Spring 2020

Account Recovery Methods for Two-Factor Authentication (2FA): An Exploratory Study

Lauren Nicole Tiller
Old Dominion University, ltill002@odu.edu

ACCOUNT RECOVERY METHODS FOR TWO-FACTOR AUTHENTICATION (2FA): AN EXPLORATORY STUDY

by

Lauren Nicole Tiller
B.S. May 2018, Old Dominion University

A Thesis Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

PSYCHOLOGY

OLD DOMINION UNIVERSITY

May 2020

Approved by:

Jeremiah D. Still (Director)

Robin Lewis (Member)

Rachel Phillips (Member)

ABSTRACT

ACCOUNT RECOVERY METHODS FOR TWO-FACTOR AUTHENTICATION (2FA): AN EXPLORATORY STUDY

Lauren Nicole Tiller
Old Dominion University, 2020
Director: Dr. Jeremiah D. Still

System administrators have started to adopt two-factor authentication (2FA) to increase user account resistance to cyber-attacks. Systems with 2FA require users to verify their identity using a password and a second-factor authentication device to gain account access. This research found that 60% of users only enroll one second-factor device to their account. If a user's second-factor becomes unavailable, systems are using different procedures to ensure its authorized owner recovers the account. Account recovery is essentially a bypass of the system's main security protocols and needs to be handled as an alternative authentication process (Loveless, 2018). The current research aimed to evaluate users' perceived security for four 2FA account recovery methods. Using Renaud's (2007) opportunistic equation, the present study determined that a fallback phone number recovery method provides user accounts with the most cyber-attack resistance followed by system-generated recovery codes, a color grid pattern, and graphical passcode. This study surveyed 103 participants about authentication knowledge, general risk perception aptitude, ability to correctly rank the recovery methods in terms of their attack-resistance, and recovery method perceptions. Other survey inquires related to previous 2FA, account recovery, and cybersecurity training experiences. Participants generally performed poorly when asked to rank the recovery methods by security strength. Results suggested that neither risk numeracy, authentication knowledge, nor cybersecurity familiarity impacted users' ability to rank recovery methods by security strength. However, the majority of participants

ranked either generated recovery codes, 39%, or a fallback phone number, 25%, as being most secure. The majority of participants, 45%, preferred the fallback phone number for account recovery, 38% expect it will be the easiest to use, and 46% expect it to be the most memorable. However, user's annotative descriptions for recovery method preferences revealed that users are likely to disregard the setup instructions and use their phone number instead of an emergency contact number. Overall, this exploratory study offers information that researchers and designers can deploy to improve user's 2FA- and 2FA account recovery- experiences.

ACKNOWLEDGMENTS

I extend many, many thanks to all the individuals who have contributed to the successful completion of my thesis. My advisor, Dr. Jeremiah Still, deserves special recognition for his continuous efforts, constructive feedback, and enthusiasm throughout the creation of this thesis. I am so grateful for your mentorship and the countless hours you have spent sharing your vast knowledge, insight, and experiences. Thank you to my committee members, Dr. Rachel Phillips and Dr. Robin Lewis, for their patience and excellent guidance to improve the research and manuscript. Thank you to my Research Assistants, Daniel Cote, for your help thoroughly coding participant data and, Melissa Atkinson, for helping with figures. Thank you to my closest friend, Paige Duplantis, for the limitless ways you have positively impacted this graduate school experience. Thank you to my PoD lab team, Janine Mator, Steven Vera, and John Hicks, and cohort friends for your awesome support through the stresses and successes. Thank you to my dad, Barry Tiller, and mom, Sheree Ludwig-Tiller, for their unwavering love, encouragement, and support. Dad, the ways you have supported me are endless, you never fail to make me laugh and always help me keep things in perspective. Mom, I am grateful for the countless hours you spent editing this manuscript, teaching me to have faith, and always being in my corner. Thank you to my son, Blake Mersinger, for always being so supportive, understanding, compassionate, loving, and making me smile; you have helped me along this journey in so many ways and you motivate me to always be better, I love you. Thank you to my Fiancé and best friend, David Krupansky, for his unconditional love and support; your belief in me, reassurance, and optimism has kept me focused and I could not ask for a better person to share life with. Thank you to Diana Richards, for all the ways you have helped me get through graduate school. This thesis could not have happened without all of you and for that, I am forever grateful; God bless you all.

TABLE OF CONTENTS

Page

LIST OF TABLES vii

LIST OF FIGURES viii

INTRODUCTION 1
- CYBER-ATTACK THREATS AND AUTHENTICATION 3
- THE IMPORTANCE OF USER ASSESSMENTS 9
- TWO-FACTOR AUTHENTICATION (2FA) 16
- NEEDS ADDRESSED BY CURRENT RESEARCH STUDY 23

METHODOLOGY 26
- PARTICIPANTS 26
- MATERIALS AND PROCEDURE 27

DATA CODING STRATEGIES 33
- OVERALL AUTHENTICATION KNOWLEDGE SCORE 33
- OPPORTUNISTIC SCORE CALCULATIONS 34
- CODING STRATEGY FOR USER'S SUBJECTIVE SECURITY RANKINGS OF ACCOUNT RECOVERY METHODS 37
- CATEGORIZATION AND CODING STRATEGY FOR USER'S ACCOUNT RECOVERY METHOD PREFERENCE REASONING 40

RESULTS 45
- DESCRIPTIVE DATA ANALYSES 45
- STATISTICAL DATA ANALYSES 60

DISCUSSION 66
- HYPOTHESES AND RESEARCH QUESTIONS 66
- DATA PATTERNS 68
- LIMITATIONS 73

CONCLUSIONS 76

REFERENCES 80

APPENDICES 97
- APPENDIX A: STUDY SURVEY 97
- APPENDIX B: INFORMED CONSENT DOCUMENT 107
- APPENDIX C: RENAUD'S (2007) OPPORTUNISTIC SCORE SPECIFICATIONS 109

VITA 122

LIST OF TABLES

Table Page

1. Frequency Table for Declared Major 26

2. Frequency Table for Type of Electronics Commonly Used 27

3. Total Opportunistic Score for Each Account Recovery Method by Researcher Assistants 37

4. Kendall's (1938) Tau Distance Score Example 38

5. Categories and Subcategories for Coding Recovery Method Preference Reasoning 42

6. Frequency Table for Cybersecurity Familiarity 46

7. Example Frequency Table for 2FA Familiarity 47

8. Frequency Table for 2FA Account Recovery Familiarity 50

9. Frequency of Categories Mentioned 55

LIST OF FIGURES

Figure Page

1. Opportunistic Score Formula Presented by Renaud (2007) 8

2. Color Grid Pattern Used for Account Recovery 22

3. Kendall's (1938) Tau Correlation Coefficient Formula 39

4. Frequency of 2FA Devices Used by Type 48

5. Frequency of the Number of 2FA Devices Enrolled per Account 49

6. Frequency of 2FA Account Recovery Methods Used by Type 51

7. Frequency of Perceived Qualities Related to Account Recovery Methods 53

8. Frequency of Recovery Method Preferred Choice 54

9. Frequency of Open-ended Response Categories Describing Recovery Methods 56

10. Frequency of High Usability Subcategories by Account Recovery Method 57

11. Frequency of Low Usability Subcategories by Account Recovery Method 58

12. Frequency of Visual Processing Preference Subcategories by Recovery Method 59

CHAPTER 1

INTRODUCTION

The most common form of authentication is the single-factor alphanumeric password (Leu, 2017; Zyiran & Haga, 1999). When users are asked to login to a system, they typically verify their identity by authenticating with "something they are" (e.g., fingerprint), "something they know" (e.g., password), or "something they have" (e.g., Swipe card; Grassie, Garcia, & Fenton, 2017). When it comes to knowledge-based authentication, users bear the responsibility of creating strong passwords to ensure the security of their online accounts (Cain & Still, 2018). The security requirements for creating a strong password are cumbersome (Ashford, 2009; Barton & Barton, 1984; Hoonakker, Bornoe, & Carayon, 2009; Labuschagne, Veerasamy, Burke, & Eloff, 2011). To overcome these cognitive burdens, users often produce passwords that reflect common patterns or strategies that are easy to recall, which may reduce their account's resistance to cyber-attacks. To increase account security and to compensate for weak or insecure account protection provided by traditional alphanumeric passwords, some companies (e.g., Microsoft, Google, and Facebook) have started to offer or require their users adopt two-factor authentication (Reese, 2018).

Two-Factor (or Multi-Factor) Authentication (2FA or MFA) is a layered authentication process that requires the user to couple their password with another type of authentication method. In 2018, federal agencies that use dot-gov domains such as the Department of Justice began to prompt officials to add the two-factor security feature to increase the system's intruder attack resistance (Shaban, 2018). However, in the event of a failed second-factor device, regaining account access can be problematic for the authorized user (Tellini & Vargas, 2017). Essentially, when the authentication process is more secure, more information is needed to prove

the user's identity. As a result, regaining account access becomes more difficult (Tellini & Vargas, 2017).

The password reset procedure for systems that only use a single-factor password is different from the 2FA account recovery processes. Renaud (2007), noted that systems using single-factor password authentication fulfill password reset requests by either asking the user to answer a particular question, emailing the user their forgotten password, or emailing the user a secure link that obliges the user to create a new password. The account recovery process for systems that implement 2FA is more complex. Even though passwords may be involved, account recovery is not the same as a basic password reset (Loveless, 2018). Systems implementing 2FA require extra steps to ensure that an account is recovered to its rightful owner. Account recovery procedures are essentially a bypass of the system's main security protocols, which necessitates systems to treat account recovery as an alternative authentication process (Loveless, 2018).

A potential solution is to provide users with an account recovery method that is used as a failsafe for 2FA in the event of a lost, broken, stolen, or unavailable second factor. The purpose of an account recovery method is to maintain the high cyber-attack resistance while still allowing the authorized user account access. There is limited research that evaluates different types of 2FA account recovery method options. In this thesis, we explored the qualities that allow account recovery methods to maintain adequate attack resistance while still permitting account access to the authorized user. We measured the objective and subjective security of four different 2FA account recovery methods. We aimed to gain a comprehensive understanding of the typical end user's knowledge of concepts and threats associated with authentication and to measure individual differences of general risk perception, and account recovery methods preference.

1.1 Cyber-Attack Threats and Authentication

There were an estimated 82,000 reported cybersecurity attacks occurring at businesses around the world in 2016 (Smith, Wilbur, & Spiezle, 2018). This estimate almost doubled to 159,700 by 2017 (Smith et al., 2018). Researchers identified weak passwords as a critical source of security failure within the infrastructure of a system (Cazier & Medlin, 2006; Dawson & Stinebaugh, 2010). Cone, Thompson, Irvine, and Nguyen (2006) found that users put their accounts at risk by creating weak passwords or leaving their computers logged in. However, cybersecurity attacks are an issue for any authentication infrastructure. Brute force, social engineering, intersection, and Over-the-Shoulder (OSA) are several cybersecurity attack techniques commonly deployed to overcome authentication schemes.

1.1.1 Brute Force Cyber-attacks

Brute force attacks are a common threat to alphanumeric passwords. A brute force attack occurs when an attacker inputs multiple password combinations until they gain access (English & Poet, 2012). To increase brute force attack resistance, users are often encouraged to create passwords that are complex and memorable (Hoonakker et al., 2009; Labuschagne et al., 2011); changed often (Barton & Barton, 1984); remain secret (Hoonakker et al., 2009); and differ for each account (Ashford, 2009; Barton & Barton, 1984; Hoonakker et al., 2009). Strong passwords should be comprised of long character strings (Florêncio, Herley, & Coskun, 2007); include upper- and lower- case letters, special characters, numbers, punctuation, and non-dictionary words (Cybersecurity and Infrastructure Security Agency, 2009; Yale, 2007). Despite the exhaustive list of strong password suggestions and requirements, research has shown that 31% of users use the same passwords for all accounts, and 43% of users have never changed their

password (Infosecurity Europe, 2008). Still, Cain, and Schuster (2017) highlighted that the recommended practices for strong password creation lead to users creating passwords that are not memorable and will ultimately force users to invent cognitive workarounds.

The task of recalling a complex alphanumeric password can be cognitively demanding due to human memory limitations (Boechler, 2006). The cognitive science research of Sweller (1988) introduced the cognitive load theory, which suggests that the human mind is analogous to an information processing system with working- and storage- memory. Hogg (2007) reiterated that working memory is limited and defined cognitive load as "the processing of information that occurs in working memory" (p. 188). Previous research has also suggested that security behavior practices exhibited by users are directly impacted by their cognitive limitations (Conklin, Dietrich, & Walz, 2004). Conklin et al. (2004) used models to depict how a user's need for a password memory aid gets intensified as the number of passwords for different accounts rises. Additionally, their results advocate that it is difficult for users to simultaneously adhere to all the password security recommendations for multiple systems due to fundamental cognitive memory limitations (Conklin et al., 2004).

To reduce the inherent cognitive burdens of strong password "rules," authentication scheme designers created graphical passcode schemes. Graphical authentication schemes take advantage of the fact that users can make more meaningful associations with the images (Madigan, 1983; Paivio, 2013). The visual features composing the graphical passcode help memorability by supporting richer encoding. For instance, the graphical authentication scheme, Passpoints, allows users to click on specific points within a picture (Wiedenbeck, Waters, Birget, Brodskiy, & Memon, 2005). The graphical scheme, Use Your Illusion (UYI), tasks users with recognizing a distorted target image from a set of distorted distractor images (Hayashi, Dhamija,

Christin, & Perrig, 2008; Cain & Still, 2018). Another cognitive benefit of deploying a graphical authentication scheme for account recovery is the employment of recognition rather than recall memory retrieval. For example, asking users to recognize their password features in front of them rather than having to recall their password from memory (c.f., Tulving & Watkins, 1973).

However, the employment of graphical authentication comes at the cost of new cyberattack vectors. Visual passcodes used in a shared space can afford other nearby easy access to the passcode at a glance. Thus, graphical schemes have emerged to prevent Over-the-Shoulder Attacks (OSAs). For instance, some use gaze-based input where users select their passcode targets using their eyes (De Luca, Denzel, & Hussmann, 2009). Other graphical schemes like What You See is What You Enter (WYSWYE; Khot, Kumaraguru, & Srinathan, 2012) and Explore-a-Nation (EaN; Tiller, Angelini, Leibner, & Still, 2019) involve translating the graphical target information to another location.

1.1.2 Social Engineering Cyber-Attacks

When users create passwords or select graphical passcode images that pertain to personal preferences, an attacker could produce an educated guess through social engineering attacks. For example, an attacker can research a user to gather useful personal information (e.g., interest, heritage) and later use the data to determine what images might have been selected for the passcode. Researchers of alphanumeric passwords reiterated the need for users to create strong passwords to prevent social engineering attacks (Ashford, 2009; Barton & Barton, 1984; Choong & Greene, 2016; Coventry, Briggs, Jeske, & van Moorsel, 2014; Cox, Connolly, & Currall, 2001; Konieczny, Trias, & Taylor, 2015; Pelgrin, 2014).

1.1.3 Intersection and Over-the-Shoulder Cyber-Attacks

Graphical authentication schemes have also been criticized for their vulnerability to intersection attacks. This type of cyber-attack occurs when an attacker takes multiple video recordings of a user logging in, then cross-references the recordings to discriminate targets from distractors (English & Poet, 2012). For graphical authentication schemes to increase intersection attack resistance, schemes should consistently deploy the target images amongst the same distractor images (Gao, Liu, Dai, Wang, & Chang, 2009; Renaud 2007). As a result, the discriminability of the targets and distractors becomes more challenging for attackers.

The most common cyber-attack threat to graphical authentication schemes is their susceptibility to OSAs, also referred to as shoulder surfing. To determine the amount of resistance a given graphical scheme provides to prevent these types of observation attacks, previous researchers have evaluated participant performance when asked to take an attacker role (Cain & Still, 2016; 2018; De Luca, Hertzschuch, & Hussmann, 2010; Sun, Chen, Yeh, & Cheng, 2016; Tiller et al., 2019; Zangooei, Mansoori, & Welch, 2012). Research conducted by Cain and Still (2018) established that four OSA defense strategies are commonly deployed (i.e., disguising the appearance of targets, grouping the targets among distractors, using gaze-based input, or translating targets to another location). Their results suggested that graphical schemes that implement translating targets to another location- or grouping- strategies provide more resistance to OSAs (Cain & Still, 2018).

Renaud (2007) also highlighted that each type of authentication scheme potentially offers different levels of security against cyber-attacks. The “user” authentication step for accessing a system is the component of the system where the system administrator maintains limited control.

However, system designers can choose different authentication schemes based on organization preference, user's needs, and target level of security.

1.1.4 Cyber-Attack Resistance and Authentication Scheme Selection

An opportunistic score is a tool that can help system designers make better decisions when selecting the best authentication scheme from a subset of options. Renaud's (2007) formula can calculate an opportunistic score for a given authentication mechanism. The level of attack resistance that is necessary to protect users' accounts is unique to each system. Essentially, some authentication schemes provide greater attack resistance, but they may place a more significant cognitive burden on the user or reflect lower authentication success rates. Renaud (2007) provides a process that guides designers in their choice and development of web authentication. Her procedure is risk-aware, considers the user's needs, asset value, and the impact of possible account intrusion. For instance, a system designer may seek to use an authentication scheme that offers both high security and usable qualities. If a designer wants to choose between viable authentication options, an opportunistic score for each scheme can be calculated to determine which scheme is best for narrowing an attacker's window of opportunity to complete a successful attack. The extent to which the authentication scheme's guessability, observability, recordability, and analyzability can be exploited will determine how difficult it will be for an attacker to penetrate the system. These four authentication weaknesses are components of the opportunistic score. The score itself is a summed function of the subcomponents of the four weaknesses divided by the systems resistibility (see Figure 1).

Figure 1

Opportunistic Score Formula Presented by Renaud (2007)

$$\text{Opportunity} = \frac{f(Guessability; Observability; Recordability; Analizability)}{Resistibility}$$

Guessability is traditionally measured as potential password strength. For example, each guess of a four-digit pin has a 1 in 10,000 chance of being correct, which is a strong mechanism to prevent unauthorized attacks. However, users must beware not to select a PIN that relates to personally identifiable information (e.g., date of significant events like birth or wedding). Observability refers to whether or not the passcode can be directly observed, a risk that is often associated with shoulder surfing attacks. Recordability relates to the risk associated with users telling others their password or writing it down. Renaud (2007) further highlighted the need for the designer to consider prevalently used electronic functions such as a print screen or screen record. On the other hand, a gaze-based password cannot be recorded even if those functions are utilized. Analyzability is related to the software vulnerabilities that can be exploited by an attacker. This component helps designers detect software risks that could make the system more vulnerable. For example, a system that requires unique usernames is less vulnerable to cyber-attacks than systems that use email addresses as usernames. However, the opportunity for the software risk factors to become apparent during an attack can be constrained by the system's resistibility. A system's attack resistibility could be increased by limiting the number of inaccurate login-attempts or sending a user an SMS notification whenever their account has been accessed. Aside from making the right system authentication security design decisions, administrators should also educate the end-users about better authentication practices and habits to help diminish the risk that comes from their authentication choices (Still et al., 2017). The

current research calculated the opportunistic score for each account recovery method to determine their cyber-attack resistance. In addition to calculating the opportunistic score associated with a given account recovery method, we considered the knowledge, behaviors, and general risk perception abilities of typical end-users.

1.2 The Importance of User Assessments

For users to practice good information security habits during authentication, some knowledge is required. One way to establish end-user's knowledge level and common authentication practices is to use surveys. Markelj and Bernik (2015) used a questionnaire to examine awareness of- and behaviors towards- security threats specific to mobile phones. They found that the most common form of protection used on mobile phones was authentication, with the majority indicating they use a PIN passcode (Markelj & Bernik, 2015). However, 56.8% of participants specified that they did not use any authentication (Markelj & Bernik, 2015).

Gratian, Bandi, Cukier, Dykstra, and Ginther (2018) used a survey to examine how characteristics of decision-making styles, risk-taking, conscientiousness, and gender predicted cybersecurity behavior intentions. One of their research findings suggested that individual differences influence password generation intentions. More specifically, the results indicated that women generate weaker passwords than men. For the demographics of established major, participants who reported an engineering major were more likely to create stronger passwords than those with a humanities major (Gratian et al., 2018). Additionally, participants who had an avoidant decision-making style were more inclined to generate strong passwords (Gratian et al., 2018).

Cain, Edwards, and Still (2018) conducted an extensive study to evaluate users' cyber hygiene knowledge of threats, concepts, and behaviors by examining cyber topics such as authentication, security software, social networking, web browsing, USB drive use, phishing scams, and Wi-Fi hotspot usage. Their results indicated that people 45 years of age and older generally practice more secure cyber behaviors. Cyber hygiene knowledge did not differ by age. However, males had more knowledge than females. Despite having more knowledge, males and females had no difference in security behaviors. Another finding of their study suggested that users who were a victim of past cyber-attacks reported behaviors and knowledge that did not differ from users who had not been subjected to a cyber-attack. Interestingly, the survey results showed that participants who indicated they had received past cybersecurity training had less knowledge and more risky behaviors than users who reported they had not received training. They found that 81% of their participants (n = 144) had received some form of cybersecurity training. Other research studies that evaluated the proportion of self-identified cybersecurity trained participants found much lower results (19% for college-age students; Aytes & Conolly, 2003; 43% for adults; National Cyber Security Alliance, 2010). The current research used some of the general authentication knowledge, demographic, and behavior questions from Cain et al. (2018) and adapted some of the Renaud (2007) opportunistic score categories to create a comprehensive survey to establish user's general concept- and threat- authentication knowledge.

1.2.1 User's Risk Perception Abilities and Related Behaviors

Risk perception and behaviors are also important individual difference factors that should be considered for cybersecurity and general authentication practices. For instance, Van Schaik et al. (2017) examined how risk perceptions (e.g., the severity of risk) lead to protective behaviors for

installing updates, using antivirus, and using firewalls. Their results indicated that feelings of control and severity of consequences were good risk perceptions categories for predicting good precautionary behavior (Van Schaik et al., 2017). Another survey indicated that users who have propensities towards risk-taking or are conscientious create weaker passwords (Gratian et al., 2018).

The National Institute of Standards and Technology defines risk as, the net negative impact that results from a vulnerability issue after accounting for the probability- and the impact- of the occurrence (Grassie et al., 2017). This study assessed risk from a cognitive perspective and evaluated decision making in terms of user's risk perception abilities. Risk perception is a critical component of risk behavior. More specifically, risk perception is the user's (a.k.a., the decision-maker) assessment of risk inherent in the given situation (Department of Homeland Security Federal Infrastructure Protection Bureau [DHS-IP], 2013). When a user conducts a risk assessment, they account for both the perceived magnitude and probability of the risk (DHS-IP, 2013). Risk perception is influenced by the risk assessment process and individual differences.

Previous research studies have pointed to the importance of establishing end-users general risk assessment abilities. An empirical study conducted by Pattinson (2012) evaluated how users' risk perceptions and their behavior when working with a computer impacted their organization's Information Security (IS). The study used a Repertory Grid Technique (RGT) as an interviewing instrument to elicit IS risk perceptions of computer users (Pattinson, 2012). More specifically, the research examined individual differences in users' IS risk perception according to the user's gender, organizational level, and InfoSec awareness level. One finding suggested that the higher the organizational position that is held by the user, the more concern they will have for the organization risks rather than their risks (Pattinson, 2012). Another finding

indicated that female computer users have a more balanced view of IS risks but do not view the damage to their organization's reputation as a serious risk (Pattinson, 2012).

