-267. doi: 10.1016/j.intcom.2011.03.007

Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50*(9), 904-908. doi: 10.1177/154193120605000909

Hart, S. G., & Staveland, L. E. (1988). *Human mental workload*: P.A. Hancock and N. Meshkati (Eds.), Amsterdam: Elsevier.

Hub, M., Capek, J., Myskovs, R., & Roudnu, R. (2010). *Usability Versus Security of Authentication*. Paper presented at the Proceedings of the 2010 international conference on Communication and management in technological innovation and academic globalization, Tenerife, Spain.

Hunt, R. R., & Elliot, J. M. (1980). The role of nonsemantic information in memory: Orthographic distinctiveness effects on retention. *Journal of Experimental Psychology, 109*(1), 49-74.

Hwang, W., & Salvendy, G. (2010). Number of People Required for Usability Evaluation: the 10+/-2 Rule. *Communications of the ACM, 53*(5), 130-133. doi: 10.1145/1735223.1735255

Ives, B., Walsh, K. R., & Schneider, H. (2004). The Domino Effect of Password Reuse. *Communications of the ACM, 47*(4), 75-78. doi: 10.1145/975817.975820

Jain, A. K., Ross, A., & Pankanti, S. (2006). Biometrics: a tool for information security. *IEEE Transactions on Information Forensics and Security, 1*(2), 125-143.

Kent, J. (2005). Malaysia car thieves steal finger. from http://news.bbc.co.uk/2/hi/asia-pacific/4396831.stm

Kirushnaamoni, R. (2013, 21-22 Feb. 2013). *Defenses to curb online password guessing attacks.* Paper presented at the Information Communication and Embedded Systems.

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General, 139*(3), 558.

Moroney, W. F., Biers, D. W., & Eggemeier, F. T. (1995). Some Measurement and Methodological Considerations in the Application of Subjective Workload Measurement Techniques. *The International Journal of Aviation Psychology, 5*(1), 87-106. doi: 10.1207/s15327108ijap0501_6

Morris, R., & Thompson, K. (1979). Password security: a case history. *Commun. ACM, 22*(11), 594-597. doi: 10.1145/359168.359172

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational psychologist, 38*(1), 63-71.

Paas, F., & Van Merriënboer, J. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review, 6*(4), 351-371.

Pilar, D. R., Jaeger, A., Gomes, C. F. A., & Stein, L. M. (2012). Passwords Usage and Human Memory Limitations: A Survey across Age and Educational Background. *PLOS ONE, 7*(12). doi: 10.1371/journal.pone.0051067

Reid, G. B., & Nygren, T. E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. *Advances in Psychology, 52*, 185-218.

Renaud, K., & Olsen, E. S. (2007). DynaHand: Observation-Resistant Recognition-Based Web Authentication. *IEEE Technology and Society Magazine, 26*(2), 22-31. doi: 10.1109/mtas.2007.371279

Ritov, I., Gati, I., & Tversky, A. (1990). Differential weighting of common and distinctive components. *Journal of Experimental Psychology: General, 119*(1), 30.

Rubio, S., Diaz, E., Martin, J., & Puente, J. M. (2004). Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychology: An International Review, 53*(1), 61-86. doi: 10.1111/j.1464-0597.2004.00161.x

Schaub, F., Walch, M., Könings, B., & Weber, M. (2013). *Exploring the design space of graphical passwords on smartphones.* Paper presented at the Proceedings of the Ninth Symposium on Usable Privacy and Security.

Schmettow, M. (2012). Sample Size in Usability Studies. *Communications of the ACM, 55*(4), 64-70. doi: 10.1145/2133806.2133824

Shay, R., Komanduri, S., Durity, A. L., Huh, P., Mazurek, M. L., Segreti, S. M., . . . Cranor, L. F. (2014). *Can long passwords be secure and usable?* Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Toronto, Ontario, Canada.

Shneiderman, B., & Plaisant, C. (2010). *Designing the user interface: strategies for effective human-computer interaction* (5th ed.): Addison-Wesley Longman Publishing Co., Inc.

Stobert, E., & Biddle, R. (2013). *Memory retrieval and graphical passwords*. Paper presented at the Proceedings of the Ninth Symposium on Usable Privacy and Security, Newcastle, United Kingdom.

Tetsuji, T., Takehito, O., & Hideki, K. (2006). *Awase-E: Recognition-based Image Authentication Scheme Using Users' Personal Photographs.* Paper presented at the Innovations in Information Technology.

Tullis, T. S., Tedesco, D. P., & McCaffrey, K. E. (2011). *Can Users Remember their Pictorial Passwords Six Years Later*. Paper presented at the CHI '11 Extended Abstracts on Human Factors in Computing Systems, Vancouver, BC, Canada.

Van Oorschot, P. C., Salehi-Abari, A., & Thorpe, J. (2010). Purely Automated Attacks on PassPoints-Style Graphical Passwords. *IEEE Transactions on Information Forensics and Security, 5*(3), 393-405. doi: 10.1109/tifs.2010.2053706

Wright, N., Patrick, A. S., & Biddle, R. (2012). *Do You See Your Password?: Applying Recognition to Textual Passwords*. Paper presented at the Proceedings of the Eighth Symposium on Usable Privacy and Security, Washington, D.C.

Yost, B., & North, C. (2006). The Perceptual Scalability of Visualization. *IEEE Transactions on Visualization and Computer Graphics, 12*(5), 837-844. doi: 10.1109/TVCG.2006.184