Another study conducted by Sharit et al. (2014) established the role that graph literacy, health literacy, and numeracy abilities had in enabling veterans to perform tasks using My HealtheVet (MHV) system (c.f., the Department of Veteran Affairs' Personal Health Record Portal). Notably, the research conducted by Sharit et al. (2014) used the Berlin Numeracy Test (BNT) to establish a veteran's risk numeracy abilities. The results indicated that higher task performance could be differentiated from a lower performance by age, health literacy, graph literacy, and BNT scores (Sharit et al., 2014). They found that veterans with little computer proficiency skills whom are older and had results reflecting low health literacy, graph literacy, and BNT scores are at a disadvantage and are subject to miss out on the health management benefits that are provided by the MHV system (Sharit et al., 2014).

Research on risk decision making has highlighted the importance of user's statistical numeracy abilities (e.g., comparing and transforming proportions and probabilities) for making informed and accurate risk decisions regarding numerical and non-numerical information (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012; Lipkus, Samsa, & Rimer, 2001; Schwartz, Woloshin, Black, & Welch, 1997). According to Cokely et al. (2012), the BNT was created to measure risk literacy, which they defined as "the range of statistical numeracy skill that is important for accurately interpreting and acting on information about risk" (p. 37). Cokely et al. (2012) presented 21 research studies that were conducted across 15 countries to assess the convergent, criterion, and discriminant validity of the BNT. One study compared the BNT to other numeracy tests developed by Schwartz et al. (1997) and Lipkus et al. (2001) to establish the BNTs criterion validity. The results demonstrated when holding constant the two alternate

strongest predictors of performance (e.g., cognitive reflection and fluid intelligence), the BNT significantly predicted the additional unique variance in risk understanding (Cokely et al., 2012). The two other numeracy tests (e.g., Schwartz et al., 1997 and Lipkus et al., 2001) lost their risk literacy predictive power when the cognitive intelligence or reflection tests were added to a hierarchical regression model (Cokely et al., 2012). The results suggested that the BNT is a psychometrically sound instrument that quickly assesses risk literacy and statistical numeracy abilities to help distinguish between educated individuals (Cokely et al., 2012, 2018).

The BNT has broad predictive power regarding individual performance differences for both numerical and non-numerical risk decision tasks (Cokely et al., 2012, 2018). Cokely et al. (2018) rationalized that statistical numeracy tests like the BNT are robust because, "effective decision making in our complex and uncertain world often requires the same kinds of reasoning and metacognitive skills (e.g., evaluating thinking, feelings, and risks) that are used when solving various practical probabilistic math problems" (p. 479). Effective, naturalistic decision making is primarily about using personally meaningful and practical inductive reasoning and self-regulation for contemplating risk and uncertainty (Cokely et al., 2018). The cognitive processes required for numeracy test questions are more analogous to the processes used for effective naturalistic risk decision making when compared to the mental processes utilized for other cognitive ability tests (Cokely et al., 2018).

The current study used the BNT to establish the relations between the end-user's general risk literacy and their perception of the security provided by different 2FA recovery methods. Research conducted by Nurse, Creese, Goldsmith, and Lamberts (2011) suggested that individuals with high-numeracy levels are more likely to attend to risk information while individuals with low-numeracy levels are more likely to draw on expert guidance and emotions.

It is common for researchers to use the BNT as a general risk literacy tool in studies that provide domain-specific numbers or probabilities that can be associated with risk (Sharit et al., 2014; Woller-Carter, Okan, Cokely, & Garcia-Retamero, 2012). In the context of authentication, it is not common for companies to provide the probability of a successful attack for a given scheme. Instead of this, it is essential to establish how the risk information for the different 2FA account recovery methods should be communicated in the current study.

1.2.2 Communication of Cybersecurity Risks

It is crucial to consider the best practices for communicating cybersecurity risk information to promote informed judgment. Nurse et al. (2011) examined the methods used for effectively expressing cybersecurity risks and proposed a list of recommendations. Nurse et al. (2011) noted that the use of visual aids is a popular method for formatting risk-communication information. They emphasized that the effectiveness of a visual varies based on the context of the application. Each risk communication presentation format (e.g., textual, visual, numeric) has unique strengths and weaknesses in facilitating productive risk communication information (Chen et al., 2018; Nurse et al., 2011). If the research situation allows, a combination of different presentation formats can also be used (Chen et al., 2018; Nurse et al., 2011). When communicating risks, Nurse et al. (2011) suggested presenting non-cybersecurity experts with a limited amount of security details to keep the communication simple and reduce cognitive load. At the same time, researchers should provide representative information when the risk perception task requires users to make educated judgments (Almuhimedi et al., 2015; Chen et al., 2018; Nurse et al., 2011).

The current research hoped to shed light on how risk literacy relates to cybersecurity risk perception of different 2FA account recovery methods. One section of our study provided users with textual and visual information for each 2FA account recovery method. The goal was to provide enough information about each method that would allow users with knowledge about the different cybersecurity authentication attack vectors (e.g., brute force, intersection) to accurately determine how secure a given method is in comparison to the other methods. For example, each recovery method description contained information regarding the number of password characters (or passcode images) that need to be input in order to successfully authenticate (see Appendix A – section 5). Tentatively, if the participant understands passwords or passcodes that are short and/or lack complexity are more susceptible to brute force attacks, they should have been able to determine which method is most secure by comparison. Combining all the attack vector information that could have been deducted for the 2FA account recovery textual and visual excerpts (c.f., password length, password can be easily recorded, an attacker would only need one observation to obtain the full password) should have allowed for users to correctly rank order the methods in terms of their security.

The current research had users rank the 2FA account recovery methods in terms of security. This task allowed us to indirectly test whether users understood the different attack vectors. This was an integral part of the study because if a participant does well on the BNT, but not well on the comprehension and ranking tasks, it is probable they may not know enough about cybersecurity authentication information. The results will guide suggestions for future efforts regarding the best ways of presenting information to improve cybersecurity authentication comprehension and authentication selection.

Theoretically, if users have high-risk literacy, are aware of the authentication risks, and practice good authentication hygiene, then cyber-attack resistance potentially increases. Opposingly, when users practice poor password hygiene (e.g., reusing passwords across multiple accounts, creating easily cracked passwords), they are inherently more susceptible to cyber-attacks (Ur et al., 2015). The risks associated with poor password habits are further compound by events like a hacker exposing a company's password database. According to McCandless (2019), there has been a large number of reported database breaches where passwords for system users were leaked. With the increased risk of having authentication credentials compromised due to massive data breaches, more organizations are offering two-factor authentication (McCandless, 2019).

1.3 Two-Factor Authentication (2FA)

In 2016, the United States Federal Government passed a law titled the Federal Cybersecurity Enhancement Act. The bill specified that agencies with "elevated privilege" personnel accounts were required to use multi-factor authentication for all the accounts with said, "elevated privileges" (Federal Cybersecurity Enhancement Act of 2016). The 2FA authentication process prompts users to provide identity verifying information for two different authentication categories (c.f., something you know, something you have, or something you are). One type of 2FA example is performing a grocery store transaction using the Point Of Sale (POS) Machine, the user inputs a debit card (something they have) and inputs a PIN (something they know). The idea is that each different category of authentication type requires different attacker capabilities and various kinds of attack strategies, which in turn increases the user's account security.

However, 2FA for computer systems is a relatively new concept that many large companies have started to offer or require users to adopt.

1.3.1 Second Factor Mechanisms

Standard practice has the first factor remaining the user's password for operating systems that enable 2FA. The second factor is often something that the user has. For instance, a common 2FA platform is Duo Mobile Application (App), which is essentially an App the user downloads on their device that is also linked to their system account (Duo, 2019). When the user reaches the second-factor stage during authentication for the given account, they can either choose to receive a push notification from the App or have the App generate a one-time code for the user to type-in to authenticate. If the user opts to receive a push notification, the user will be required to approve the login action. If the user selects the "reject" push notification option, the authentication attempt will be blocked (Loveless, 2018).

Another Universal Second Factor (U2F) is in the form of a USB thumb drive, for example, the Yubico Security Key, which is an implementation of Fast Identity Online (FIDO; Srinivas, Balfanz, Tiffany, & Czeskis, 2017). The Yubico security key communicates with the system via a USB stick that has a built-in touch sensor that the user must touch to authenticate (Lang, Czeskis, Balfanz, Schilder, & Srinivas, 2017). A one-button hardware token is also a primary FIDO 2FA device a user might physically possess. After a hardware token is registered to a user's account, pushing the button will generate a time-based (TOTP) or hash-based (HOTP) one-time (6- or 8- digit) passcode that is only valid for a limited amount of time (Goldberg, 2018). Similarly, users can also opt to receive a one-time passcode by downloading 2FA software on to a personal computer. For example, our university offers WinAuth to faculty and

students as an option for obtaining a software-generated code as their second factor (ODU Information Technology Services, 2019). WinAuth is an open-source authenticator software for Windows that generates time-based codes like that of hardware tokens (Mackie, 2017). Despite the extra layer of security provided by 2FA to increase the infrastructure of the system's attack resistance, the 2FA usability research findings revealed a mixture of strengths and weaknesses regarding the 2FA processes and devices (Colnago et al., 2018; Das, Dingman, & Camp, 2018).

1.3.2 User Studies Surrounding 2FA

Das et al. (2018) conducted research that compared user acceptance of the USB Yubico keys (e.g., YubiKey 4, YubiKey 4 Nano, YubiKey 4C, and YubiKey NEO) in a two-part study. They gathered users' usability and acceptability data both before and after interface interactions modifications were made to improve usability. Despite the Yubico improvements, the second study revealed that participants continued to express their belief in password strength alone. They concluded their study with a warning that stated, "Even the best-designed hardware will not be used if the benefits are not apparent" (Das et al., 2018, p. 15).

Colnago et al. (2018) explored the behaviors and opinions of 2FA adoption at Carnegie Mellon University (CMU). The results indicated that users believed it provided their account with more security, and it was reasonably easy to use. However, many noted that 2FA was annoying. Additionally, they found that users' experience with the CMU Duo App led to positive perceptions. The issues associated with the Duo mobile App were also evaluated. Interestingly, the results indicated that users commonly reported problems such as forgetting one's second factor, having it too far away, losing one's phone, having a dead phone battery, having no data connection, and the hardware token desynchronizing. They noted that the frequency of users

experiencing any of the aforementioned problems significantly impacted both the usability and security constructs. The user's perceptions were negatively affected as the frequency of the issues increased. When these problems occurred, users reported consequences such as not being able to do homework and participate in class; not having access to one's email or computer system; not having access to one's dorm or office; and the interruption of a current task.

1.3.3 2FA Account Recovery

In emergencies where 2FA issues prevent the authorized user from gaining account access, an alternate account recovery option that does not require a registered device or downloaded software should be available to users. Essentially, an account recovery authentication option is an account feature that some systems with 2FA make available for users to set up before losing a second-factor device. Some organizations with 2FA (e.g., Reddit, GitHub, and Google) are currently offering precautionary account recovery options (Loveless, 2018; Prins, 2018; Wallen, 2018). Other websites such as Apple, Evernote, Twitter, and Coinbase inform account holders that in the event of a lost second factor, it may take several business days to regain account access (Afonin, 2016; Coinbase Support – Account Management, n.d.; Ravenscraft, 2014). To regain account access to a LinkedIn account when the second factor is unavailable, the user is required to complete a multi-part form and submit a copy of a government-issued ID (Loveless, 2018).

Several recovery options are used by website systems when users need to regain account assess. Loveless (2018) conducted an informal exploratory evaluation of authentication practices for 2FA and 2FA account recovery for several websites (e.g., Facebook, Amazon, Apple ID, GitHub, Reddit, Yahoo, Twitter, LinkedIn, Gmail, Kraken, Live, and Coinbase). The article covered several recovery options that are used by organizations, which include backup email or

phone, backup recovery codes, offsite or downloadable codes, Master Key (passwords), and a valid government ID. None of the companies that were evaluated provided users with all of the recovery options.

One secure account recovery method is a fallback phone number that allows users to specify a phone number where a special access code can be received (Loveless, 2018). Some websites provide users with the option to have the code sent to the emergency contact via text or phone call. Loveless (2018) reported that GitHub was the only website that allowed a fallback phone number to be used as an account recovery option if the 2FA device failed. However, GitHub did not prompt users to enter an emergency contact number instead of their phone number (GitHub Help, 2019). Presumably, if the account holder's phone is unavailable, the mobile app for 2FA cannot be used, and they will be denied account access. If the user sets up their phone number as the fallback number and their phone is not accessible, the fallback recovery system is also rendered useless, and the user would still be without immediate account access.

The Coinbase website only allowed a backup phone number as account recovery if the account holder was using SMS-based 2FA (Loveless, 2018). Alternatively, so it seems, even if a user adds a second phone number to their account as an account recovery option, Coinbase does not provide the user with the option to select which account recovery phone number receives the SMS 2FA recovery code (Coinbase Support – Account Management, n.d.). If the user follows the link "unable to submit a one-time code" the only options given to the user are "I no longer own the phone ending in +x xxx xxx xx01" or "cancel sign-in" (Coinbase Support – Account Management, n.d.). If the user follows the "no longer owns phone" option, the user is advised that account recovery could take 48-72 hours (Coinbase Support – Account Management, n.d.).

Additionally, users are informed that they will need to: recall and input the old phone number associated with the account; provide the system with their new number; send in a valid and current form of ID; and take a selfie to accompany the ID being sent (Coinbase Support – Account Management, n.d.).

An ideal way to implement a fallback phone number as an account recovery option is for the system to instruct the user to provide an emergency contact phone number (e.g., significant other, family member) during their initial setup of 2FA. When this account recovery option is used, the user could be asked to accurately recognize and select the last four digits of the number that corresponds to their emergency contact's phone number (e.g., ***-***-9540) from a set of distractor phone number endings or asked to recall the last four digits from memory. According to Wickelgren and Norman (1966), users are better at performing recognition tasks than recall tasks. They propose that recognition over recall is a major tenet of good design and is explained well by the Strength Theory. This theory highlights the fact that recall and recognition involve the same memory task but proposes that recognition requires a lower threshold of strength. Thus, accomplishing the memory task becomes easier because recognition necessitates less cognitive resources.

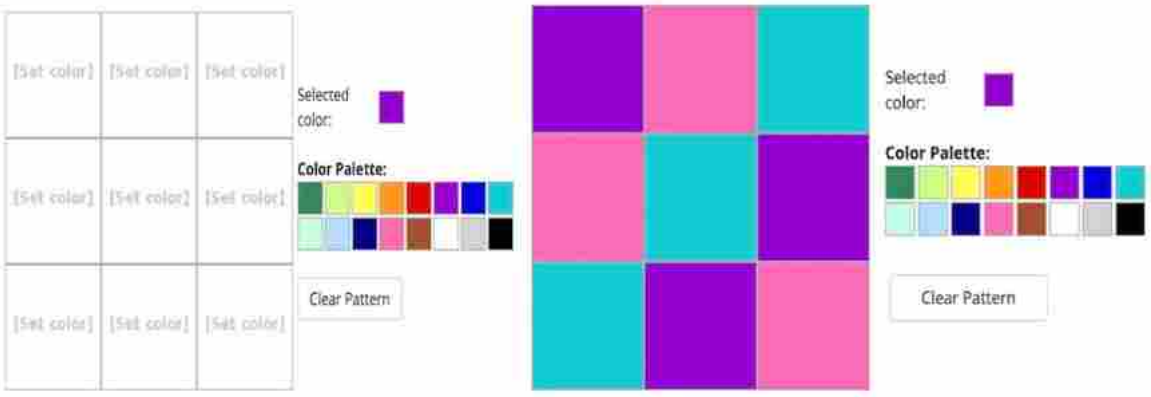
Generated account recovery codes are another secure account recovery method. Account recovery codes are often provided when the user enables 2FA on their account (Loveless, 2018). The recovery codes (e.g., WHZ-23156) are generated by the system and are unique for each user. When recovery codes are provided, it is common for the system to prompt the user to remember the codes; save the codes to an offline server or encrypt the file; or print a hardcopy and store it in a safe place (Loveless, 2018; GitHub Help, 2019). When a user attempts to recover the

account, the system will ask the user to input one of their account recovery codes (GitHub Help, 2019).

A color grid pattern is another type of account recovery method. During the initial setup, users are asked to create a pattern of colors on a 3x3 grid that they will be able to remember and recreate (see Figure 2). At setup, the user chooses the colors they want to use from a system provided color pallet and create a pattern on the grid. The system requires that the user chooses at least two different colors. During account recovery, users are prompted to recreate their color pattern on the grid (MIDAS & Monarch-Key, 2019). Similar to the color grid pattern, another viable account recovery method that could be offered to account holders is a graphical passcode scheme.

Figure 2

Color Grid Pattern Used for Account Recovery



In general, graphical passcodes provide users with a system-generated set of three to five picture icons that are assigned as the user's passcode icons (CHC; Wiedenbeck, Waters, Sobrado, & Birget, 2006; UYI; Hayashi et al., 2008; WYSWYE; Khot et al., 2012; EaN; Tiller et al.,

2019). During the user's initial interaction with the authentication scheme, they are instructed to remember their unique set of icons. For the account recovery, the user will be asked to recognize and select their target passcode icons from a set of distractor icons. Note that the order in which users select their icons during login often does not matter. Theoretically, account recovery options will only be used intermittently. Therefore, it is important that the recovery method that an organization chooses is not only secure but also memorable.

According to Shneiderman and Plaisant (2010), first-time users know the given task concept but lack knowledge of the interface. Expert users are identified as those who seek rapid task completion due to the familiarity of the task and concepts (Shneiderman & Plaisant, 2010). When authentication with a particular scheme is frequent, users can become experts (Shneiderman & Plaisant, 2010). On the other hand, some tasks only require intermittent authentication (e.g., taxes or account recovery). When a system is not used for a long time, the user can forget significant portions of what they knew, thus becoming an intermediate user (Cooper, Reimann, & Cronin, 2007). This is likely to be the case for authentication recovery schemes since there will often be an elapsed period between the initial setup and the actual application of the account recovery method. Therefore, memorability needs to be an inherent feature of the account recovery method that is implemented. The current research deployed a survey question that addressed the memorability of the 2FA recovery methods from a metacognitive perspective.

1.4 Needs Addressed by Current Research Study

The overall research goal was to achieve a better understanding of what factors lead to a more secure account recovery by comparing the four different 2FA account recovery methods. We

examined end-users' risk perception abilities and the authentication choices users make to secure their information. Ultimately, the results aid in determining which 2FA account recovery methods are perceived as secure, memorable, easy to use, and are preferred.

This research looked at descriptive survey information. The survey aimed to evaluate the relations between what typical users know about general- and secure- authentication practices and their perception of account recovery. Individual participant differences and risk literacy were examined. We calculated the opportunistic score for each authentication recovery method to objectively rank the methods according to the level of security they provide. More specifically, we tested four critical hypotheses.

First, we evaluated the Kendall's tau correlations (1938) between the objective opportunistic rankings of the account recovery methods and user's subjective security rankings of the account recovery methods for each participant. We expected that the tau correlation coefficients (r_τ) that result from comparing objective and subjective account recovery security rankings would be less than $r_\tau = .33$, thus suggesting that the user's comprehension of recovery method security is below optimal performance. Secondly, if there was variance in participant's general risk literacy (BNT) scores, we predicted that there would be a positive relationship between user's BNT score and their ability to correctly rank the recovery methods according to the amount of security they provide. For instance, we expected that participants who have a higher BNT score to generate fewer ranking discrepancies when they are asked to order the account recovery methods by their security. For the third hypothesis, we predicted that the average general concept- and threat- authentication knowledge score for individuals who indicate that they have received cybersecurity training would not be significantly different from the average score of those who indicate they have not received training. The fourth hypothesis

predicted that scores regarding general concept- and threat- authentication knowledge would be higher for males.

In conjunction with the four hypotheses, other descriptive survey data were also explored. A general research inquiry examined whether or not the characteristics of experience with -2FA and -account recovery predicted secure account recovery method selection. We assessed whether or not having cybersecurity training impacted participant's ability to rank account recovery methods by security strength. Lastly, we explored the relationship between a participant's ranking ability and their overall authentication knowledge score. The survey allowed us to establish how users perceive account recovery methods.

CHAPTER 2

METHODOLOGY

2.1 Participants

A total of 113 undergraduate students (females = 78, males = 35) were recruited through the Old Dominion University (ODU) SONA Experiment Management System and were compensated with one research credit. After examination of the data inclusion criteria, specified in the following section, data from 10 participants were omitted resulting in a final sample of 103 participants (females = 73, males = 30). Ages ranged from 18 to 50 years (M = 21.50, SD = 6.10). Reported daily computer use ranged from 2 to 23 hours (M = 8.35, SD = 3.99). The number of 2FA devices participants registered to any given account enrolled in 2FA ranged from 1 to 5 (M = 1.58, SD = 0.92). All participants reported corrected to normal vision. See Table 1 and 2 for additional demographic information.

Table 1

Frequency Table for Declared Major

Declared Major	*n*	%
Sciences	44	40.7
Education	21	19.4
Arts and Letters	18	16.7
Health Science	15	13.9
Engineering and Technology	5	4.5
Business	3	2.8
No Major Declared	2	1.9

Note. N = 103.

Table 2

Frequency Table for Type of Electronics Commonly Used

Electronic Type	*n*	%
Desktop Computer	41	39.8
Laptop Computer	100	97.1
Smartphone	101	98.1
Smartwatch	39	37.9
Tablet	23	22.3
Other – Gaming System	4	3.9

Note. Participants were given the option to select more than one electronic type, which resulted in the collection of 308 responses. The percentage represents the proportion of participants that reported typically using a given technology ($N = 103$).

2.2 Materials and Procedure

This research used a 42-question survey that took participants approximately 35 minutes to complete. Previous research has noted that self-report is a valid measure for the topics that are covered by the survey of the current study (Cain et al., 2018; Russell, Weems, Ahmed, & Richard III, 2017). According to Russell et al. (2017), when users do not behave securely, the reports of their non-secure behaviors still result in honest reporting.

Participants viewed an informed consent document and were asked to accept the terms of the study before participating (see Appendix B). Participants were instructed to read the questions and statements thoroughly and to provide an honest and accurate answer. Participants were encouraged to select an answer for all questions. For instance, if an item went unanswered, the system alerted and prompted them to choose an answer before proceeding. However, it was not mandatory for the participants to select an answer on any survey question.

The survey itself consisted of five sections of questions: demographics, general knowledge of authentication concepts, knowledge of authentication threats, the Berlin Numeracy Test, and perception of different account recovery methods (see Appendix A).

2.2.1 Demographics Section 1

The survey asked participants 18 demographic questions to get a better description of the sample. The demographic questions consisted of Yes/No, fill-in-the-blank, check all that apply, and multiple-choice questions. Questions one, two, four, and six through 12 were directly from the cyber hygiene article produced by Cain et al. (2018). Question three was a self-report item regarding a participant's eyesight. Item five addressed participants' declared college major. Question 13 was the attention check question, "Is your heart beating?". The last five questions were created to reflect participants' experience with 2FA and account recovery.

2.2.2 Knowledge of Authentication: General Concepts Section 2 and Threats Section 3

To develop the content for the general concepts- and threats- authentication knowledge sections 2 and 3, we chose seven questions (two general and five threats) from the Cain et al. (2018) article that related to authentication practices. Cain et al. (2018) noted that their survey was developed based on government website topics that highlight the best cybersecurity practices and referred to previous literature that also evaluated these topics. The knowledge of general authentication concepts, section 2, consisted of three questions about authentication security concepts and one attention check question. The first two questions were from the cyber hygiene survey (Cain et al., 2018); these questions focused on capturing the participants' knowledge of common authentication terminology. The new third question was created to assess whether users

understand the operational definition of authentication (e.g., the process of verifying the identity of the user). The fourth inquiry was an instructed attention check, “We care about data quality, to ensure you are currently paying attention can you please select the color option “Yellow”?”. All four questions in the authentication concepts section were multiple-choice and had four possible answer choices, with only one answer being correct.

The knowledge of authentication threats, section 3, provided seven statements, which included one attention-check prompt. The first five knowledge of threats statements were from the cyber hygiene article (Cain et al., 2018); these statements focused on capturing the participants' knowledge regarding common threats, behaviors, or outcomes associated with secure authentication practices. An example of a threats statement is, “It is safe to share a password with others”. Participants were asked to choose from the choices Strongly Agree, Agree, Neither Agree nor Disagree, Disagree, and Strongly Disagree. The sixth statement was the attention check statement, “Please respond with Strongly Agree to this question.” Renaud’s (2007) opportunistic score prompts inspired the last statement. Specifically, the statement focused on addressing participants' knowledge about the threat associated with authentication error messages.

2.2.3 Berlin Numeracy Test Section 4

The numeracy test, section 4, consists of the full seven question Berlin Numeracy Test (BNT; Cokely et al., 2012; Schwartz et al., 1997). The traditional BNT consists of four relatively difficult items suited for moderate-to-highly numerate individuals (Cokely et al., 2012). However, to provide additional discriminability for low-to-moderate numerate individuals, it is suggested that the 3 relatively easy Schwartz et al. (1997) items be added to the 4-item BNT

(Cokely, Ghazal, & Garcia-Retamero, 2014; Petrova et al., 2019). Previous research has shown that using the combination of both easy and challenging items shows better discriminability than using the tests alone (Petrova et al., 2019). The numeracy test was used as a predictor of the participant's comprehension of everyday risk. Previous research indicates that the numeracy test has adequate internal consistency with Cronbach α scores ranging from .70 to .75 (Cokely et al., 2012).

For this section of the survey, participants are asked to fill-in-the-blank with the correct answer with only one answer being correct. A participant's risk literacy score was calculated by counting the number of correct answers given and summing of all the participant's correct answers (possible range of 0 – 7).

2.2.4 Perception Section 5

Before inquiring about account recovery method perceptions, the participants were given images and a short description for each of the four different account recovery methods (e.g., fallback phone numbers, generated account recovery codes, graphical passcodes, and color grid pattern; see Appendix A – section 5). When account recovery methods are applied in real situations, typically, there is not any information provided about the different methods. However, some methods may have been unfamiliar to participants, so a brief description of each method was provided. Also, the short descriptions provided participants with information hinting at the security of the given method.

To establish whether or not participants attended to the account recovery descriptions, the following Yes/No question was added:

> It is vital to our research that we only include responses from people that devoted their attention to this study. In your honest opinion, should we use your data for *this section* in our analyses? (You will receive credit for this study even if you provide a negative response)

Several studies have noted that participants will honestly report whether or not they think their data should be included (Brühlmann & Mekler, 2018; Cunningham, Godinho, Kushnir, & Bertholet, 2017; Curran, 2016; Meade & Craig, 2012). The results of Meade and Craig's (2017) study indicated that this question, in combination with the other attention check questions, is a good strategy for weeding out careless responders. We found a moderate rate of participants that responded with "No" do not include their data ($N = 25$). We used this question as a quasi-experimental variable to evaluate the raw data for each survey question to compare participants that indicated "Yes" to those who selected "No". Significant differences between the "Yes" and "No" groups were found for the BNT score variable. This group difference is examined more closely in the statistical analyses results section that included the BNT score as a variable. The raw data for the groups did not significantly differ for any other variable evaluated in the results section.

In the perception section 5, participants were asked to rank order the recovery methods for three different perceived qualities: security, memorability, and ease of use. The last two survey questions regarded account recovery method preferences. First, participants were asked to choose their preferred account recovery method. The last question was open-ended and asked participants, "Why would you prefer to use this recovery method?" Upon completing the survey, participants were thanked and provided with a debriefing statement (see Appendix A).

2.2.5 Participant Exclusion Criteria

Each survey section had a different level of importance to the study. Therefore, missing data cases were handled uniquely for each section to help with data preservation. To further secure the integrity of our data, the survey included three instructional attention check prompts. If a participant responded incorrectly to two or more of the attention checks, their data were omitted from the data analysis ($N = 1$).

The BNT, section 4, was a critical study variable. If a participant failed to answer all seven numeracy questions, their data were omitted for the data analysis ($N = 1$). Additionally, it is possible that participants might have been inclined to cheat and use the internet as a source for answers because this study was conducted online. If this occurred, the validity of the BNT scores would be threatened. To account for this threat, the average amount of time spent on the BNT and the standard deviation were evaluated ($M = 389.49$s, $SD = 351.59$s). Data from 5 participants suggested a response time that was more or less than two standard deviations from the mean, which resulted in omitting their data from all analyses.

Participants' perceived security of the recovery methods, in section 5, was also a critical study component. If a participant failed to answer the security ranking question, their data were omitted from all data analyses ($N = 3$). These data cleaning criteria resulted in the exclusion of 10 participants from all data analyses.

There were two additionally cases of missing data for the two nonessential study variables of ranking recovery methods by ease of use ($N = 1$) and memorability ($N = 1$); however, these two participants were not excluded from any data analyses.

CHAPTER 3

DATA CODING STRATEGIES

The following sections detail the processes used to code the data in preparation for statistical analysis. We highlight the process used to calculate a participant's overall authentication knowledge score. We establish the coding strategies used to determine the objective security ranking of the recovery methods by detailing opportunistic score calculations. Additionally, we walk through the process used to assess participants' ability to correctly rank order the recovery methods in terms of their security. The last section consists of the strategies used to categorize and code the participant's account recovery method preference reasoning.

3.1 Overall Authentication Knowledge Score

The scores for the concept- and threat- authentication knowledge sections were combined to create an overall knowledge score. The three concept authentication knowledge questions only had one correct answer, and they were scored dichotomously. In order to combine the concept- with the threat- authentication knowledge scores, the 5-point Likert scale score used for the six questions in the threat section was changed to a binary score. Depending on the question, due to the reverse coding of some of items, participant answers that were marked Strongly Disagree or Disagree were scored as incorrect or correct responses and the answers of Strongly Agree or Agree were scored as incorrect or correct responses (see Appendix A – section 4). The response Neither Agree nor Disagree was scored as incorrect for all questions. The number of correct concept- and threat- authentication knowledge responses were summed to determine a participant's overall authentication knowledge score (possible range of 0 – 9).

Cain et al. (2018) performed a similar type of data coding strategy wherein they combined their concepts and threat knowledge sections. Cain et al. (2018) reported a Cronbach α of .50 for the concept knowledge questions section and Cronbach α of .78 when the concept and threat knowledge questions sections were combined. The current study found a Cronbach α of .60 after combining the concept and threat authentication knowledge items in order to obtain participants' overall authentication knowledge scores. Notably, Cain et al. (2018) survey covered a wider range of cybersecurity topics and deployed a total of 34 threat and concept knowledge questions. The current study only employed a subset of their questions that pertained to the topic of authentication security (e.g., two concept and five threat questions). This study's low number of concept and threat knowledge questions (e.g., nine total) might have impacted the Cronbach α we found.

3.2 Opportunistic Score Calculations

Renaud (2007) provides a detailed criterion for calculating the opportunistic score for a given authentication scheme. Following the prompts and scoring guide for each opportunistic category, a score for each recovery method was determined (Renaud, 2007). The number of prompts for each category is as follows: one guessability, one recordability, two observability, eight analyzability, and six resistibility prompts. It is important to note that only seven out of the 18 opportunistic prompts regard the security protocols of the authentication or recovery scheme. The other 11 prompts regard the authentication security protocol of the system itself. Specifically, this was the case for six analyzability prompts (e.g., a – f) and five resistibility prompts (e.g., a & c – f; see Appendix C). For the 11 prompts that inquired about the authentication security of the system, the same security score was consistently assigned for all

the recovery methods. The score for each of the 11 system prompts was assigned according to the prompt score that suggested the system was implementing the best attack resistant strategy. For example, the "(a) Usernames" prompt for analyzability states:

> "Users should be assigned usernames rather than email addresses because email addresses are too easily obtainable and make it easier for hackers to gain access to the system. A 1 is assigned if the system uses email addresses as usernames and a 0 if unique usernames are used and not visible to other users" (Renaud, 2007, p. 18).

This prompt suggests that systems that assign usernames are more resistant to attacks, thus indicating that each recovery method should receive a score of 0 for this prompt.

The other seven opportunistic prompts that considered the security of the given method rely on the designer's interpretation. Specifically, this was the case for the: guessability prompt, recordability prompt, observability prompts, two analyzability prompts (e.g., g & h), and one resistibility prompt (e.g., b; see Appendix C). For example, the recordability prompt states:

> "As regards recordability, the systems can be assigned values as follows: 1 if the code is easily recorded; 0.5 if it was harder to record or describe, or if recording of the key does not provide an observer with the full key; 0 if it is difficult or impossible to record or describe, such as, for example, a biometric" (Renaud, 2007, p. 17).

The term "easily recorded" is not operationally defined, suggesting that designers are to determine for themselves what they consider to be an easily recordable interaction.

In general, to calculate an overall opportunistic score for each account recovery method, the score for each prompt in the guessability, recordability, observability, and analyzability categories are added together and divided by the sum of the prompts for the recordability category. Lower opportunistic scores indicate higher authentication attack resistance.

For the current study, the principal investigator and four other research assistants rated each recovery method on the seven subjective prompts to avoid bias in the interpretation of the Renaud's (2007) opportunistic score instructions. The research assistants were given task instructions, the Renaud (2007) article to read, and the same account recovery descriptions that were provided in this study's survey (see Appendix A – section 5). The researchers received an excel sheet with all 18 opportunistic prompts with score guides (see Appendix C). The 11 system prompts were pre-scored, and the seven remaining prompts had empty cells for scoring each recovery method. To allow for a more detailed discussion about rating discrepancies, research assistants were asked to explain why they assigned values to the given opportunist category prompts. The researcher's ratings were assessed on a descriptive level to evaluate the recovery method ranking orders. The researchers discussed and attempted to resolve any coding discrepancies, but consistent ratings could not be achieved. However, four out of the five researchers found that the total opportunistic scores for the fallback phone number and account recovery codes suggested they were ranked either 1st or 2nd most secure. Whereas the color grid pattern and graphical passcode scores suggested they ranked as less secure in either the 3rd or 4th position. In lieu of this, the total opportunistic score for each recovery method was averaged across raters to find the final objective recovery method rank order.

The averaged opportunistic score calculations resulted in the account recovery methods ranked by most attack resistance as follows: 1st fallback phone number, 2nd generated account recovery codes, 3rd color grid pattern, and 4th graphical passcode (see Table 3).

Table 3

Total Opportunistic Score for Each Account Recovery Method by Researcher Assistants

Researcher	Account Recovery Methods – Total Opportunistic Score			
	Fallback Phone Number	Generated Recovery Codes	Color Grid Pattern	Graphical Passcode
Principal Investigator	0.125	0.214	0.286	0.25
Research Assistant 1	0.139	0.219	0.219	0.563
Research Assistant 2	0.071	0.214	0.357	0.149
Research Assistant 3	0.214	0.179	0.429	0.464
Research Assistant 4	0.214	0.286	0.179	0.179
Average Score	0.153	0.222	0.294	0.320

Note. The average score row reflects the average opportunistic score for a given recovery method when total scores were averaged across raters.

3.3 Coding Strategy for User's Subjective Security Ranking of Account Recovery Methods

The survey question that instructed participants to, "rank order the account recovery methods from safest (1) to least safe (4)", was coded using Kendall's tau rank method to derive a ranking distance score and tau correlation coefficient (Kendall, 1938). The tau rank correlation is a non-parametric statistical test that was used to establish the degree of agreement between the objective opportunistic score ranking and the participant's subjective security ranking of the account recovery methods (Kendall, 1938). As a comparison, Spearman's rho (ρ; 1904) rank correlation focuses on rank distances, Kendall's tau correlation (τ; 1938) is used for evaluating rank differences. The opportunistic score results suggested that the target rank order is: (1) fallback phone number, (2) account recovery codes, (3) color grid pattern, and (4) graphical passcode. We used Kendall's tau distance measure to find the relative order difference between

the two rank orders (e.g., objective opportunistic ranks and participant's subjective ranking). To calculate the distance, we compared each given positional rank to the ranks that proceed it. A score of +1 was assigned when the relative order of the pair was correct (e.g., concordant pair), and a score of -1 was assigned if the relative order was incorrect (e.g., discordant pair). Table 4 provides an example for a participant that ranked the account recovery methods as 2314.

Table 4

Kendall's (1938) Tau Distance Score for a Participant that Ranked the Methods as 2314

Rank Position				Ranking Distances	Tau Distance Score	Tau Correlation Coefficient
2	**3** = +1	**1** = -1	**4** = +1	+1	2	r_τ = .33
3	**1** = -1	**4** = +1	-	0		
1	**4** = +1	-	-	+1		

Note. N = 89.

The tau rank correlation uses Kendall's tau distance score divided by the number that represents a "perfect" score (see Figure 3). The n in the equation represents the number of forced-rank positions (e.g., n = 4). For the current research, we found that 6 represents a "perfect" score for the formula denominator. A tau correlation coefficient will be calculated for each participant and will result in a coefficient between 1 to -1. A correlation coefficient of 1 would result from a perfect match between a participant's ranking and the objective opportunistic ranking. A completely independent relationship between the ranks would yield a 0, and the worst possible relation between the ranks would yield a -1. For the example of the participant that ranked the account recovery methods as 2314, we would find the number of

concordant pairs (4) minus the number of discordant (2) pairs equals 2. The tau correlation of this example would yield a coefficient of $r_\tau = (2/6) = .33$.

Figure 3

Kendall's (1938) Tau Correlation Coefficient Formula

$$r_\tau = \frac{(number\ of\ concordant\ pairs) - (number\ of\ discordant\ pairs)}{\frac{1}{2}n(n-1)}$$

It is important to note that the interpretation of the strength of Kendall's tau rank correlation coefficient values is not standardized (Bachmann & Bernstein, 2010). Bachmann and Bernstein (2010) specified that Kendall's tau typically generates lower correlation values and cannot be meaningfully compared to other rank correlation values. When Kendall's tau rank correlation (1938) is used in circumstances where there are only four ranking positions, there are only 7 possible tau correlation coefficient values that can result (e.g., 1, 0.66, 0.33, 0, -0.33, -0.66, -1). For this research, we rationalize for the strength of the tau correlation coefficient values to be interpreted as follows: 0 to +/- .33 as a weak correlation; -0.33 to -0.66 as a moderate negative correlation; 0.33 to 0.66 as a moderate positive correlation; -0.66 to -1 as a strong negative correlation; and 0.66 to 1 as a strong positive correlation. Interpreting the strength of the correlations in this way allowed us to determine the threshold that indicates users are comprehending the security provided by the recovery methods. The correlation strength parameters are used for hypothesis one, which predicts that the correlation between the objective security ranking and participants' perceived security will be less than .33 on average. This

suggests the objective versus subjective rankings have less than a moderate positive correlation on average.

3.4 Categorization and Coding Strategy for User's Account Recovery Method Preference Reasoning

The last two survey items regard participant's account recovery method preference. First, participants were asked which recovery method they would choose to use. Then they were asked to explain why. Participant's preference reasoning data were used to explore what type of account recovery method attributes lead to method choice. The author and a research assistant used Glaser and Strauss's (1967) Grounded Theory strategies to code participant's account recovery method preference reasoning into nominal scale categories. Grounded Theory is a systematic approach for conducting qualitative data research wherein research questions are established a posteriori based on the specific research data that is collected (Glaser & Strauss, 1967).

The researchers conducted a thematic research analysis by independently reading each participant's annotative reasoning and creating a list of category themes based on the observed responses. The researchers discussed the themes and merged the findings to create a consolidated category coding list. The final coding list resulted in 11 categories, wherein 4 categories were assigned subcategories. In sum, there were 29 possible categories and subcategories used to code responses for each recovery method. Notably, participant's answers were fairly straight forward; thus, the majority of the 29 subcategories represent the explicit words that participants used to describe a given method. The 11 main categories were created to umbrella the subcategory descriptors using human factors terminology. For example, the term "High Usability" is often

used to indicate that an interaction or product possess qualities such as being, “easy to use”, “efficient”, “effective”, “reliable”, “convenient”, etc. (Hornbaek, 2006; Mator et al., 2020). For a more detailed report on the operational definitions of the categories and subcategories, see Table 5. The researchers noticed that in addition to describing a preferred method, participants often addressed aspects of other recovery methods. To maximize the breadth of findings, each narrative was coded to capture the details of all recovery methods referenced during the second round of coding.

Both researchers analyzed the participant’s responses independently a second time (Egelman et al., 2014; Kraus, Schmidt, Walch, Schaub, & Möller, 2017). The researchers used the established category and subcategory list to code the participant’s preference responses. Specifically, to code a narrative, the researchers would mark which recovery method was being described and assign the detail to a given category and, when applicable, subcategory. The same two researchers were employed for both rounds of coding to allow for the raters to be well-calibrated on what to look for. This trained rater approach is similar to previous research methods that have been used when evaluating qualitative data in phases (Angeli, 2013; Greenhow, Li, & Mai, 2019).

To compare the coded results, a fully crossed design was employed, and we defined interrater reliability as the propensity for any two human factors researchers to assign an account recovery method annotation to the same category and subcategory (Hallgren, 2012). To determine a beyond chance interrater agreement coefficient, we used Cohen’s (1960) version of Kappa (κ). The Landis and Koch (1977) benchmark scale was used to establish the level of agreement between raters (Alonso, 2013; Hallgren, 2017). Specifically, the scale characterizes the raters beyond chance agreement level according to different range values of κ (Hallgren,

2017; Landis & Koch, 1977). Generally, κ values of .81 or more reflect almost prefect to perfect agreement, values between .60 and .80 typically reflect good agreement, values .41 and .59 typically reflect moderate agreement, and values between .21 and .40 reflect fair agreement and κ values below .20 are interpreted as poor agreement (Fleiss, Levin, & Paik, 2003; Hallgren, 2017; Hornbaek, 2006; Landis & Koch, 1977). The results of our interrater analysis suggested good agreement between raters, $\kappa = .728$, $p < .001$. Before statistically analyzing the data, all coding discrepancies between raters were resolved by discussing until a unanimous agreement was reached ($N = 217$).

Table 5

Categories and Subcategories for Coding Recovery Method Preference Reasoning

Categories	Definition	Subcategories
High Usability	• If the participant referenced a recovery method as having any quality that is expressed in the subcategory list.	• Easy to Use • Simple • Effective • Efficient • Reliable • Requires Less Effort • Convenient • Sufficient • Versatile

Categories	Definition	Subcategories
Low Usability	- If the participant referenced a recovery method as having any quality that is expressed in the subcategory list.	- Difficult - Busy - Less Effective - Less Efficient - Less Reliable
Visual Processing Preference	• If the participant indicated that they are a visual learner or they expressed the method aesthetically pleasing by using descriptions such as, “I like the visual aspect”, “it's pretty”, or “I like colors”.	• Visual Learner • Aesthetically Pleasing
Recovery Method Setup Instructions Might be Disregarded	- If the participant indicated that they would use their phone number or they always have their phone on them, they were classified in the likely misunderstood method description. - If the participant displayed knowledge of the recovery methods and noted that the method's weakness consists of people disregarding the 2FA setup instructions by using their phone number, they were classified as likely understood method description.	- Misunderstood Description - Understood Description
Familiar	• The participant referred to a method as being familiar, common, conventional, traditional, or worked in the past.	
Unfamiliar	- Participants referred to a method as being unfamiliar or uncommon.	

Categories	Definition	Subcategories
Easy to Remember	• Participants implied that the method enhances their ability to more easily recall or remember the associated password or passcode. The participant indicated that they could recall/reference the location of pertinent information. Participants specified they don't have to remember anything. The participant's description suggested a methods ease of use was attributed to its high memorability.	
Difficult to Remember	- The participant implied they would not be able to remember a method's associated password or passcode.	
Secure	• The participant indicated that a given method is safer to use than the others, the safest or most secure, provides good account protection, hard to hack, or the password or passcode cannot be guessed easily.	
Unsecure	- The participant indicated that a given method is not the most secure, is not the safest, it can be easily hacked, or it can be easily accessed.	
Intrusive to Emergency Contact	• The participant indicated that the method elicits an inconvenience, burden, or it can be intrusive to an emergency contact.	

CHAPTER 4

RESULTS

This study asked participants to answer survey questions regarding their general authentication knowledge, knowledge of threats, risk literacy abilities, and perception of different account recovery methods. In the following result sections, we report the descriptive data and statistical findings. All statistical tests used an α level of .05 to indicate the probability of rejecting the null hypothesis.

4.1 Descriptive Data Analyses

4.1.1 Demographics

First, we present the descriptive findings, in text and figures, related to participants' cybersecurity familiarity and experience with -2FA and -2FA account recovery. Twenty percent of participants indicated they had previous exposure to some type of educational cybersecurity material (see Table 6). When the survey data were collected, it was mandatory for all ODU students to have their accounts enrolled in 2FA. However, only 89% of participants indicated that they used 2FA to protect any personal accounts. This may suggest that the conceptual meaning of 2FA might not be apparent to some users. Eighty-one percent of participants indicated they use a smartphone or tablet app as their second factor, and 60% of participants only have one 2FA device enrolled per account ($M = 1.58$, $SD = 0.92$). For more reports on 2FA familiarity, see Table 7 and Figures 4 and 5. Forty-three percent of participants indicated they had previously set up a 2FA account recovery option for cases such as a lost or stolen second factor wherein 56% indicated they set up a secondary email as an account recovery option. For a complete report on 2FA account recovery familiarity, see Table 8 and Figure 6.

Table 6

Frequency Table for Cybersecurity Familiarity

Variable	*n*	%
Received Cybersecurity Training		
Yes	11	10.7
No	92	89.3
Training Location		
Work	8	7.5
School	5	4.7
Online	2	1.9
Other – Military	2	1.9
N/A	89	84.0
Taken a Class with Cybersecurity Topics		
Yes	16	15.5
No	87	84.5
Cybersecurity Expert		
Yes	1	1.0
No	102	99.0
Target of a Cybersecurity Attack		
Yes	17	16.5
No	86	83.5

Note. $N = 103$.

Table 7

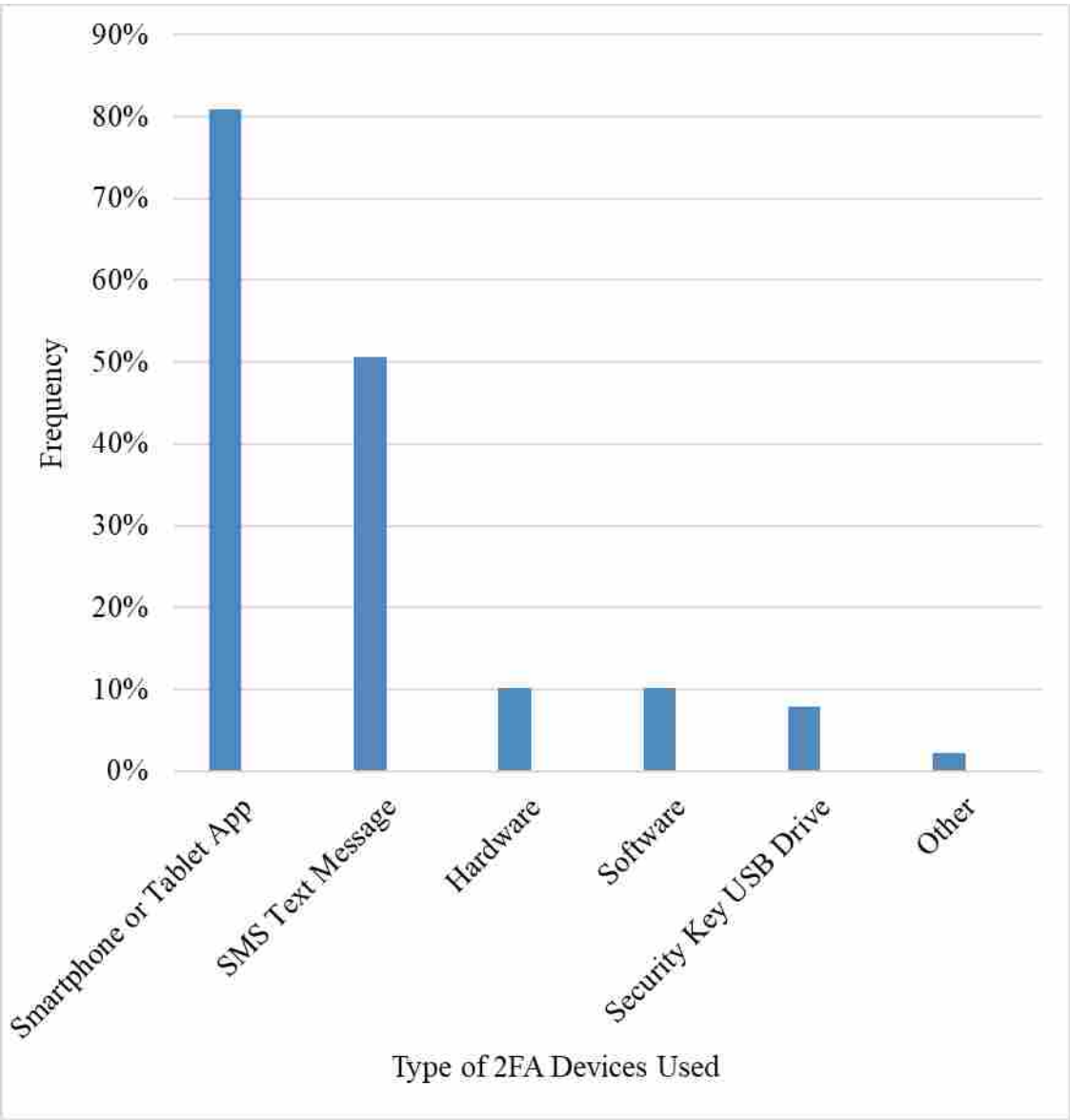
Frequency Table for 2FA Familiarity

Variable	*n*	%
Use 2FA to Protect Personal Accounts		
Yes	89	86.4
No	14	13.6

Note. $N = 103$.

Figure 4

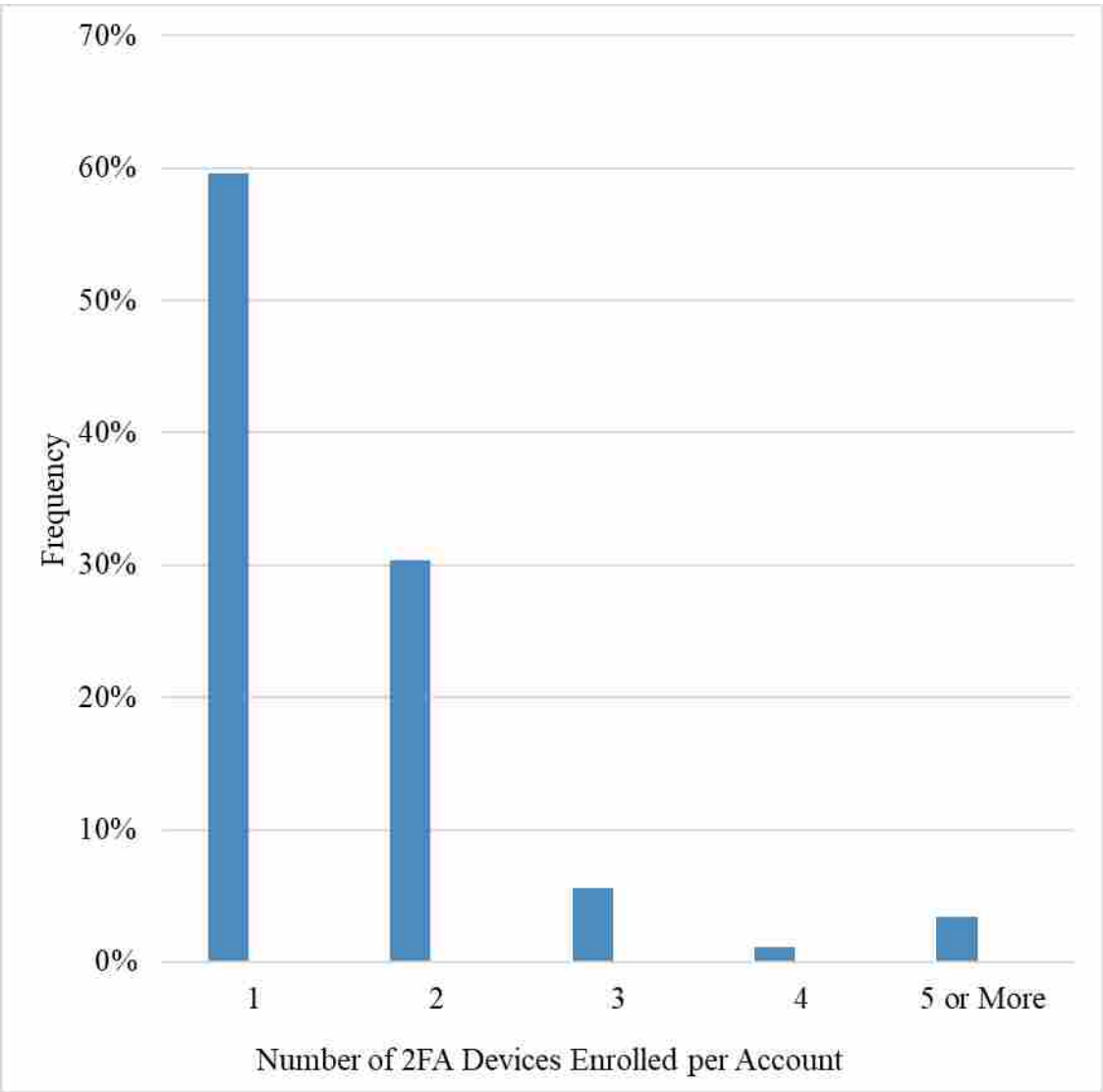
Frequency of 2FA Devices Used by Type



Note. The 89 participants who reported Yes to using 2FA were asked to indicate the type of 2FA devices they had used. Participants had the option to select more than one kind of 2FA device, which resulted in $N = 144$ responses. The percentage represents the proportion of participants out of the 89 participants who reported they use 2FA.

Figure 5

Frequency of the Number of 2FA Devices Enrolled per Account



Note. The percentage represents the proportion of participants out of the 89 participants who reported they use 2FA.

Table 8

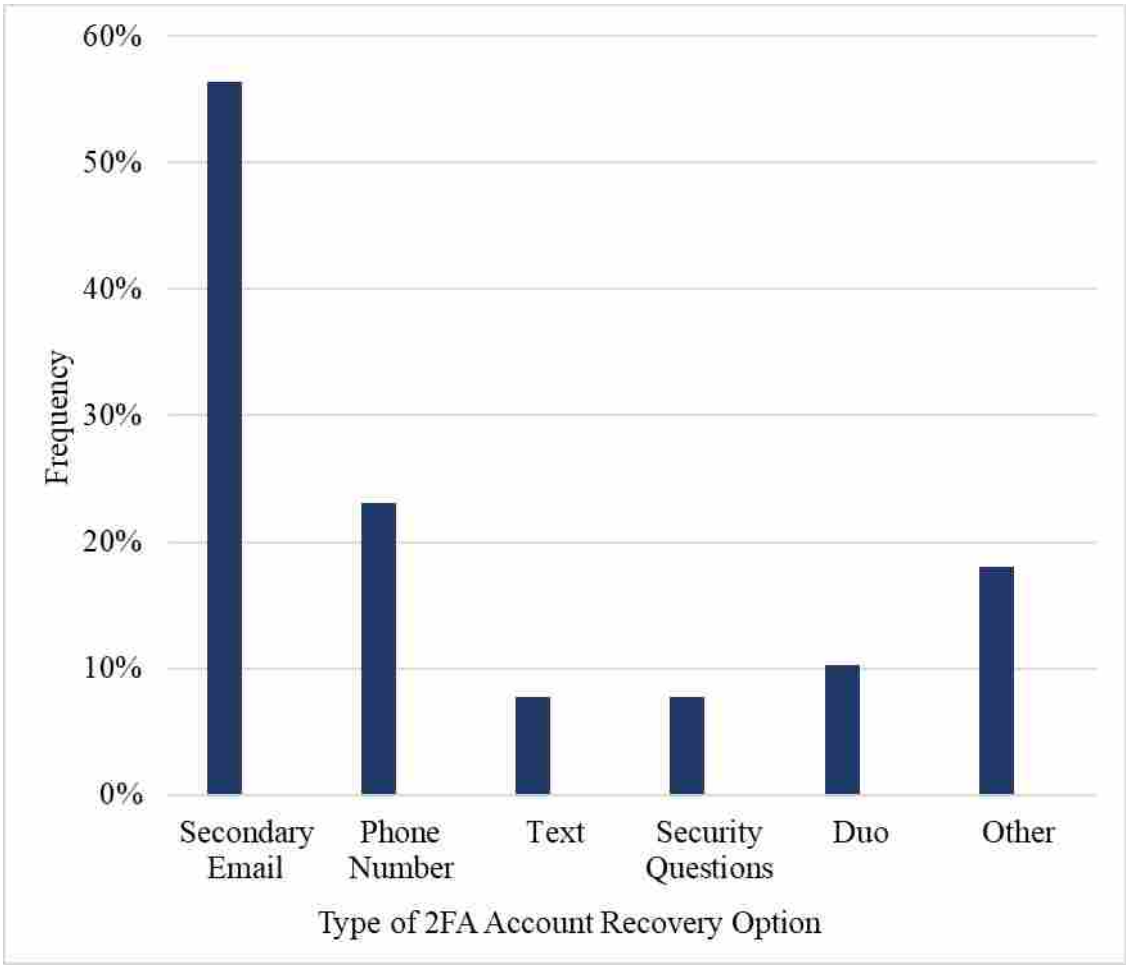
Frequency Table for 2FA Account Recovery Familiarity

Variable	*n*	%
Setup 2FA Account Recovery Option		
Yes	39	43.8
No	50	56.2

Note. $N = 89$.

Figure 6

Frequency of 2FA Account Recovery Methods Used by Type



Note. The 39 participants that reported they had set up 2FA account recovery were asked to indicate the type of option they had setup. Participants had the opportunity to select more than one type of recovery option, which resulted in 48 responses. The percentage represents the proportion of participants out of the 39 participants that reported using 2FA account recovery.

4.1.2 Perception and Preference Qualities Associated with Account Recovery Methods

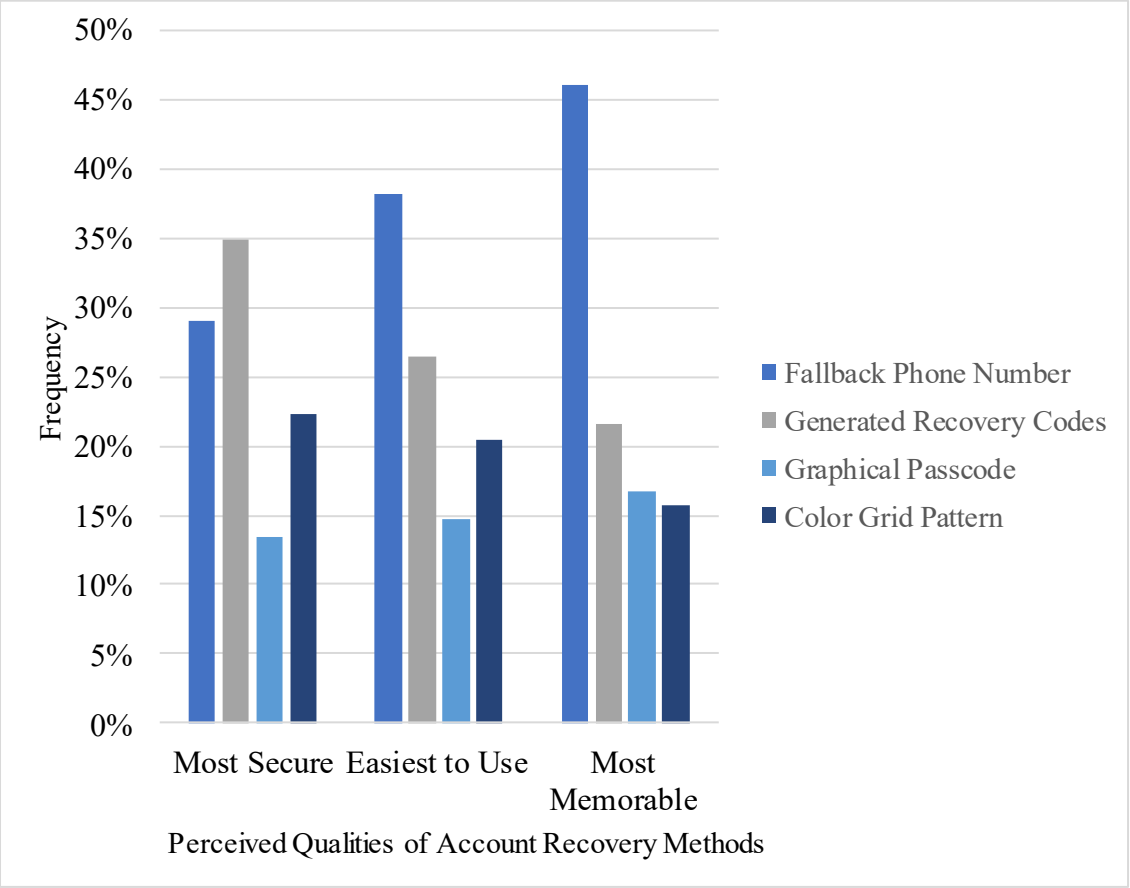
This section presents the percentage of participants that perceived each account recovery method as being most -secure, -memorable, and -easiest to use. We highlight the recovery methods that would be preferred and why. These findings help accentuate characteristics that may aid in determining viable 2FA account recovery solutions.

Thirty-five percent of participants indicated that the generated recovery codes method was the safest. Thirty-eight percent of participants indicated that the fallback phone number method would presumably be the easiest to use and 46% of participants indicated that it would be the most memorable. For a more detailed report on recovery method perception qualities, see Figure 7.

Of the given recovery method options, 45% of participants indicated that they would choose the fallback phone number as their account recovery method. For a more detailed report on recovery method preference, see Figure 8.

Figure 7

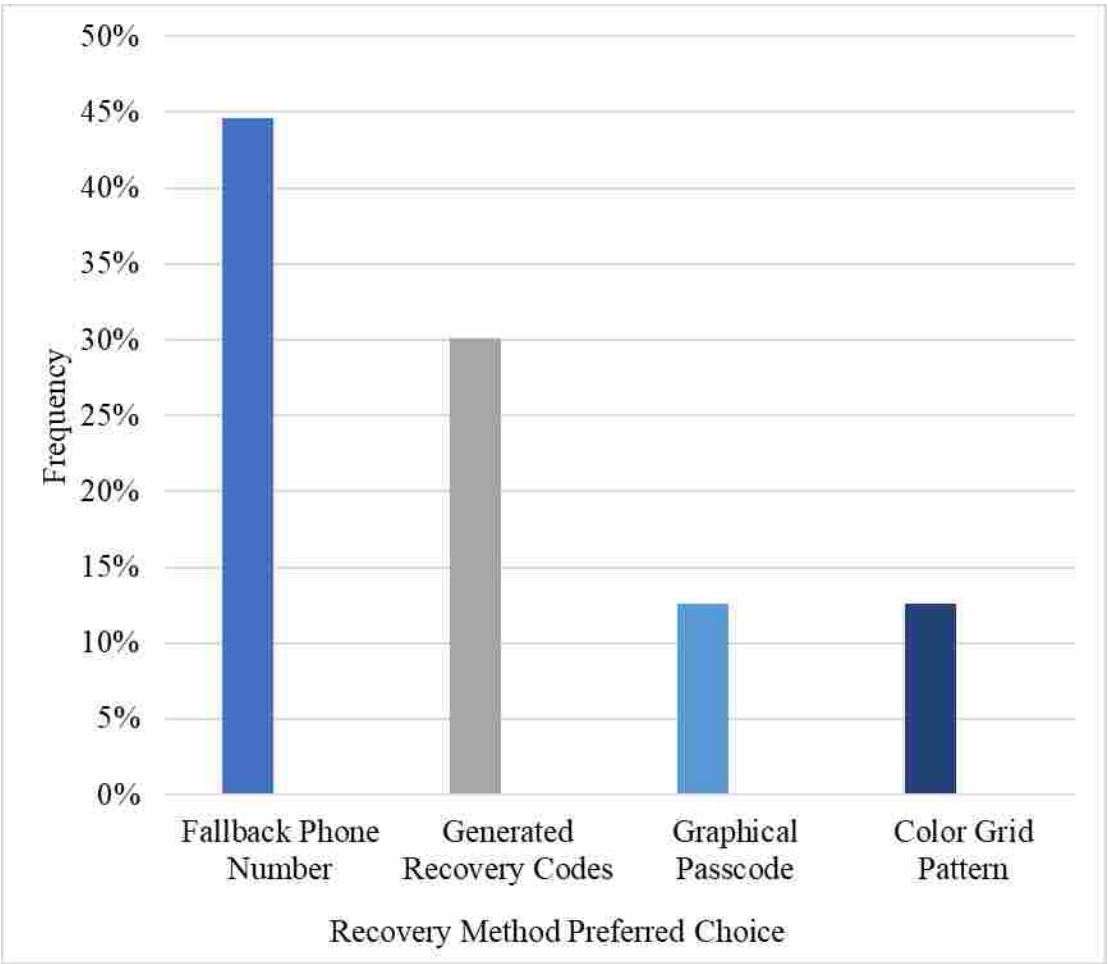
Frequency of Perceived Qualities Related to Account Recovery Methods



Note. Most Secure ($N = 103$), Easiest to Use ($N = 102$), and Most Memorable ($N = 102$).

Figure 8

Frequency of Recovery Method Preferred Choice



Note. $N = 103$.

The last question of the survey invited participants to explain why they preferred a given recovery method. The 103 participants provided a total of 217 open-ended descriptions expressing account recovery method preferences and opinions. Specifically, 95 responses mentioned the fallback phone number, 53 the generated recovery codes, 36 the color grid pattern, and 33 the graphical passcode.

We further evaluated the frequency of category mentions by collapsing across the recovery methods. We found that a majority of the 217 responses, 34%, regarded the high usability aspects. The second most referenced category was easy to remember, captured by 13% of the responses. For a more detailed report on the frequency of category descriptions, see Table 9. Figure 9 depicts the relative mentions of a category concerning the account recovery method being described. As previously noted, 4 categories consisted of subcategories. See Figures 10 through 12 for more reports on subcategory mentions with respect to the given account recovery method.

Table 9

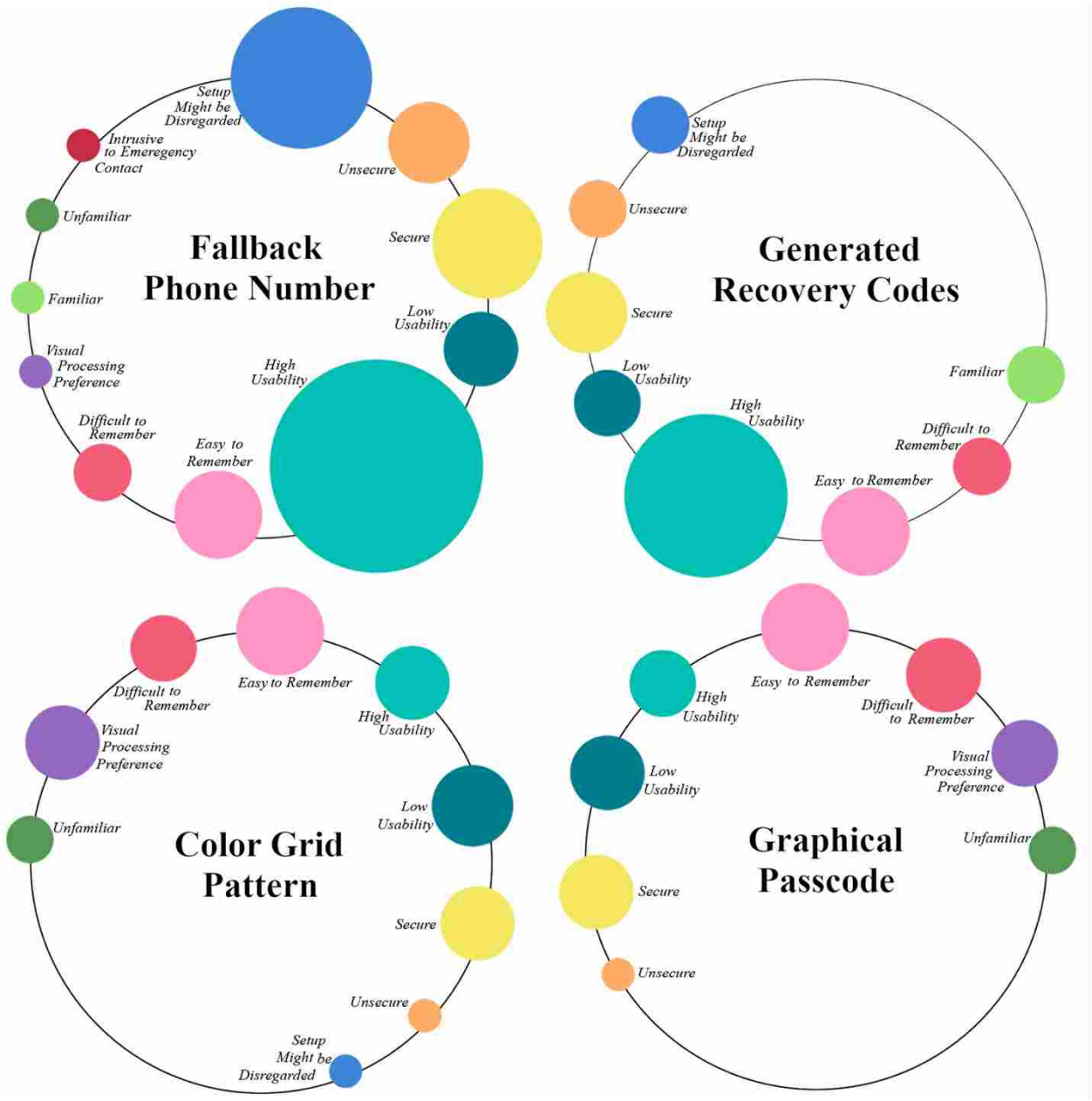
Frequency of Categories Mentioned

Category	*n*	%
High Usability	74	34.1
Low Usability	20	9.2
Easy to Remember	28	12.9
Difficult to Remember	15	6.9
Secure	27	12.4
Unsecure	11	5.1
Familiar	4	1.8
Unfamiliar	5	2.3
Visual Processing Preference	10	4.6
Setup Instructions Might be Disregarded	22	10.1
Intrusive to Emergency Contact	1	0.5

Note. $N = 217$.

Figure 9

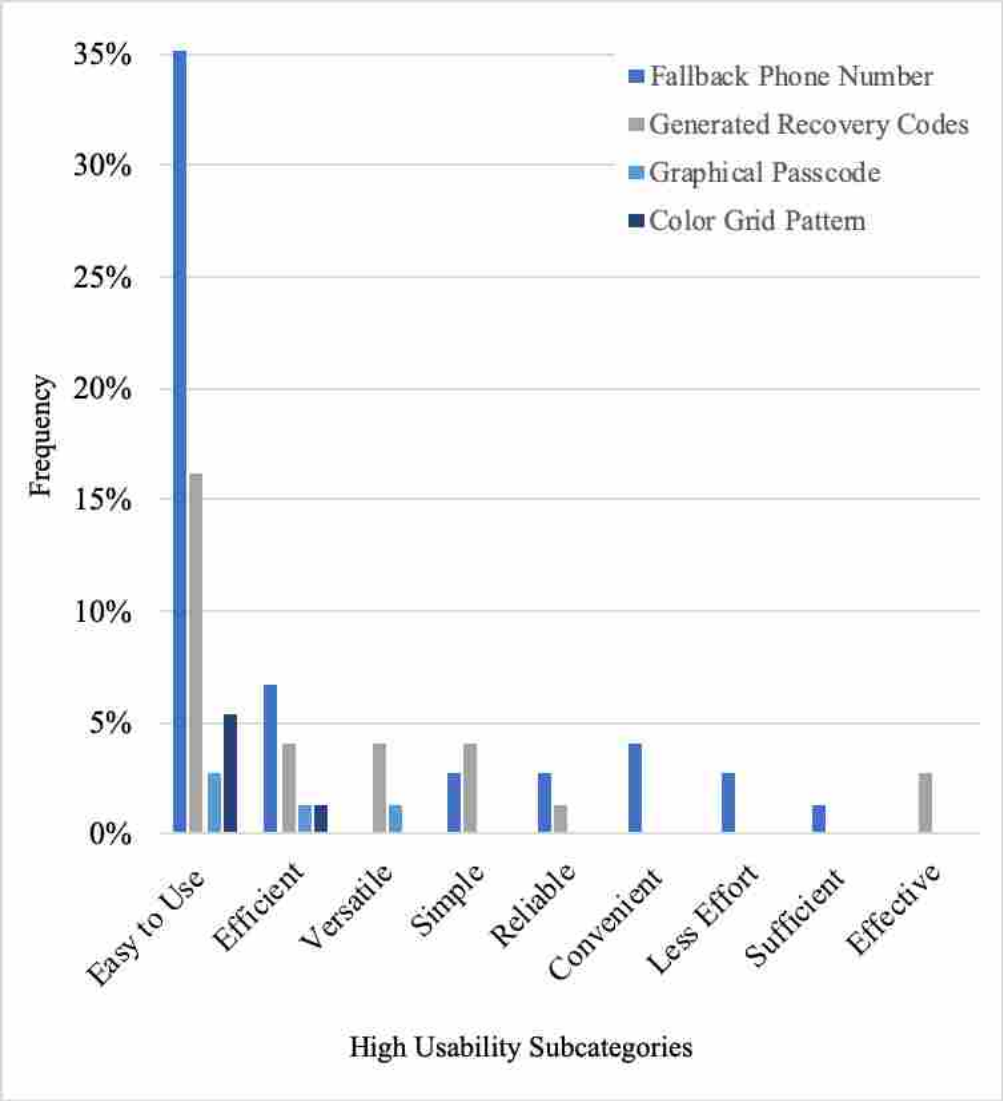
Frequency of Open-ended Response Categories Describing Account Recovery Methods



Note. $N = 217$.

Figure 10

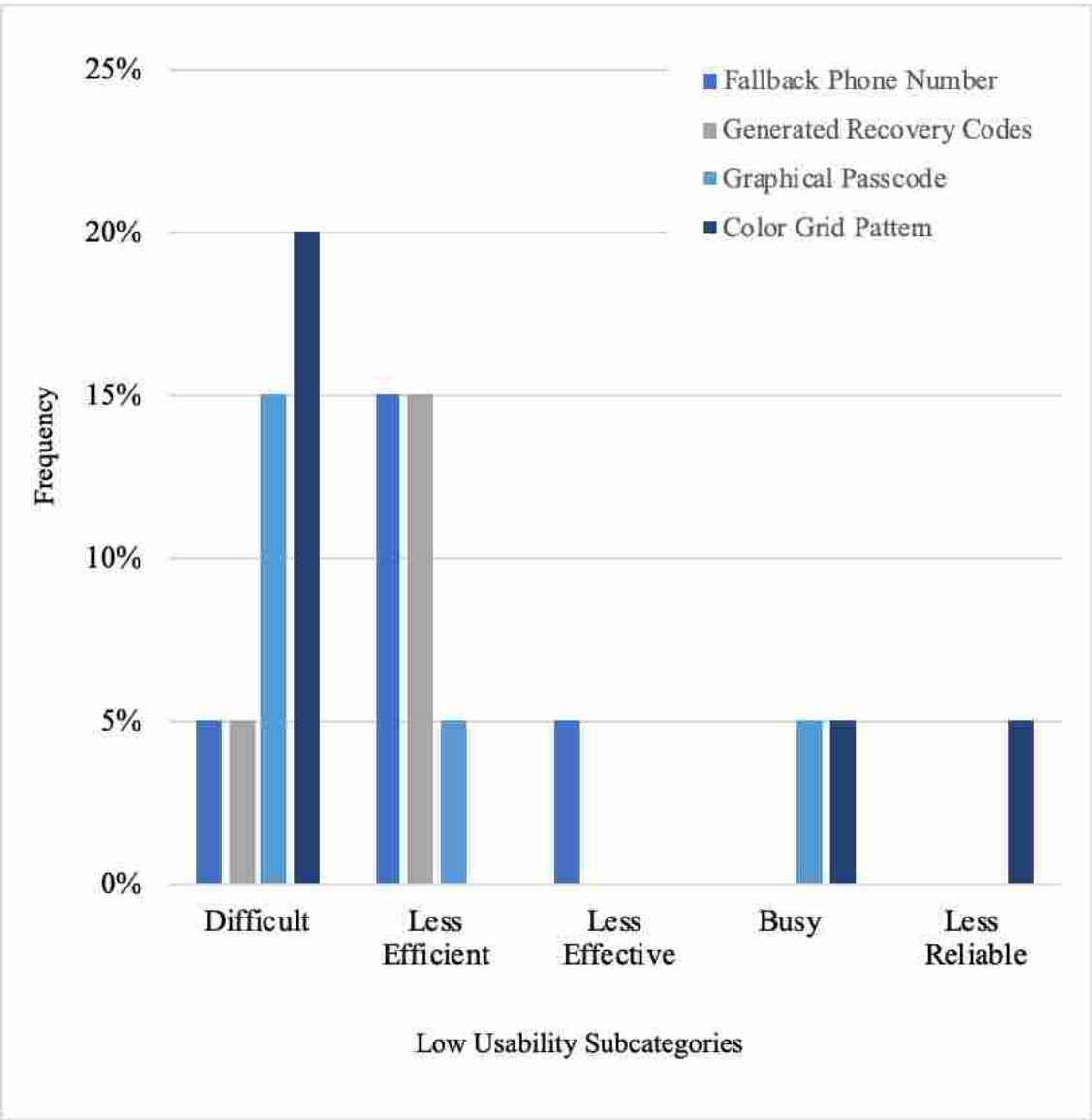
Frequency of High Usability Subcategories by Account Recovery Method



Note. The percentage represents the proportion of subcategory responses out of the 74 total High Usability responses for all recovery methods.

Figure 11

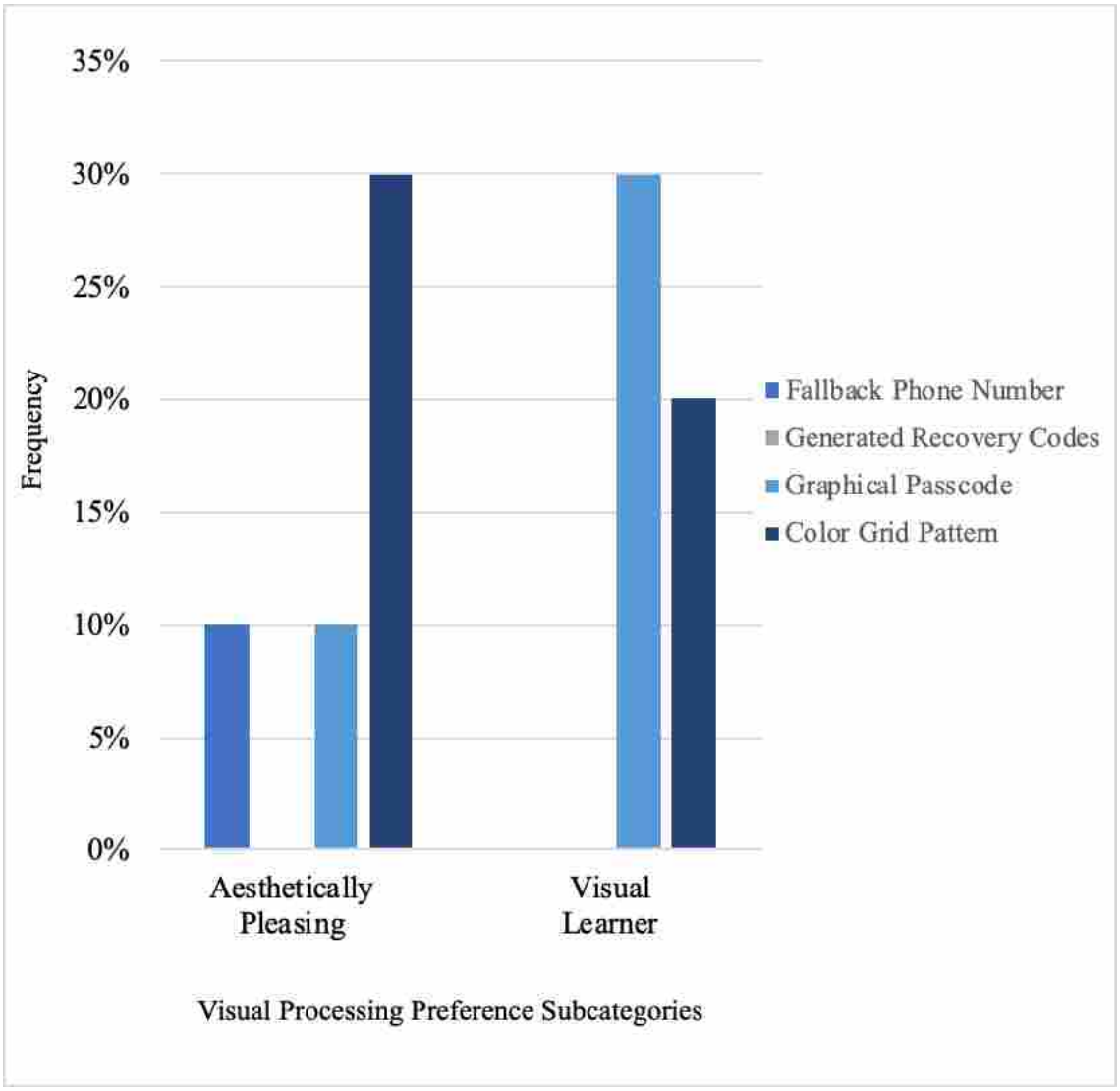
Frequency of Low Usability Subcategories by Account Recovery Method



Note. The percentage represents the proportion of subcategory responses out of the 20 total Low Usability responses for all recovery methods.

Figure 12

Frequency of Visual Processing Preference Subcategories by Account Recovery Method



Note. The percentage represents the proportion of subcategory responses out of the 10 total Visual Processing Preference responses for all recovery methods.

4.2 Statistical Data Analyses

In this statistical analysis section, we tested our four critical survey hypotheses and two general research questions. The critical hypothesis evaluations consist of (1) the similarity between account recovery methods ranked safest to least safe according to the opportunistic score objective ranking and participant's subjective ranking. (2) An analysis of the relations between BNT scores and account recovery security ranking abilities. This research will also present the findings of participants' overall authentication knowledge as it relates to (3) cybersecurity training and (4) gender. The statistical analysis for general research inquiries consist of exploring the relationship between the participant's ability to rank account recovery methods in terms of security to participant's overall authentication knowledge score and to their cybersecurity training experience. Additionally, we explore characteristics that may or may not predict secure account recovery method selection, including experience with 2FA and experience with account recovery. These findings highlight characteristics that may predict participants' ability to determine the security associated with a 2FA account recovery solutions.

4.2.1 Kendall Tau Correlation Significance

Our first hypothesis proposed that the tau correlations coefficients (r_τ) that results from comparing participant's subjective security rank order of the account the recovery methods and the objective recovery method rank order that was determined using Renaud's (2007) opportunistic score will be less than moderate $r_\tau = .33$ on average. A one-sample *t*-test was conducted to explore the tau correlation coefficients. Results indicated participant's average tau correlation coefficient ($M = 0.11$, $SD = 0.56$) was significantly less than the moderate positive correlation lower-end cutoff of $r_\tau = .33$, $t(102) = -4.08$, $p < .001$, $d = -0.402$. The Shapiro-Wilk

assumption check of normality was violated; we plotted the distribution of the values which revealed that the tau coefficients were very scattered and slightly bimodal.

4.2.2 Risk Numeracy and Kendall Tau Correlations for the Subjective Security Ranking of Account Recovery Methods

The second hypothesis proposed that there would be a positive relationship between user's BNT scores and account recovery security ranking abilities. We expected that participants who have a higher risk perception BNT score would reflect fewer ranking discrepancies when they are asked to order the account recovery methods by their security. To test this hypothesis, a linear regression was conducted to determine how well the risk numeracy scores predict the tau correlation coefficients. A linear regression model revealed that the BNT score ($M = 2.51$, $SD = 1.57$) did not statistically significantly predict a tau correlation coefficient, $F(1, 101) = 0.153$, $p = .696$, $R^2 = 0.002$. For every 1 BNT score increase, the tau correlation coefficient is expected to increase by 0.014. This suggests a person's numeracy ability cannot predict their ability to correctly rank order the account recovery methods according to the security they provide. The assumption of linearity was violated, which may indicate an inefficient model; however, all other assumptions were met. A violation of linearity may have occurred due to both variables having different distribution patterns; specifically, BNT scores were positively skewed, and the Tau coefficients were scattered.

As previously noted, we evaluated the raw data for each variable to compare the responses of participants that indicated "Yes", use their data, to those who selected "No", do not use their data. The results of an independent samples *t*-test suggested BNT scores for participants who selected "No" ($M = 1.80$, $SD = 0.91$, $N = 25$) were significantly lower than those who

selected “Yes” (M = 2.73, SD = 1.67, N = 78), t(101) = -2.66, p = .009, d = -0.612. However, the linear regression results did not change when the participants who selected “No” were included in the analysis. As a result, the analysis for this hypothesis included the data of both the participants who indicated “Yes” and those who indicated “No”.

4.2.3 Overall Authentication Knowledge and Cybersecurity Training

The third hypothesis proposed that prior cybersecurity training would not significantly impact the participant’s authentication knowledge. The results revealed that out of the nine concept- and threat- authentication knowledge questions, participants answered more than half correctly on average (M = 6.91, SD = 1.70). However, a low number of participants indicated that they had, “… received training in cybersecurity…” (N = 11). To make group sizes more equal, the “yes” cybersecurity training group was expanded to include participants that indicated they are a cybersecurity expert, or they have taken classes that covered cybersecurity topics. Specifically, we added the participants that selected “yes” to the questions, “Have you taken classes covering the topic of cybersecurity in the past?” or “Do you consider yourself an expert in cybersecurity?” (N = 21; see Appendix A – section 1).

To test our third hypothesis, an independent samples t-test (Cybersecurity training: yes and no) was used to explore the relationship between overall authentication knowledge and training in cybersecurity. Overall authentication knowledge scores for participants who had not received any form of cybersecurity training (M = 6.83, SD = 1.71, N = 82) were not significantly different from the scores of participants who had received cybersecurity training (M = 7.24, SD = 1.67, N = 21), t(101) = -0.98, p = 0.165, d = -0.240. Levene’s test suggested equal variances (F = .48, p = 0.492). The Shapiro-Wilk assumption check of normality was violated. The researchers

plotted the data and observed that participants without cybersecurity training reflected a positively skewed distribution, and participants with cybersecurity training had a negatively skewed distribution, which likely caused the normality violation.

4.2.4 Overall Authentication Knowledge and Gender

The fourth hypothesis proposed that concept- and threat- authentication knowledge scores would be higher for males. To test this hypothesis, an independent samples t-test (Gender: male and female) was conducted to explore the relations between overall authentication knowledge and gender. Results suggested the opposite of the proposed hypothesis. Females (M = 7.22, SD = 1.32, N = 73) had significantly higher overall authentication knowledge scores than males (M = 6.17, SD = 2.26, N = 30), $t(37) = 2.39$, $p = .002$, $d = 0.616$. Levene's test suggested unequal variances ($F = 9.26$, $p = 0.003$), degrees of freedom were adjusted from 101 to 37. The Shapiro-Wilk assumption check of normality was also violated. It is likely that these violations occurred because four male participants reflected near floor performance, which was different from the distribution of performance values for females and the other males. However, the data revealed that the performance of the four males on other measures (e.g., BNT, Tau Ranking) was near average, which implies that removing their data from all analyses would not be justified.

4.2.5 Account Recovery Methods Perceived as Safest and Experience with 2FA

A 2 (2FA experience: yes and no) x 4 (Account method that was selected as the safest: fallback phone numbers, generated account recovery codes, color grid, and graphical passcode) chi-square test was conducted to explore the relations between past experience with 2FA and the

frequency participants selected a given account recovery method as the safest. The relationship was not significant, X^2 (3, N = 103) = 4.01, p = .260.

4.2.6 Account Recovery Methods Perceived as Safest and Experience with 2FA Account Recovery

A 2 (2FA account recovery experience: yes and no) x 4 (recovery method that was selected as the safest: fallback phone numbers, generated account recovery codes, color grid, and graphical passcode) chi-square test was conducted to explore the relations between past experience with 2FA account recovery and account recovery method selection. The relationship was not significant, X^2 (3, N = 89) = 1.56, p = .668.

4.2.7 Kendall Tau Ranking Coefficients and Cybersecurity Training

An independent samples *t*-test (Cybersecurity training: yes and no) was used to explore the relations between the tau correlations coefficients that resulted from ranking the account recovery methods by security strength and cybersecurity training experience. The tau correlation coefficients for participants who had not received any form of cybersecurity training (M = 0.11, SD = 0.57, N = 82) were not significantly different from the tau coefficients of participants who had received cybersecurity training (M = 0.10, SD = 0.49, N = 21), t(101) = 0.11, p = 0.915, d = 0.026. Levene's test suggested equal variances (F = 1.49, p = 0.225). The Shapiro-Wilk assumption check of normality was violated; upon reviewing the distribution of the data the normality violation likely occurred because cybersecurity training groups reflected oppositely skewed distributions.

4.2.7 Kendall Tau Ranking Coefficients and Overall Authentication Knowledge Scores

A Pearson correlation test was used to establish the relations between overall authentication knowledge scores and the tau correlation coefficients. The results suggested there is not a significant relationship between the two variables, $r(101) = -.09$, $p = .346$.

CHAPTER 5

DISCUSSION

5.1 Hypotheses and Research Questions

This research presented the descriptive findings of what users know, perceive, and prefer with respect to 2FA and account recovery. Also, the impact of risk numeracy aptitude, authentication knowledge, and cybersecurity familiarity on the ability to rank order account security strength of recovery methods correctly were explored.

We examined participants' understanding of the security strength of popular account recovery methods. Further, individual differences characteristics that might impact their understanding were examined. For instance, our findings regarding the impact of an individual's overall authentication knowledge and risk numeracy on ranking abilities are insightful. We compared participant's subjective ranking of the recovery methods by security strength to the objective security ranking. We found evidence to support our first hypothesis, which proposed that on average tau correlation coefficients would reflect a less than moderately positive performance. This suggests that users cannot accurately gauge the amount of security that a given recovery method provides. Our second hypothesis examined whether higher risk numeracy abilities could predict a more accurate ranking of the account recovery methods by security strength. Presumably, if a user has higher risk numeracy abilities, they are better able to identify general risks and will be able to abstract recovery method information that implies lower account security. We did not find evidence to support our second hypothesis, and results suggested that risk numeracy aptitude neither benefits nor harms the user's ability to determine the security strength provided by account recovery authentication methods.

Furthermore, we found that there was not a relation between the participant's tau correlation coefficient and their overall authentication knowledge score. These results suggested that neither high-risk numeracy abilities nor knowing basic authentication practices improve user's diagnostic ability to determine which recovery methods are comparatively safer. This information can be important to designers and system administrators when creating training or when selecting and implementing these account recovery methods.

Our hypotheses examined whether gender or having prior cybersecurity training would impact overall authentication knowledge scores. When considering gender's impact on authentication concept- and threat- knowledge, we found females had more overall authentication knowledge than males. This result is the reverse of our proposed hypothesis and does not concur with Cain et al. (2018), finding that males have more knowledge than females. A caveat is that their study used a wider range of cyber hygiene knowledge questions beyond authentication. Additionally, we found that there was a sample size disparity across the gender groups (e.g., males = 30, females = 73), which might have contributed to our inverse findings for this analysis.

When considering cybersecurity training, we found that 20% of participants indicated they had either received cybersecurity training, take a class covering cybersecurity topics, or identified as a cybersecurity expert. This proportion is similar to the findings of Aytes and Conolly's (2003) research, which reported that 19% of their college-age student sample self-identified as being cybersecurity trained. However, it is much lower than more current research by Cain et al. (2018) that reported 81% of participants had received some form of cybersecurity training and the National Cyber Security Alliance (2010), who found 43% of adults had received training. Specifically, our hypothesis proposed that prior cybersecurity training will not

significantly impact the participant's authentication knowledge. We found evidence to support this hypothesis. Our findings agreed with Cain et al. (2018) previous claims that training did not impact users' cyber hygiene behaviors or knowledge. Another exploratory analysis also suggested cybersecurity training experience does not affect a participant's ability to rank the methods by security strength correctly.

We explored whether having previous experience with 2FA or 2FA account recovery would influence which recovery method users select as the safest. Our results suggested that neither type of previous 2FA experience impacted which method was perceived as the safest.

5.2 Data Patterns

The priority with any type of authentication is account security, but usability problems are very common with security-related software (Garfinkel & Lipford, 2014; Schultz, Proctor, Lein, & Salvendy, 2001). In general, users want to use a system to accomplish their goals, and authentication is not typically the primary task on which a user wants to spend effort and time (Garfinkel & Lipford, 2014). Colnago et al. (2018) pointed out that users will have an even lower tolerance when authentication security protocols distract or prevent users from completing their primary task. 2FA and 2FA account recovery are unique because they require additional steps and physical artifacts in order for an authorized user to access their account. The following discussion section acknowledges potential factors and situations that would moderate the use of 2FA devices and 2FA account recovery methods to regain account access. Additionally, we use participant demographics, perceptions, and open-ended preference responses to consider the other expectations and reservations users have regarding the recovery authentication methods.

Pragmatically, there are two ways a user could regain account access if a user's primary 2FA device becomes inaccessible. (1) The user can either enroll more than one type of 2FA device to their account or (2) they can set up an account recovery method. A caveat is that both options require that the user set up their account to accept these supplementary authentication options prior to an event where their primary 2FA device is unavailable. However, both options possess usability shortcomings that merit further examination.

Systems implementing 2FA give users the option to enroll as many 2FA devices to their accounts as they desire. However, we found that 60% of participants indicated they typically have only one 2FA device enrolled per personal account ($M = 1.58$, $Md = 1.00$, $SD = 0.92$). This finding is aligned with the previous findings of Colnago et al. (2018), which suggests users use an average of 1.3 ($Md = 1.00$) 2FA devices. This suggests most users would not have the option to instruct the system to use a different 2FA device to gain account access in the event the user's primary 2FA device becomes unavailable. Potentially, designers can encourage users to enroll more than one 2FA device. Still, it is necessary to determine if having more than one 2FA device enrolled supports prompt account access. Let us consider the user that is conscientious and enrolls more than one 2FA device. For example, a user can set up their account's 2FA to verify their identity via a primary mobile app or a secondary physical hardware key. When the primary device is unavailable, they can simply prompt the system to send the code to their secondary device. However, for their secondary device to be useful, it must be immediately available. Another plausible scenario is a user that has enrolled 2FA -SMS text and -mobile app; both are presumably received on the same phone. If the phone itself becomes inaccessible, both enrolled options are rendered useless. Correspondently, we found that 81% of participants used a mobile or tablet app as their 2FA device, and 51% indicated that they received an SMS text message.

Only 10% of participants stated that they used either a hardware or a software device, and only 8% used a USB security key (see Figure 4). Colnago et al. (2018) found that as the frequency of experiencing 2FA problems increases (e.g., forgetting one's second factor), user's perceptions are negatively impacted, as well as, the usability and security constructs. Future training material should consider ways to convince users that it is necessary to enroll additional 2FA devices, other than just their phone. Upon closer inspection of these 2FA device scenarios, it is evident that there is a security and usability tradeoff when using 2FA. Essentially, account security is increased when the user opts into 2FA (McCandless, 2019). However, all 2FA devices are physical objects, and the user needs to physically have immediate access to all of the enrolled 2FA devices to safeguard immediate account access in all potential circumstances. In the unfortunate event, a user's primary 2FA device is not accessible, and the secondary 2FA device is also not readily available, account access remains problematic.

Alternatively, a website can also offer 2FA account recovery methods to allow participants to safeguard immediate account access. We found that 39 participants had previously used some type of account recovery wherein the majority, 56%, were familiar with using a secondary email as a means to regain account access. Our study proposed that fallback phone numbers and generated account recovery codes are the most secure methods to offer. We also found that the majority of participants perceived these two methods as being safest, 39% for recovery codes, and 25% for the fallback phone number. However, both the fallback phone number and generated recovery codes options require users to have access to additional physical resources to obtain the account recovery verification information. For example, to use the system generated recovery codes method, users would have to physically locate the codes if they were printed, find them on a device if they were stored electronically, or recall a code from memory.

Similarly, the fallback phone number method would require users to either remember the emergency contact's phone number or have a way to obtain their phone number, and they must get in contact with them to receive the code.

Nevertheless, the majority of participants, 38%, perceived the fallback phone as the easiest recovery method to use. Additionally, we found that that 45% of participants indicated they would prefer using the fallback phone number for account recovery. However, 18% of the responses suggested users might disregard the fallback phone number setup instructions by using their phone number instead of an emergency contact. Conversely, 1% understood the fallback phone number instructions and noted that a potential method weakness was, "…most people would likely use their own phone number which goes against the directions…". Presumably, if the users' primary 2FA device is the app, and they initiate the fallback phone number or account recovery codes method, it is likely the case their phone itself is unavailable. Thus, if the user provides their number or stores the generated recovery codes on only their phone, and the phone itself becomes unavailable, these account recovery methods cannot be executed and are no longer useful. Future research should evaluate how users will go about setting up a fallback phone number or recovery codes method.

Additionally, researchers should determine how users obtain their emergency contact number or retrieve their electronically stored generated recovery codes. Will users often have emergency numbers committed to memory? Or will they rely on accessing their electronic files such as a phonebook, recovery code images, or codes stored as encrypted documents? If so, how will they access the files?

In addition to recovery security and usability aspects, memorability is also an especially important factor that must be considered when selecting an account recovery method. Most

likely, users will need to remember their recovery method information after a significantly delayed amount of time. This infers that for a recovery method to be usable, it needs to provide the user with account accessibility (Still et al., 2017). If an account recovery method that is not memorable is implemented, the user's experience and the usability of the 2FA recovery process will be negatively impacted. Our research found that 46% of participants ranked the fallback phone number method as most memorable when asked to rank the methods by perceived memorability. Additionally, we found that 20% of the participant responses touched on the memorability aspect of the recovery methods. Specifically, each method received seven responses indicating that the given method seemed easy to remember, 13%, and 3 to 5 responses that suggested it was difficult to remember, 7%.

Al Ameen (2016) noted that the cognitive abilities for both encoding and retrieval stages of memory are leveraged when users can view targets. However, the fallback phone number and recovery codes methods take a form that is more like a strong traditional alphanumeric password because they require a user to provide a code without any system aids. This suggests these recovery methods rely on a pure-recall process, which is analogous to asking a user to complete an essay exam question by pulling the correct information from memory (Tulving & Watkins, 1973). On the other hand, adopting recovery methods like the color gird pattern or a graphical passcode scheme could provide cognitive benefits that are analogous to asking users to select the correct answer from a set of multiple-choice answers (Tulving & Watkins, 1973). Presumably, recovery methods like these take advantage of the humans' affinity for encoding and recognizing visual objects (e.g., picture superiority effect; Paivio, 2013). Paivio's (2013) research states that unlike letters and numbers, images are encoded both visually and semantically into long-term memory. Studies conducted by Cain and Still (2018) and Chiassion, Forget, Stobert, Van

Oorschot, and Biddle (2009) suggest graphical passcodes have superior retention to alphanumeric and PIN-based authentication. Cain and Still (2018) reported that 100% of participants could not recall their strong system-assigned password after a three-week delay. We found that 13% of the 69 participant responses for the graphical passcodes and color grid methods mentioned that they preferred the method because they had a visual processing preference, and the method was either aesthetically pleasing or they perceive themselves as visual learners. Tentatively, both the color grid pattern and a graphical passcode scheme are good candidates for memorable account recovery options, but they are not as familiar as the alternatives.

5.2 Limitations

There are limitations to this exploratory study. First, it is important to note that participants were asked to self-report their cybersecurity training experiences (e.g., past training, cybersecurity expert, or taken courses covering cybersecurity topics). Unfortunately, this does not provide insight into the topics that participants were training on, and it is unclear whether those topics included authentication material. Participants may have lacked the depth of knowledge required to make informed authentication decisions. This would include knowledge about attack vectors and cyber hygiene best practices. Future research should aim to evaluate what topics are being covered by cybersecurity training material, and the effectiveness of these lessons. Additionally, this research did not inquire about how participants were gauging the security associated with a given recovery method. Conceivably, participants might have been deploying a different heuristic to gauge the security strength of the recovery methods and may not have been explicitly considering properties that indicate secure authentication. Future research should evaluate the

cognitive processes that users employ when comparing the account security strength of different authentication or recovery methods.

This study attempted to objectively rank the recovery methods by security strength using Renaud's (2007) opportunistic score. Specifically, a small group of research assistants independently evaluated each recovery method using the opportunistic score template. However, ranking discrepancies were found among the raters, which led us to use the average opportunistic score for each method to establish our final objective ranking of the recovery methods by security (see Table 3). These scoring discrepancies are potentially a limitation to this study. However, the researchers did observe ways to improve the opportunistic scoring template. For example, a more detailed explanation about the math required for calculating the guessability associated with a given method is needed. Additionally, we detected that some of the questions needed the rater to subjectively score methods. For example, it is the rater's opinion on whether they think they can observe and recall a full key (e.g., recovery code or graphical passcode images or color pattern) after one viewing, multiple viewings, or it cannot be observed (see Appendix C – Observability a). For the graphical recovery methods, the researchers struggled with distinguishing between what constitutes a distractor image versus a background image (see Appendix C – Analyzability g & h). We found that researchers had different interpretations regarding how generated recovery codes would be stored by users and observed by attackers which resulted in recordability and observability scoring discrepancies. For example, most websites note that each recovery code from the set can only be used once, which suggests observing or recoding the user type in that one code would be invaluable to an attacker. However, if the rater considered that attacker observed or recorded the whole list of recovery codes, they scored the method differently. This suggests further research is needed to evaluate

the reliability and validity of Renaud's (2007) opportunistic score. We needed the opportunistic score for determining our ranking of the authentication methods, and we believe others like system administrators will find the tool useful as well. However, future research should attempt to improve the approach to overcome the identified issues.

CHAPTER 6

CONCLUSIONS

This is the first study to consider potential account recovery method solutions for systems that implement 2FA. We found that some people are using 2FA and may not even be aware they are using it. Shockingly, only 40% of users have more than one 2FA device enrolled per personal account. Additionally, the majority of the users sampled opt to use 2FA devices that are executed using their phone, 81% use a mobile app, and 51% receive an SMS text message. This suggests regaining account access could be challenging for users in situations when their primary 2FA device is unavailable, especially if their phone itself becomes unavailable. Ultimately, user's perceptions, as well as, the system's security- and usability- constructs will decrease, as the frequency of 2FA problem that prevents users from accessing their accounts increase (e.g., a broken phone; Colnago et al., 2018). As more systems begin to adopt 2FA, instances that necessitate account recovery will become more prevalent. If companies do not adequately prepare for such occurrences, it could be costly and increase the overhead for their information technology departments. Currently, some companies with 2FA warning users that account recovery could take as long as several business days (Afonin, 2016; Coinbase Support – Account Management, n.d.; Ravenscraft, 2014). If organization employees experience this account access delay, it could be disruptive and prevent them from accessing the services they need to complete their work. Other companies like Reddit, Github, and Google are currently offering precautionary account recovery options (Loveless, 2018; Prins, 2018; Wallen, 2018). Essentially, these organizations deploy 2FA and offer users an account recovery authentication option that can be set up prior to losing a second-factor device. We found that 56% of participants who indicated they had experienced setting up a 2FA account recovery method were most familiar

with a secondary email address method. Presumably, account recovery occurs intermittently; therefore, it is critical for memorability to be an inherent feature of the recovery method being implemented, or the method will be rendered useless. A common pattern across memorability studies suggests that graphical authentication schemes that allow users to recognize their passcode elicit better memory (Cain & Still, 2018; Wiedenbeck et al., 2006). This research considered users' perceived security, usability, memorability, and preference for a graphical passcode-, a user created color grid pattern-, generated recovery codes-, and a fallback phone number- 2FA account recovery methods.

Account security is always a priority when selecting a 2FA account recovery method. This study used Renaud's (2007) opportunistic equation to objectively rank the recovery methods by security strength. We established that a fallback phone number recovery method provides user accounts with the most security followed by system-generated recovery codes, a color grid pattern, and a graphical passcode. Our research provided users with recovery method descriptions that hinted at the method's account security strengths and weaknesses. We asked participants to subjectively rank the recovery methods by security strength. This allowed us to indirectly measure whether or not participants have knowledge about the different authentication attack vectors. We found that participants performed poorly on this task. Further, we investigated three individual difference variables that might impact performance on this task.

Previous research suggests that user's ability to accurately interpret and act on risk information can be revealed by their statistical risk numeracy abilities (Cokely et al., 2012; 2014; Petrova et al., 2019; Schwartz et al., 1997). This study deployed the BNT test to capture the participant's general risk understanding (Cokely et al., 2012; Schwartz et al., 1997). We found variability among participant scores, but on average, scores suggested low to moderate numeracy

abilities. On the other hand, we found that participants exhibited high concept- and threat-authentication knowledge and, presumably, this should have aided in their ability to evaluate the security of a given recovery method. We also inquired about participant's exposure to cybersecurity educational material through previous training, a course covering cybersecurity topics, or if they considered themselves an expert. Surprisingly, we found that these individual difference variables of risk numeracy aptitude, authentication knowledge, and cybersecurity familiarity had no impact on participants' ability to correctly rank the recovery methods by security strength.

Even though our results suggested that cybersecurity training did not impact user's security ranking abilities or authentication knowledge, it is important to highlight that 80% of participants indicated that they had no prior experience with formal cybersecurity training material. Also, our ability to further examine the other 20% of participants that were familiar with cybersecurity material was limited because we did not inquire about the topics that they had been exposed to nor the breadth of their training. Presumably, participant's poor ranking abilities might indirectly suggest that they lack a deeper understanding of the different authentication attack vectors, despite their moderate general authentication knowledge. Or users were possibly applying other usability heuristics to determine recovery method security strength.

The findings that the aforementioned individual differences do not impact user's security diagnostic abilities and the other insights provide useful guidance for communicating recovery method security information. From a practitioner's perspective, we propose the following list of recommendations. First, companies that use 2FA should encourage users to enroll at least two different types of 2FA devices (e.g., a mobile app and a USB key). Companies should also provide users with the option to set up an account recovery method to circumvent problems that

may otherwise prevent users from accessing their accounts. We recommend that companies educate users about recovery options by considering the best practices for communicating recovery method security information to promote informed judgment. For example, if the company wants to deploy the fallback phone number recovery option, we recommend for their goal to focus on drawing users away from using their phone number (e.g., offer narratives highlighting negative use cases). This aligns with previous research, which suggests that a company's communication should focus on leading the user away from making a risky security decision when they want to avoid user issues (Nurse et al., 2011). Lastly, we found that risk literacy abilities do not benefit nor harm a user's appraisal of recovery method security. We are not aware of any companies that provide users with representative authentication security information. We recommend companies be more transparent by providing users with this information using a visual (Cokely et al., 2018; Nurse et al., 2011). For instance, they could provide a comparison depicting the probabilities of a successful attack occurring when using the given recovery methods. Providing users with this explicit security information will aid the user in making a more informed judgment.

REFERENCES

Angeli, C. (2013). Examining the effects of field dependence–independence on learners' problem-solving performance and interaction with a computer modeling tool: Implications for the design of joint cognitive systems. *Computers & Education*, *62*, 221-230. doi: 10.1016/j.compedu.2012.11.002

Afonin, O. (2016, December 20). The ugly side of Two-Factor Authentication [Web blog post]. Retrieved from https://blog.elcomsoft.com/2016/12/the-ugly-side-of-two-factor-authentication/

Al Ameen, M. N. (2016). *The impact of cues and user interaction on the memorability of system-assigned random passwords* (Doctoral dissertation). Available from The University of Texas at Arlington (UTA) Library Database. Retrieved from https://rc.library.uta.edu/uta-ir/bitstream/handle/10106/25773/ALAMEEN-DISSERTATION-2016.pdf?sequence=1&isAllowed=y

Almuhimedi, H., Schaub, F., Sadeh, N., Adjerid, I., Acquisti, A., Gluck, J., Cranor, L., & Agarwal, Y. (2015). Your location has been shared 5,398 times! A field study on mobile app privacy nudging. *Proceedings of the 33rd Annual CHI Conference on Human Factors in Computing Systems*, AMC, 787-796. doi: 10.1145/2702123.2702210

Alonso, O. (2013). Implementing crowdsourcing-based relevance experimentation: An industrial perspective. *Journal of Information Retrieval*, *16*(2), 101-120. doi: 10.1007/s10791-012-9204-1

Ashford W. (2009, September 7). Millions of web users at risk from weak passwords [Web blog post]. Retrieved from http://www.computerweekly.com/Articles/2009/09/07/237569/Millions-of-web-users-at-risk-from-weak-passwords.htm?printerfriendly=true 07

Aytes, K., & Conolly, T. (2003). A research model for investigating human behavior related to computer security. *Proceedings of Americans Conference on Information Systems*, *9*, 2027-2031. Retrieved from https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1726&context=amcis2003

Barton, B. F., & Barton, M. S. (1984). User-friendly password methods for computer-mediated information systems. *Computers & Security*, *3*(3), 186-195. doi: 2027.42/24740/0000162

Bachmann, A., & Bernstein, A. (2010). When process data quality affects the number of bugs: Correlations in software engineering datasets. *Proceedings of the 7th Working Conference on Mining Software Repositories*, IEEE, 62-71. doi: 10.1109/MSR.2010.5463286

Boechler, P. M. (2006). Understanding cognitive processes in educational hypermedia. In C. Ghaoui (Ed.), *Encyclopedia on Human Computers Interaction,* 648-651. doi:10.4018/978-1-59140-562-7.ch097

Brühlmann, F., & Mekler, E. D. (2018). Surveys in games user research. *Games User Research*, 141-162.

Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence, and kappa. *Journal of Clinical Epidemiology*, *46*(5), 423-429. doi: 10.1016/0895-4356(93)90018-V

Cain, A. A., Edwards, M. E., & Still, J. D. (2018). An exploratory study of cyber hygiene behaviors and knowledge. *Journal of Information Security and Applications*, *42*, 36-45. doi: 10.1016/j.jisa.2018.08.002

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

Cain, A. A., & Still, J. D. (2018). Usability comparison of over-the-shoulder attack resistant authentication schemes. *Journal of Usability Studies*, *13*(4), 196-219. Retrieved from http://uxpajournal.org/wp-content/uploads/sites/8/pdf/JUS_Cain_August2018.pdf

Cazier, J. A., & Medlin, B. D. (2006). Password security: An empirical investigation into ecommerce passwords and their crack times. *Journal of Information Systems Security*, *15*(6), 45–55. doi: 10.1080/10658980601051318

Chiasson, S., Forget, A., Stobert, E., Van Oorschot, P. C., & Biddle, R. (2009). Multiple password interference in text passwords and click-based graphical passwords. *Proceedings of the 16th ACM conference on Computer and communications security*, 500-511. doi: 10.1145/1653662.1653722

Chen, J., Ge, H., Moore, S., Yang, W., Li, N., & Proctor, R. W. (2018). Display of major risk categories for android apps. *Journal of Experimental Psychology: Applied*, *24*(3), 306-330. doi: 10.1037/xap0000163

Choong, Y., & Greene, K. (2016). What's a special character anyway? Effects of ambiguous terminology in password rules. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *60*(1), 760-764. doi: 10.1177/1541931213601174

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Journal of Educational and Psychological Measurement*, *20*(1), 37–46. doi: 10.1177/001316446002000104

Coinbase Support – Account Management. (n.d.). I have lost or need to update my phone or 2-Factor Authentication Device [Web support page]. Retrieved from

https://support.coinbase.com/customer/en/portal/articles/2488794-troubleshooting-2-factor-authentication

Cokely, E. T., Feltz, A., Ghazal, S., Allan, J. N., Petrova, D., & Garcia-Retamero, R. (2018). Decision making skill: From intelligence to numeracy and expertise. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (2nd Eds.), *Cambridge Handbook of Expertise and Expert Performance*, 476-505. Retrieved from http://www.riskliteracy.org/files/1414/8312/0711/2016_Cokely_et_al_Decision_Making_Skill_CamHandbook_Final_Pre-Print.pdf

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The berlin numeracy test. *Journal of Judgment and Decision Making*, *7*(1), 25-47. Retrieved from https://psycnet.apa.org/record/2012-03055-003

Cokely, E. T., Ghazal, S., & Garcia-Retamero, R. (2014). Measuring numeracy. In B. L. Anderson & J. Schulkin (Eds.), *Numerical Reasoning in Judgments and Decision Making about Health*, 11-38. Retrieved from https://books.google.com/books?hl=en&lr=&id=pCCmAwAAQBAJ&oi=fnd&pg=PA1&dq=Cokely+ET,+Ghazal+S,+Garcia-Retamero+R.+Measuring+numeracy.+In:+Anderson+BL,+Schulkin+J,+editors.+Numerical+Reasoning+in+Judgments+and+Decision+Making+about+Health.+Cambridge,+UK:+Cambridge+University+Press%3B+2014.+pp.+11%E2%80%9338.&ots=b7aa83UXLL&sig=Df_21ygzTYn8ga9p8Wd-7r0rOEo#v=onepage&q&f=false

Colnago, J., Devlin, S., Oates, M., Swoopes, C., Bauer, L., Cranor, L., & Christin, N. (2018). “It's not actually that horrible”: Exploring adoption of two-factor authentication at a

university. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems,* ACM, 1-11. doi: 10.1145/3173574.3174030

Cone, B. D., Thompson, M. F., Irvine, C. E., & Nguyen, T. D. (2006). Cyber security training and awareness through game play. In S. Fischer-Hübner, K. Rannenberg, L. Yngström, & S. Lindskog (Eds.), *Proceedings of IFIP International Information Security Conference*, *201*, 431-436. doi: 10.1007/0-387-33406-8_37

Conklin, A., Dietrich, G., & Walz, D. (2004). Password-based authentication: A system perspective. *Proceedings of the 37th Annual Hawaii International Conference on System Sciences*, IEEE, 1-10. doi: 10.1109/HICSS.2004.1265412

Cooper, A., Reimann, R., & Cronin, D. (2007). *About face 3: The essentials of interaction design*. John Wiley & Sons. Retrieved from https://fall14se.files.wordpress.com/2017/04/about_face_3__the_essentials_of_interaction_design.pdf

Coventry, L., Briggs, P., Jeske, D., & van Moorsel, A. (2014). Scene: A structured means for creating and evaluating behavioral nudges in a cyber security environment. *Proceedings of the International Conference of Design, User Experience, and Usability*, *8517*, 229-239. doi: 10.1007/978-3-319-07668-3_23

Cox, A., Connolly, S., & Currall, J. (2001). Raising information security awareness in the academic setting. *VINE Journal of Information and Knowledge Management Systems*, *31*(2), 11-16. doi: 10.1108/03055720010803961

Cunningham, J. A., Godinho, A., Kushnir, V., & Bertholet, N. (2017). What does it mean when people say that they have received expressions of concern about their drinking or advice to cut down on the AUDIT scale?. *BMC medical research methodology*, *17*(1), 158-163.

Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology*, *66*, 4-19.

Cybersecurity and Infrastructure Security Agency (2009, May 21). Choosing and protecting passwords. *National Cyber Awareness System: Security Tip ST04-002*. Retrieved from http://www.us-cert.gov/cas/tips/ST04-002.html

Das, S., Dingman, A., & Camp, L. J. (2018). Why Johnny doesn't use two factor a two-phase usability study of the FIDO U2F security key. *Proceedings of the International Conference on Financial Cryptography and Data Security*. doi: 10.1007/978-3-662-58387-6

Dawson, L. A., & Stinebaugh, J. (2010). *Methodology for prioritizing cyber-vulnerable critical infrastructure equipment and mitigation strategies* (No. SAND2010-1845). Sandia National Laboratories. Retrieved from https://prod-ng.sandia.gov/techlib-noauth/access-control.cgi/2010/101845.pdf

De Luca, A., Denzel, M., & Hussmann, H. (2009). Look into my eyes! Can you guess my password?. *Proceedings of the 5th Symposium on Usable Privacy and Security,* ACM, 160-179. doi: 10.1145/1572532.1572542

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

Department of Homeland Security Federal Infrastructure Protection Bureau [DHS-IP], Carnegie Mellon Software Engineering Institute. (2013). *Unintentional Insider Threats: A Foundational Study* (Contract No. A8721- 05-C-0003). Retrieved from https://resources.sei.cmu.edu/asset_files/TechnicalNote/2013_004_001_58748.pdf

Duo. (2019). *Secure two-factor authentication app*. Retrieved from https://duo.com/product/trusted-users/two-factor-authentication/duo-mobile

Egelman, S., Jain, S., Portnoff, R. S., Liao, K., Consolvo, S., & Wagner, D. (2014). Are you ready to lock?. *Proceedings of the SIGSAC Conference on Computer and Communications Security*, AMC, 750-761. doi: 10.1145/2660267.2660273

English, R., & Poet, R. (2012). The effectiveness of intersection attack countermeasures for graphical passwords. *Proceedings of the 11th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, IEEE, 1-8. doi: 10.1109/TrustCom.2012.271

Federal Cybersecurity Enhancement Act of 2016, 40 U.S.C § 11331 *et seq*. (2016)

Fleiss, J. L., Levin, B., & Paik, M. C. (2003). The measurement of interrater agreement. Statistical methods for rates and proportions, *Statistical Methods for Rates and Proportions*, 598-626. doi: 10.1002/0471445428

Florêncio, D., Herley, C., & Coskun, B. (2007). Do strong web passwords accomplish anything? *HotSec*, *7*(6). Retrieved from https://www.usenix.org/legacy/event/hotsec07/tech/full_papers/florencio/florencio.pdf

Gao, H., Liu, X., Dai, R., Wang, S., & Chang, X. (2009). Analysis and evaluation of the colorlogin graphical password scheme. *Proceedings of the International Conference on Image and Graphics*, IEEE, *5*, 722-727. doi: 10.1109/ICIG.2009.62

Garfinkel, S., & Lipford, H. R. (2014). Usable security: History, themes, and challenges. *Synthesis Lectures on Information Security, Privacy, and Trust*, *5*(2), 1-124. doi: 10.2200/S00594ED1V01Y201408SPT011

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Retrieved from http://www.sxf.uevora.pt/wp-content/uploads/2013/03/Glaser_1967.pdf.

Goldberg, J. (2018). What does "MFA" mean?. *Proceedings of the International Conference Workshop Who Are You?! Adventures in Authentication Workshop*, USENIX SOUPS, (4). Retrieved from https://wayworkshop.org/2018/papers/way2018-goldberg.pdf

GitHub Help. (2019). *Recovering your account if you lose your 2FA credentials*. Retrieved from https://help.github.com/en/articles/recovering-your-account-if-you-lose-your-2fa-credentials#authenticating-with-a-fallback-number

Grassie, P. A., Garcia, M. E., & Fenton, J. L., (2017). Digital identity guidelines [Special Publication 800-63-3]. *National Institute of Standards and Technology*. doi: 10.6028/NIST.SP.800-63-3

Gratian, M., Bandi, S., Cukier, M., Dykstra, J., & Ginther, A. (2018). Correlating human traits and cyber security behavior intentions. *Journal of Computers & Security*, *73*, 345-358. dio: 10.1016/j.cose.2017.11.015

Greenhow, C., Li, J., & Mai, M. (2019). From tweeting to meeting: Expansive professional learning and the academic conference backchannel. *British Journal of Educational Technology*, *50*(4), 1656-1672. doi: 10.1111/bjet.12817

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in quantitative methods for psychology*, *8*(1), 23-34. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3402032/#

Hayashi, E., Dhamija, R., Christin, N., & Perrig, A. (2008). Use your illusion: Secure authentication usable anywhere. *Proceedings of the 4th Symposium on Usable Privacy and Security*, ACM, 35-45. doi: 10.1145/1408664.1408670

Hogg, N. (2007). Measuring cognitive load. In R. A. Reynolds, R. Woods, & J. D. Baker (Eds.), *Handbook of Research on Electronic Surveys and Measurements*, 188-194. doi: 10.4018/978-1-59140-792-8.ch020

Hoonakker, P., Bornoe, N., & Carayon, P. (2009). Password authentication from a human factors perspective: Results of a survey among end-users. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *53*(6), 459-463. doi: https://doi.org/10.1177/154193120905300605

Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, *64*(2), 79-102. doi: 10.1016/j.ijhcs.2005.06.002

Information Technology Services. (2019). Two-Factor Authentication. *Old Dominion University*. Retrieved from https://www.odu.edu/ts/access/two-factor-authentication

Infosecurity Europe. (2008). Woman 4 times more likely than men to give passwords for chocolate. *Press Release*. Retrieved from http://www.infosec.co.uk/page.cfm/T=m/Action=Press/PressID=1071

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, *30*(1/2), JSTOR, 81-93. doi: 10.2307/2332226

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

Konieczny, F., Trias, E., & Taylor, N. J. (2015). SEADE: Countering the futility of network security. *Journal of Air & Space Power*, 4-14. Retrieved from https://apps.dtic.mil/dtic/tr/fulltext/u2/a624282.pdf

Kraus, L., Schmidt, R., Walch, M., Schaub, F., & Möller, S. (2017). On the use of emojis in mobile authentication. In S. De Capitani di Vimercati & F. Martinelli (Eds.), *ICT Systems Security and Privacy Protection*, IFIPAICT, *502*, 265-280. doi: 10.1007/978-3-319-58469-0_18

Labuschagne, W. A., Veerasamy, N., Burke, I., & Eloff, M. M. (2011). Design of cyber security awareness game utilizing a social media framework. *Proceedings of Information Security South Africa*, 1-9. doi: 10.1109/ISSA.2011.6027538

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159-174. Retrieved from https://www.jstor.org/stable/2529310

Lang, J., Czeskis, A., Balfanz, D., Schilder, M., & Srinivas, S. (2017). Security keys: Practical cryptographic second factors for the modern web. *International Conference on Financial Cryptography and Data Security*, *9603*, 422-440. doi: 10.1007/978-3-662-54970-4_25

Leu, E. (2017, June 8). Authentication trends for 2017. *Upwork Global*. Retrieved from https://www.upwork.com/hiring/for-clients/authentication-trends/

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly-educated samples. *Journal of Medical Decision Making*, *21*, 37–44. doi: 10.1177/0272989X0102100105

Loveless, M. (2018, March 6). How popular web services handle account recovery. *Duo Security*. Retrieved from https://duo.com/decipher/reality-of-online-account-recovery

Mackie, C. (2017). Portable open-source authenticator for Windows. *WinAuth*. Retrieved from https://winauth.github.io/winauth/index.html

Madigan, S. (1983). Picture memory. In J. C. Yuille (Ed.), *Imagery, Memory and Cognition: Essays in Honor of Allan Paivio*, 65-89. doi: 10.4324/9781315774787

Markelj, B., & Bernik, I. (2015). Safe use of mobile devices arises from knowing the threats. *Journal of Information Security and Applications*, *20*, 84-89. doi: 10.1016/j.jisa.2014.11.001

Mator, J. D., Lehman, W. E., McManus, W., Powers, S., Tiller, L. N., Unverricht, J. R., & Still, J. D. (2020). Usability: Adoption, Measurement, Value. *Human Factors*.

McCandless, D. (2019). World's biggest data breaches & hacks. *Information is Beautiful*. Retrieved from https://informationisbeautiful.net/2018/worlds-biggest-data-breaches-hacks-updated/

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological methods*, *17*(3), 437-455.

MIDAS & Monarch-Key. (2019, February 16). MIDAS (midas.odu.edu). *Old Dominion University*. Retrieved from https://www.odu.edu/ts/access/monarchkey#tab5=0

National Cyber Security Alliance. (2010). *NCSA / Norton by Symantec Online Safety Study*. Retrieved from http://www.staysafeonline.org/download/datasets/2064/FINAL+NCSA+Full+Online+Safety+Study+2010%5B1%5D.pdf

Nurse, J. R., Creese, S., Goldsmith, M., & Lamberts, K. (2011). Trustworthy and effective communication of cybersecurity risks: A review. *Proceedings of the 1st Workshop on*

Socio-Technical Aspects in Security and Trust (STAST), IEEE, 60-68. doi: 10.1109/STAST.2011.6059257

Siegel, S., & Castellan, N. J. (1988). *Nonparametric Statistics for the Behavioral Sciences (2nd Ed.)*. New York: McGraw-Hill.

Smith, M., Wilbur, J., & Spiezle, C. (2018, January 25). Cyber incident & breach trends report. *Online Trust Alliance of the Internet Society*. Retrieved from https://www.internetsociety.org/wp-content/uploads/2019/04/2018-cyber-incident-report.pdf

Paivio, A. (2013). *Imagery and Verbal Processes*. Psychology Press. doi: 10.4324/9781315798868

Petrova, D., Mas, G., Navarrete, G., Rodriguez, T. T., Ortiz, P. J., & Garcia-Retamero, R. (2019). Cancer screening risk literacy of physicians in training: An experimental study. *PloS ONE*, *14*(7). doi: 10.17605/OSF.IO/QN9A2

Pattinson, M. R. (2012). *An examination of information system risk perceptions using the repertory grid technique* (Doctoral thesis, University of Adelaide, Adelaide, South Australia). Retrieved from https://digital.library.adelaide.edu.au/dspace/bitstream/2440/87355/8/02whole.pdf

Pelgrin, W. (2014). A model for positive change: Influencing positive change in cyber security strategy, human factor, and leadership. In M. E. Hathaway (Ed.), *Best Practices in Computer Network Defense: Incident Detection and Response*, 107-117. doi: 10.3233/978-1-61499-372-8-107

Prins, C. W. (2018, April 15). 2-Factor authentication recovery codes [Web blog post]. *4me*. Retrieved from https://www.4me.com/blog/two-factor-authentication/two-factor-authentication-recovery-codes/

Ravenscraft, E. (2014, December 09). What happens if I use two-factor authentication and lose my phone? [Web blog post]. *lifehacker*. Retrieved from https://lifehacker.com/what-do-i-do-if-i-use-two-factor-authentication-and-los-1668727532

Reese, K. R. (2018). *Evaluating the usability of two-factor authentication* (Master's thesis). Available from All Theses and Dissertations Database. (UMI No. 6869)

Renaud, K. (2007). A process for supporting risk-aware web authentication mechanism choice. *Journal of Reliability Engineering & System Safety*, *92*(9), 1204-1217. doi: 10.1016/j.ress.2006.08.008

Russell, J. D., Weems, C. F., Ahmed, I., & Richard III, G. G. (2017). Self-reported secure and insecure cyber behaviour: Factor structure and associations with personality factors. *Journal of Cyber Security Technology*, *1*(3-4), 163-174. doi: 10.1080/23742917.2017.1345271

Schultz, E. E., Proctor, R. W., Lien, M. C., & Salvendy, G. (2001). Usability and security an appraisal of usability issues in information security methods. *Computers & Security*, *20*(7), 620-634.

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine*, *127*(11), 966−972. Retrieved from https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1001.5503&rep=rep1&type=pdf

Shaban, H. (2018, October 8). The government is rolling out 2-factor authentication for federal agency dot-gov domains. *The Washington Post*. Retrieved from https://www.washingtonpost.com/technology/2018/10/08/government-is-rolling-out-factor-authentication-federal-agency-gov-domains/?noredirect=on&utm_term=.0b2701f0143b

Sharit, J., Lisigurski, M., Andrade, A. D., Karanam, C., Nazi, K. M., Lewis, J. R., & Ruiz, J. G. (2014). The roles of health literacy, numeracy, and graph literacy on the usability of the VA's personal health record by veterans. *Journal of Usability Studies*, *9*(4), 173-193. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.568.6582&rep=rep1&type=pdf

Shneiderman, B., & Plaisant C. (2010). Designing the user interface. *Strategies for Effective Human-Computer Interaction 5th edition*. Pearson Addison-Wesley Publishers.

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, *15*(1), 72–101. Retrieved from https://books.google.com/books?hl=en&lr=&id=1Me3AAAAIAAJ&oi=fnd&pg=PA72&dq=Spearman,+C.+(1904).+The+Proof+and+Measurement+of+Association+Between+Two+Things.+The+American+Journal+of+Psychology,+15(1),+72–101&ots=jt2DI1QvWF&sig=fQRQKw44ZwLSSctrS43IgO_2qy0#v=onepage&q&f=false

Srinivas, S., Balfanz, D., Tiffany, E., & Czeskis, A. (2017, April 11). Universal 2nd Factor (U2F) overview. FIDO Alliance Proposed Standard. Retrieved from https://fidoalliance.org/specs/fido-u2f-v1.2-ps-20170411/FIDO-U2F-COMPLETE-v1.2-ps-20170411.pdf

Still, J. D., Cain, A. A., & Schuster, D. (2017). Human-centered authentication guidelines. *Journal of Information & Computer Security*, *25*(4), 437-453. doi: 10.1108/ICS-04-2016-0034

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

Sweller, J. (1988). Cognitive load during problem solving: Effect on learning. *Journal of Cognitive Science*, *12*(2), 257-285. doi: 10.1016/0364-0213(88)90023-7

Tellini, N., & Vargas, F. (2017). *Two-Factor Authentication: Selecting and implementing a two-factor authentication method for a digital assessment platform* (Bachelor's Thesis, KTH Royal Institute of Technology, Stockholm, Sweden). Retrieved from http://www.diva-portal.org/smash/get/diva2:1104740/FULLTEXT01.pdf

Tiller, L. N., Angelini, C. A., Leibner, S. C., & Still, J. D. (2019). Explore-a-Nation: Combining graphical and alphanumeric authentication. In A. Moallem (Ed.), *HCI for Cybersecurity, Privacy and Trust*, HCII, *11594*, 81-95. doi: 10.1007/978-3-030-22351-9_6

Tulving, E., & Watkins, M. J. (1973). Continuity between recall and recognition. *The American Journal of Psychology*, *86*(4), 739-748. doi: 10.2307/1422081

Ur, B., Noma, F., Bees, J., Segreti, S. M., Shay, R., Bauer, L., Christin, L., & Cranor, L. F. (2015). "I added '!' at the end to make it secure": Observing password creation in the lab. *Proceedings of the 11th Symposium On Usable Privacy and Security* (SOUPS), 123-140.

Van Schaik, P., Jeske, D., Onibokun, J., Coventry, L., Jansen, J., & Kusev, P. (2017). Risk perceptions of cyber-security and precautionary behavior. *Journal of Computers in Human Behavior*, *75*, 547-559. doi: 10.1016/j.chb.2017.05.038

Wallen, J. (2018, August 7). How to retrieve your Google 2FA backup codes (and make more)[Web blog post]. *TechRepublic*. Retrieved from https://www.techrepublic.com/article/how-to-retrieve-your-google-2fa-backup-codes-and-make-more/

Wickelgren, W. A., & Norman, D. A. (1966). Strength models and serial position in short-term recognition memory. *Journal of Mathematical Psychology*, *3*(2), 316-347. doi: 10.1016/0022-2496(66)90018-6

Wiedenbeck, S., Waters, J., Birget, J. C., Brodskiy, A., & Memon, N. (2005). PassPoints: Design and longitudinal evaluation of a graphical password system. *International Journal of Human-Computer Studies*, *63*(1-2), 102-127. doi: 10.1016/j.ijhcs.2005.04.010

Wiedenbeck, S., Waters, J., Sobrado, L., & Birget, J. C. (2006). Design and evaluation of a shoulder-surfing resistant graphical password scheme. *Proceedings of the Working Conference on Advanced Visual Interfaces*, ACM, 177–184. doi: 10.1145/1133265.1133303

Woller-Carter, M. M., Okan, Y., Cokely, E. T., & Garcia-Retamero, R. (2012). Communicating and distorting risks with graphs: An eye-tracking study. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, HFES, *56*(1), 1723-1727. doi: 10.1177/1071181312561345

Yale (2007, April 13). Selecting good passwords. *Yale University Guide 1610 GD.01*. Retrieved from http://hipaa.yale.edu/sites/default/files/files/1610-GD01-Selecting-Good-Passwords.pdf

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

Zyiran, M., & Haga, W. J. (1999). Password security: An empirical study. *Journal of Management Information Systems*, *15*(4), 161-185. doi: 10.1080/07421222.1999.11518226

APPENDICES

APPENDIX A: STUDY SURVEY

Section 1: Demographics

1. What is your age? _____
2. What is your gender?
    a. Female
    b. Male
3. Are you colorblind?
    a. Yes
    b. No
4. Is your major in a technical field? (For example: engineering, science, or applied math)
    a. Yes
    b. No
    a. N/A
5. What is your declared college major? _____
6. Have you been trained in cyber security in the past?
    a. Yes
    b. No
7. If you have received training in cyber security, where did this training take place? Check all that apply
    a. Work
    b. School
    c. Online
    d. Other _____
    e. N/A
8. Have you taken classes covering the topic of cyber security in the past?
    a. Yes
    b. No
9. Do you consider yourself an expert in cyber security?
    a. Yes
    b. No
10. Have you ever been the target of a cybersecurity attack in the past?
    a. Yes
    b. No
11. What type of electronics do you typically use? Check all that apply
    a. Desktop
    b. Laptop
    c. Smart Phone
    d. Tablet
    e. Smart Watch
    f. Other _____

12. On average, how many hours do you spend using technology a day?
13. Is your heart beating?
    a. Yes
    b. No
14. Do you currently use two-factor authentication (2FA) to protect any of your personal accounts?
    a. Yes
    b. No

If yes, answer the following 3 questions:

15. Which 2FA devices have you used for your personal accounts? Select all that apply.
    a. SMS text messaging
    b. Smartphone or tablet App (ex. Duo Mobile)
    c. Hardware token (a small physical token, requires you push a button to receive a one-time 6- or 8- digit OATH-HOTP passcode)
    d. Software token (software that has been downloaded onto your computer that use TOTP or HOTP authentication to generate a one-time passcode)
    e. Security Key (aka. U2F token) (ex. YubiKey)
    f. Other _____
16. How many 2FA devices do you have enrolled for one account, on average?
    a. 1
    b. 2
    c. 3
    d. 4
    e. 5 or more
17. Have you ever been asked to setup a 2FA account recovery option, for cases such as a lost or stolen second factor?
    a. Yes What type of recovery method was setup? _____
    b. No

Section 2: Knowledge of Authentication Concepts (bolded answers indicate correct responses)

1. Which of these is the strongest password?
    a. Penguin$123
    b. DoG?99
    c. Magicusa1990
    d. DomMom390#
2. What are over the shoulder attacks/ shoulder surfing?
    a. When a person attacks from over the shoulder and steals their device
    b. When a person watches the user type in their login information and password, in order to gain access to their accounts
    c. When a hacker in a public place infects the user's computer with a virus without making physical contact, just standing over their shoulder
    d. When a hacker overrides the user's antivirus software and gains access to their account

3. What is authentication?
    a. **The process of verifying the identity of the user**
    b. The act of protecting a user's identity
    c. Having multiple accounts with the same password
    d. Assigning the user a unique user name
4. We care about data quality. To ensure you are currently paying attention can you please select the color option "Yellow"?
    a. Green
    b. Purple
    c. Yellow
    d. Blue

Section 3: Knowledge of Threats (bolded answers indicate correct responses)

1. Posting personal information on social media can be dangerous
    a. Strongly Disagree
    b. Disagree
    c. Neither Agree nor Disagree
    d. Agree
    e. Strongly Agree
2. It is safe to share a password with others
    a. Strongly Disagree
    b. Disagree
    c. Neither Agree nor Disagree
    d. Agree
    e. Strongly Agree
3. Use of a strong password can decrease the user's vulnerability to an attack
    a. Strongly Disagree
    b. Disagree
    c. Neither Agree nor Disagree
    d. Agree
    e. Strongly Agree
4. Using the same passwords on multiple websites decreases the user's vulnerability to an attack
    a. Strongly Disagree
    b. Disagree
    c. Neither Agree nor Disagree
    d. Agree
    e. Strongly Agree
5. It is not important to password protect your mobile device
    a. Strongly Disagree
    b. Disagree
    c. Neither Agree nor Disagree
    d. Agree

   e. Strongly Agree
6. Respond with "Strongly Agree" to this question
   a. Strongly Disagree
   b. Disagree
   c. Neither Agree nor Disagree
   d. Agree
   e. Strongly Agree
7. Error messages provided by the website during authentication can prove to be valuable clues to an attacker
   a. Strongly Disagree
   b. Disagree
   c. Neither Agree nor Disagree
   d. Agree
   e. Strongly Agree

Section 4: Berlin Numeracy Test

Instructions: Please answer all the math questions that follow. **Do NOT** use a calculator but feel free to use scratch paper for notes.

1. Imagine that we flip a fair coin 1,000 times. What is your best guess about how many times the coin would come up heads in 1,000 flips?
2. Imagine we are throwing a five-sided die 50 times. On average, out of these 50 throws how many times would this five-sided die show an odd number (1, 3 or 5)? (____ out of 50 throws)
3. In the BIG BUCKS LOTTERY, the chance of winning a $10 prize is 1%. What is your best guess about how many people would win a $10 prize if 1000 people each buy a single ticket to BIG BUCKS?
4. In ACME PUBLISHING SWEEPSTAKES, the chance of winning a car is 1 in 1,000. What percent of tickets to ACME PUBLISHING SWEEPSTAKES win a car? (____%)
5. Out of 1,000 people in a small town 500 are members of a choir. Out of these 500 members in a choir 100 are men. Out of the 500 inhabitants that are not in a choir 300 are men. What is the probability that a randomly drawn man is a member of the choir? (____%)
6. Imagine we are throwing a loaded die (6 sides). The probability that the die shows a 6 is twice as high as the probability of each of the other numbers. On average, out of these 70 throws how many times would the die show the number 6? (____out of 70 throws)
7. In a forest 20% of mushrooms are red, 50% brown and 30% white. A red mushroom is poisonous with a probability of 20%. A mushroom that is not red is poisonous with a probability of 5%. What is the probability that a poisonous mushroom in the forest is red? (____%)

Correct Answers: 1) *500* or *50%* or *½* 2) *30 throws* 3) *10* 4) *0.10%* 5) *25%* 6) *20 throws* 7) *50%*

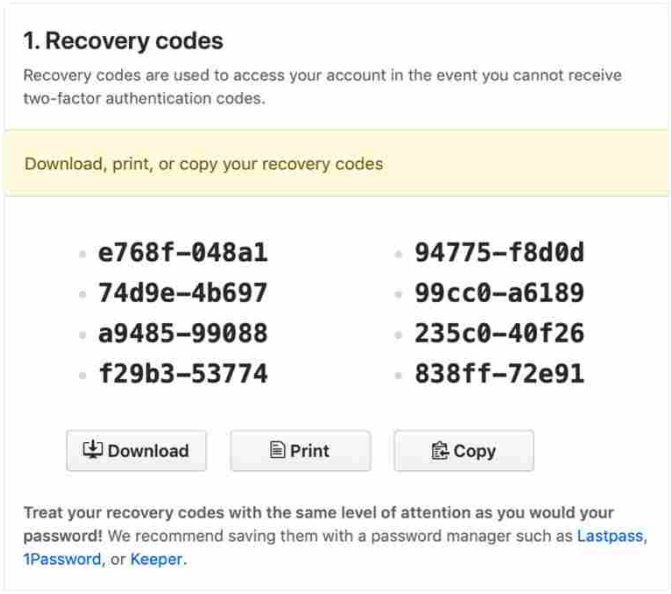
Section 5: Account Recovery Method Descriptions

During the initial setup of two-factor authentication on a personal account, an account recovery method can also be setup as a failsafe in the event of a lost or stolen second factor. When this type of unexpected event occurs, the user would be required to accurately authenticate using their 1st factor alphanumeric password and their account recovery method in order to gain account access. The following information provides a short description of four potential account recovery options. Please read the account recovery descriptions and answer the following questions

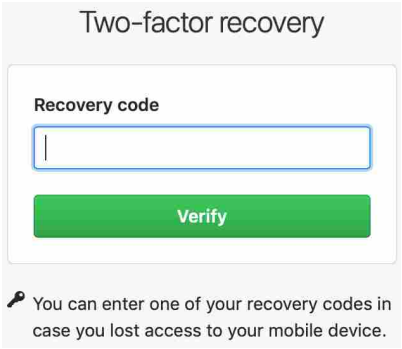
1. ***Generated Recovery Codes***-

During the initial setup of the user's account recovery, the system will generate a unique set of 8 or more different system-assigned account recovery codes. Each recovery code is a 10-character complex string of numbers and letters. Systems often advise the user to do the following: remember at least 1 of the codes; save the codes to an offline server or to an encrypted file; or print a hardcopy and store in a safe place.

During account recovery login, the user is asked to correctly input 1 or more of the recovery codes they received originally. The figures below depict an example of system prompts the user would typically see when setting up or executing account recovery using system generated recovery codes.



Generated recovery codes the user receives when setting up their 2FA account recovery method.



A system prompt requesting the user to enter one of the recovery codes they were assigned when initially setting up their 2FA account recovery method.

2. ***Graphical Scheme Passcode-***

During setup, the system generates a set of 3 picture icons that are assigned to the user as their passcode icons. The system prompts the user to remember their icons.

During account recovery login, the user is asked to recognize and select their passcode icons from a set of distractor icons. Note that the order in which participants select their icons during login often does not matter.

The figures below depict Convex Hull Click (CHC) as an example of a graphical scheme that can be used for account recovery. The user's passcode consists of 3 icons. During login, the user's passcode icons always form a triangle shaped selection region on the grid. To login successfully, the user must click one time anywhere inside the triangular region created by their 3 passcode icons. The same passcode and non-passcode icons are shown during each new account recovery session; however, the icons' location on the grid changes to keep their passcode more secure. Users are also instructed not to hover the mouse cursor over- nor click directly on- a passcode icon.

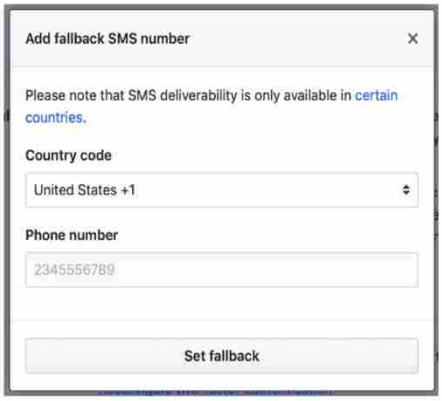
The triangle created by the passcode icons reflects the area that can be clicked in order for the user to successfully login.

The CHC interface that the user would interact with.

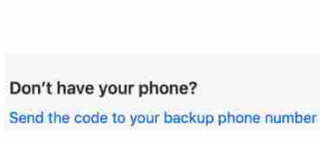
3. ***Fallback Phone Number-***

During the initial setup, the user provides a 10-digit emergency contact phone number (e.g., significant other, family member) where a one-time verification code can be received. Systems often provide the user with the option to have the code sent to the emergency contact via text or phone call. Since a personal cell phone can be used as a medium device for 2FA through an app or SMS, systems often advise the user not to provide their own phone number. Additionally, systems often advise the user to choose a fallback contact phone number that they can recall or obtain without referring to their own personal cell phone (e.g., if your phone is lost, your contact list might not be available).

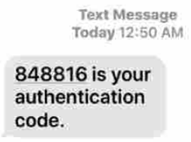
During account recovery login, the system informs the user that a verification code was sent to the fallback phone ending in +x xxx xx01, for example. The user is required to get in touch with the fallback contact to obtain the code. The system then prompts the user to correctly enter the verification code. The figures below depict an example of system prompts the user would typically see when setting up or executing account recovery using a fallback phone contact verification code.



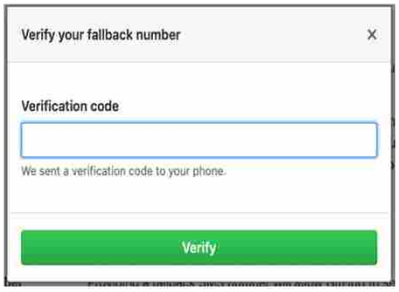
The system prompt that the user receives when setting up a fallback phone number contact as their 2FA account recovery method.



A system link that allows users to send a verification code to their fallback contact.



A verification code received by the fallback contact via SMS.

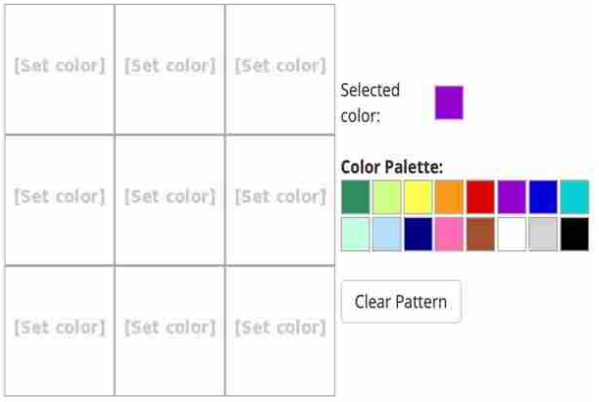


The system prompt requesting the user to enter the verification code that was sent to the fallback contact phone number.

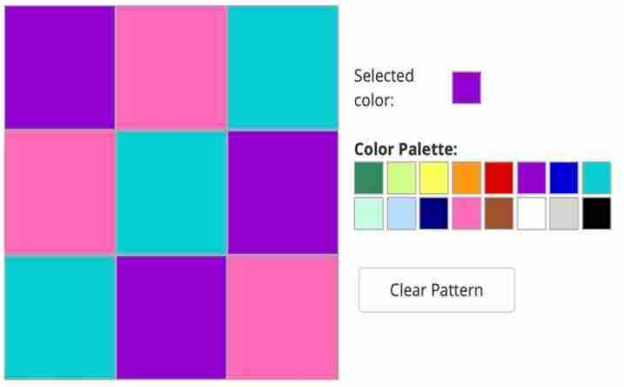
4. ***Color Grid*** –

During setup, the user is asked to create a color grid pattern that they can remember and recreate using 9 squares on a 3x3 grid. To create a pattern, the user chooses colors from a system provided color palette. To increase account security, the system requires that the user chooses at least two different colors.

During account recovery login, users are prompted to recreate their color pattern on the 9 square grid. The figures below depict an example of the interface the user would typically see when setting up or executing account recovery using a color grid pattern.



A blank 3x3 grid ***before*** the user has set up or executed their account recovery color pattern.



A color grid ***after*** the user has set up or executed their pattern for account recovery.

Perception Questions

1. Rank order the recovery methods from safest (1) to least safe (4)
   ____ Generated Recovery Codes

   ____ Graphical Passcode

   ____ Fallback Phone Number

   ____ Color Grid Pattern

2. Rank order the recovery methods from easiest to use (1) to most difficult to use (4)
   ____ Generated Recovery Codes

   ____ Graphical Passcode

   ____ Fallback Phone Number

   ____ Color Grid Pattern

3. Rank order the recovery methods from most memorable (1) to least memorable (4)
   ____ Generated Recovery Codes

____ Graphical Passcode

____ Fallback Phone Number

____ Color Grid Pattern

4. It is vital to our research that we only include responses from people that devoted their attention to this study. In your honest opinion, should we use your data for *this section* in our analyses? (You will receive credit for this study even if you provide a negative response)

    a. Yes
    b. No

Recovery Method Preference

1. If given the option, which account recovery method would you choose to use?
    a. Generated Recovery Codes
    b. Graphical Passcode
    c. Fallback Phone Number
    d. Color Grid Pattern
2. Why you would prefer to use this recovery method? Please explain.

Participant Debriefing Statement

Many companies are implementing Two-Factor Authentication (2FA) to increase the security of user accounts. When 2FA issues (e.g., 2nd factor is lost, stolen, or unavailable) prevent an authorized user from gaining account access, an alternate account recovery option that does not require a downloaded software or registered device should be available. Essentially, an account recovery option is an account feature that can be made available to 2FA users and it is set up before losing a second-factor device. The purpose of this study is to gain a comprehensive understanding of the novice end user's: general authentication knowledge; behaviors; risk literacy abilities; and measure their preference and perceived security of different 2FA account recovery methods. Researching account recovery is important because literature has shown that current account recovery processes for 2FA can be cumbersome and some companies' caution that it may take several business days for the user to regain account access. We predict the results of this study will help us determine which 2FA account recovery methods are preferred, and which are perceived as –secure and –memorable. Our goal is to offer the research community a comparison of the attributes of the different account recovery methods and provide website designers with better direction when selecting prospective account recovery options.

If you have any questions or comments, feel free to contact the researcher using the contact information that is given for this study in the SONA system. We plan to publish the results within the next year and they will be available for viewing at http://www.psychofdesign.com/publications.htm .

Thank you for your participation!

APPENDIX B: INFORMED CONSENT DOCUMENT

OLD DOMINION UNIVERSITY

PROJECT TITLE: Account Recovery for Two-Factor Authentication

INTRODUCTION

Please take your time in deciding if you would like to participate in this study by reading this notice carefully. This page will record your consent to participate in this study, "Account Recovery for Two-Factor Authentication," on Qualtrics through the SONA research system.

DESCRIPTION OF RESEARCH STUDY

This study aims to gain a comprehensive understanding of the typical knowledge and behaviors regarding authentication. We are exploring the idea of two-factor authentication (2FA) account recovery methods. An account recovery method is a failsafe for 2FA in the event of a lost or unavailable second factor. We want to help to increase the security of user privacy data while still allowing the authorized user account access. Our goal is to determine if the typical end-user knows the concepts and threats associated with authentication, and to measure security feature preference for account recovery.

You will be asked to complete a 42-question survey. The survey is composed mainly of multiple questions with only a few fill-in-the-blanks. Math will be required to answer seven of the survey questions. For each item you should select the answer(s) that you think is best based on your knowledge, experiences, and opinions. The survey should take about 30-45 minutes to complete.

There will be approximately 400 participants completing the Qualtrics survey through the SONA research system.

EXCLUSIONARY CRITERIA

To participate in this study, you must be at least 18 years old and you must have corrected to normal vision.

RISKS AND BENEFITS

RISKS: If you decide to participate in this study, there is a potential risk of eye strain from interacting with a computer-based system. This risk is similar to typical computer usage. And, as with any research, there is some possibility that you may be subject to risks that have not yet been identified.

BENEFITS: There are no immediate benefits to participants for participating in this study. The overall benefit will be to help improve authentication systems in terms of usability and security.

COSTS AND PAYMENTS

Your decision to participate in this study must be voluntary. And, we recognize that your participation, poses some inconveniences. Therefore, you will receive 1 research credit through the SONA system for participation.

CONFIDENTIALITY

The results of the study will not be associated with you in any way. No records are kept that allow your name to be associated with your responses in the study or on the survey. Your responses will be anonymous. The outcome of this research may be used in reports, presentations, and publications. But, again we will not identify you personally. Of course, your records may be subpoenaed by court order or inspected by government bodies with oversight authority.

WITHDRAWAL PRIVILEGE

Your participation in this study is completely voluntary and you may refuse to participate. If you agree to participate, you have the right to stop at any time or the right to skip any survey question that you do not wish to answer.

COMPENSATION FOR ILLNESS AND INJURY

If you say agree to participate, then your consent does not waive any of your legal rights. However, in the event of any harm arising from this study, neither Old Dominion University nor the researchers are able to give you any money, insurance coverage, free medical care, or any other compensation for such harm. In the event that you suffer some type of harm as a result of participation in any research project, you may contact Dr. Jeremiah Still at 757-683-6424, Dr. Tancy Vandecar-Burdin the current IRB chair at 757-683 3802 at Old Dominion University, or the Old Dominion University Office of Research at 757-683-3460 who will be glad to review the matter with you.

VOLUNTARY CONSENT

By continuing with the study, you are saying several things. You are saying that you have read this information, that you are satisfied that you understand this information, the research study, and its risks and benefits. If you have any questions later on, then the researchers should be able to answer them: Dr. Jeremiah Still at 757-683-6424.

If at any time you have any questions about your rights, then you should call Dr. Tancy Vandecar-Burdin, the current IRB chair, at 757 683 3802, or the Old Dominion University Office of Research, at 757 683 3460.

Electronic Consent

By clicking the “I agree” button you indicate that you meet the study’s requirements and consent to the study. You may print a copy of this screen for your records.

APPENDIX C: RENAUD'S (2007) OPPORTUNISTIC SCORE SPECIFICATIONS

Opportunistic Scoring Rubric Given Research Assistant to Assess Account Recovery Methods

Opportunist Category	Question(s) or Prompt(s)	Account Recovery Methods			
		Fallback Phone Number	Generated Recovery Codes	Graphical Passcode	Color Grid Pattern
Guessability	This is the traditional measure of strength of an authentication key: the size of the dictionary space. The guessability of a four-digit PIN is 1 in 10,000 since there are 10,000 four-digit numbers to choose from hence any key that is as strong or stronger than this is assigned a 0. Weaker keys will be assigned a proportionally higher guessability figure.				

Opportunist Category	Question(s) or Prompt(s)	Account Recovery Methods			
		Fallback Phone Number	Generated Recovery Codes	Graphical Passcode	Color Grid Pattern
Recordability	As regards recordability, the systems can be assigned values as follows: - 1 if the code is easily recorded - 0.5 if it was harder to record or describe, or if recording of the key does not provide an observer with the full key - 0 if it is difficult or impossible to record or describe, such as, for example, a biometric It should be noted that recordability is an extremely difficult weakness to counteract, mainly due to inbuilt operating system features such as the “Print Screen” button and browser print functionality, which allows the user to print the authentication screen and mark off the required images to offset memory lapses.				

Opportunist Category	Question(s) or Prompt(s)	Account Recovery Methods			
		Fallback Phone Number	Generated Recovery Codes	Graphical Passcode	Color Grid Pattern
Observability	*Observation of the code involves two equally important features:* (a) being able to actually see the key on the screen, and to use it—we assign: - 0.5 if the key can be used if observed only once to obtain the full key - 0.25 if key entry needs to be observed multiple times to obtain the key - 0 if the key cannot be observed (b) being able to judge the position of the key based on where the person is pointing at the screen or on the keyboard — we assign: - 0.5 if observation of the key location is meaningful - 0 if not				

Opportunist Category	Question(s) or Prompt(s)	Account Recovery Methods			
		Fallback Phone Number	Generated Recovery Codes	Graphical Passcode	Color Grid Pattern
Analyzability	(a) Usernames — users should be assigned usernames rather than email addresses because email addresses are too easily obtainable and make it easier for hackers to gain access to the system — we assign: - 1 if the system uses email addresses as usernames - 0 if unique usernames are used and not visible to other users.	**0**	**0**	**0**	**0**

Opportunist Category	Question(s) or Prompt(s)	Account Recovery Methods			
		Fallback Phone Number	Generated Recovery Codes	Graphical Passcode	Color Grid Pattern
Analyzability	(b) Error messages — error messages need to be provided on various levels. The developer obviously needs a different kind of error message that will enable him/her to analyze problems with the web site. Once the site is deployed, however, the messages should become targeted at the needs of the user, and no longer inform as to actual failure codes or database errors, but rather in terms of actions the user needs to take to recover. This limits the usefulness of error messages to the potential intruder. Hence — we assign: - 0 if error messages have been tailored in this way - 1 otherwise.	**0**	**0**	**0**	**0**
	(c) Default Keys — this particularly bad practice earns a rating of 1 because many users will not redefine their key or an intruder can take advantage of the default setting before the user logs in for the first time.	**0**	**0**	**0**	**0**

Opportunist Category	Question(s) or Prompt(s)	Account Recovery Methods			
		Fallback Phone Number	Generated Recovery Codes	Graphical Passcode	Color Grid Pattern
Analyzability	(d) Forced changes — the reasoning behind this is that a leaked authentication key will only be useful to an intruder for a limited period of time. However, routine forced renewals actually decrease guessability since users need to come up with new passwords every time it is renewed, and they eventually start choosing easy-to-remember passwords. Hence an application with this policy earns a 1.	**0**	**0**	**0**	**0**

Opportunist Category	Question(s) or Prompt(s)	Account Recovery Methods			
		Fallback Phone Number	Generated Recovery Codes	Graphical Passcode	Color Grid Pattern
Analyzability	(e) Key retrieval — Forgotten keys should never be emailed. This is simply too easy for an intruder to intercept. The current practice of asking the user to confirm the answer to a particular question reduces the authentication key space to a very small space indeed and one that can probably be uncovered by a research-based attack. A better mechanism is to reset the password and email the user a secure link, which requires the user to set a new password. For a secure site a more secure option may be required, such as, perhaps, sending an SMS message to the user and requiring her to confirm the request via SMS before the secure link is emailed. A policy that emails authentication keys or uses confirmation questions to confirm identity earns a 1 for analyzability.	**0**	**0**	**0**	**0**

Opportunist Category	Question(s) or Prompt(s)	Account Recovery Methods			
		Fallback Phone Number	Generated Recovery Codes	Graphical Passcode	Color Grid Pattern
Analyzability	(f) Backward browsing—intruders can often try to obtain information by using the back button, which is impossible to disable. Hence, we will assign 0 only if the system ensures that authentication pages expire immediately after they are processed. A weakness of 1 will be assigned otherwise.	**0**	**0**	**0**	**0**

Opportunist Category	Question(s) or Prompt(s)	Account Recovery Methods			
		Fallback Phone Number	Generated Recovery Codes	Graphical Passcode	Color Grid Pattern
Analyzability	(g) Choice of distractor images (if applicable) — some recognition-based authentication mechanisms rely on the user choosing one image from a group of distractor images. If distractor images are varied at each attempt it is a simple matter for the intruder to observe the interaction over an extended period of time to identify the target images, or to refresh the display repeatedly. It is more secure to fix distractors for a particular user and to use these repeatedly. — we assign: - 0 if the scheme follows the aforementioned policy - 1 if there is a policy of varying distractors - NA if this type of policy is not applicable	_____ Please explain why you chose this "choice of distract-or images" score?	_____ “ ”	_____ “ ”	_____ “ ”

Opportunist Category	Question(s) or Prompt(s)	Account Recovery Methods			
		Fallback Phone Number	Generated Recovery Codes	Graphical Passcode	Color Grid Pattern
Analyzability	(h) Choice of background image (if applicable) — some mechanisms make use of a single large image, which needs to have particular characteristics: In this case the image needs to have many features which can be chosen by the user, but not too many, which could cause confusion. Hence — we assign: - 1 to an image with fewer than 10 identifiable features - 0 to an image with more than 1000 features (Numbers in between are assigned on a proportional basis) - NA if choice of background image does not apply to this authentication scheme	_____ Please explain why you chose this "choice of backgr-ound images" score?	_____ “ ”	_____ “ ”	_____ “ ”
Resistibility	(a) Lockout policy (i.e. only allows a set amount of login attempts)— we assign: - 1 if there is a strikeout policy - 0 otherwise	**1**	**1**	**1**	**1**

Opportunist Category	Question(s) or Prompt(s)	Account Recovery Methods			
		Fallback Phone Number	Generated Recovery Codes	Graphical Passcode	Color Grid Pattern
Resistibility	(b) Key Strength — stronger keys are less prone to brute-force attacks so many systems enforce password policies that require passwords to have a specific length, a digit, upper- and lower-case letters and special characters. Unfortunately, people cannot remember long and complicated strings and this increases the likelihood that it will be recorded. A much better way of strengthening a key is by using length rather than complication. For example, password users can be encouraged to write a whole sentence rather than a simple word. — we assign: - 0 if a key complicating policy is applied - 1 if a key lengthening policy is applied - NA if this policy cannot be applied	_____ Please explain why you chose this "key strength " score?	_____ “ ”	_____ “ ”	_____ “ ”

Opportunist Category	Question(s) or Prompt(s)	Account Recovery Methods			
		Fallback Phone Number	Generated Recovery Codes	Graphical Passcode	Color Grid Pattern
Resistibility	(c) Timeouts—If an authentication takes too long it is likely that an intruder is trying to determine which the target images are by doing some kind of research. The legitimate user can be expected to home in on her images very quickly. Hence a time limit should be applied to the authentication step. — we assign: - 1 if there is a policy - 0 otherwise	**1**	**1**	**1**	**1**
	(d) Auditing (i.e., System administrators spend time scanning the information logs of the authentication mechanism to check for hacker activity) — we assign: - 0 if no regular auditing takes place - 1 if auditing takes place at weekly intervals - 2 if it occurs more often than that	**2**	**2**	**2**	**2**

Opportunist Category	Question(s) or Prompt(s)	Account Recovery Methods			
		Fallback Phone Number	Generated Recovery Codes	Graphical Passcode	Color Grid Pattern
Resistibility	(e) Evidence— we assign: - 1 if previous login attempts are displayed to the user at login time - 2 if the user is apprised by email or SMS when someone logs into their account - 0 if no historical data is provided	**2**	**2**	**2**	**2**
	(f) Ease of change — if you make it easy for users to change their authentication keys, they are more likely to do so. Hence — we assign: - 1 if this is easy to do, but only if they have to authenticate themselves before changing it.	**1**	**1**	**1**	**1**
	Total Opportunistic Score				

Note. All questions and prompts were directly quoted from Renaud (2007). The bolded scores reflect the assigned scores that suggested the system was implementing the best attack resistant strategy for the prompts that regarded the authentication security protocol of the system itself.

VITA

Lauren Nicole Tiller

Old Dominion University Department of Psychology, Norfolk, VA 23529

Email: lauren.tiller@icloud.com
Website: lntilleruxresearcher.wixsite.com/portfolio
LinkedIn: lauren-n-tiller-ux-researcher

Education

Old Dominion University, Norfolk, VA
M.S., Experimental Psychology – Human Factors Concentration 2018 – 2020
B.S., Psychology (Statistics Minor; *Magna Cum Laude*) 2016 – 2018

Tidewater Community College, Norfolk, VA
A.A.S., Humanities & Business Administration (*Summa Cum Laude*) 2014 – 2015

Rudy & Kelly Academy a Paul Mitchell Partner School, Virginia Beach, VA
Cosmetologist License 2009 – 2010

Research Experience

Graduate Researcher, Old Dominion University
Psychology of Design (PoD) Lab, Advisor: Dr. Jeremiah Still 2016 – Current

Undergraduate Research Assistant, Old Dominion University
Research Environment for Alarm & Complex Task Simulation (REACTS) Lab, 2016 – 2018
Principle Investigator: Dr. James Bliss

Patent

Tiller, L. N., Angelini, C. A., Leibner, S. C., & Still, J. D. (2019, August; under review). *System and method incorporating graphical aids for the creation and retrieval of alphanumeric passwords.* United States Patent and Trademark Office, Nonprovisional Application NO. US 16550450.

Publications

Tiller, L. N., Angelini, C. A., Leibner, S. C., & Still, J. D. (2019). Explore-a-Nation: Combining graphical and alphanumeric authentication. In *A. Moallem (Ed.), HCI for Cybersecurity, Privacy and Trust*, HCII, *11594*, 81-95.

Mator, J. D., Lehman, W. E., McManus, W., Powers, S., **Tiller**, L. N., Unverricht, J. R., & Still, J. D. (2020). Usability: Adoption, Measurement, Value. Human Factors.

Tiller, L. N., Cain, A. A., Potter, L. N., & Still, J. D. (2018). Graphical authentication schemes: Balancing amount of image distortion. *Proceedings of the International Conference on Applied Human Factors and Ergonomics*, 88-98.

Bliss, J. P., & **Tiller**, L. N. (2018). An examination of close calls reported within the International Association of Fire Chiefs database. *Proceedings of the International Conference on Applied Human Factors and Ergonomics*, 184-194.

Tiller, L. N., & Bliss, J. P. (2017). Categorization of near-collision close calls reported to the Aviation Safety Reporting System. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *61*(1), 1866-1870.

Cain, A., Griner, J., **Tiller**, L. N., Unverricht, J., & Still, J. (2017). Comparing measurement approaches of over-the-shoulder-attack resistance for graphical authentication. *Poster Session Presented at the Southeastern Human Factors Applied Research Conference*, Raleigh, NC.