Spring 2015

Password strength analysis: User coping mechanisms in password selection

Brian Thomas Curnett
Purdue University

Graduate School Form 30
Updated 1/15/2015

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Brian Thomas Curnett

Entitled

PASSWORD STRENGTH ANALYSIS: USER COPING MECHANISMS IN PASSWORD SELECTION

For the degree of Master of Science

Is approved by the final examining committee:

Melissa Dark
Chair

Christopher Foreman

Jeffery Karpicke

Robert Proctor

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Melissa Dark

Approved by: Eugene Spafford 4/24/2015

Head of the Departmental Graduate Program Date

PASSWORD STRENGTH ANALYSIS: USER COPING MECHANIMS IN PASSWORD SELECTION

A Thesis

Submitted to the Faculty

of

Purdue University

by

Brian Thomas Curnett

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

May 2015

Purdue University

West Lafayette, Indiana

Dedication

I dedicate this thesis to my parents for their unwavering support and endless love.

ACKNOWLEDGEMENTS

I would like to give special thanks to my friend, Miguel Cedeño Agamez for his technical support in developing the website used for data collection which has made this thesis possible. In addition, I would like to thank Teri Flory and Paul Duselis for their help as partners while working on password strength analysis in the Information Security Research and Education (INSuRE) project.

I would like to thank Professor Melissa Dark for making the Information Security Research and Education (INSuRE) program possible, and for her guidance, patience, and advice throughout this process.

TABLE OF CONTENTS

Page
LIST OF TABLES vi
LIST OF FIGURES vii
ABSTRACT viii
CHAPTER 1: INTRODUCTION 1
1.1. Background 1
1.2. Significance 2
1.3. Statement of Purpose 2
1.4. Research Questions 3
1.5. Assumptions 3
1.6. Limitations 3
1.7. Delimitations 4
1.8. Definitions 4
CHAPTER 2: LITERATURE REVIEW 6
2.1. Passwords: An Origin and why we need them after all these years 6
2.2. Entropy: How to Measure Passwords 7
2.3. Threats to Passwords 9
2.4. Cognitive Psychology Aspects of Memory: Recall 11
2.5. Self Generation 12
2.6. Cueing – Context Dependency 12
2.7. Chunking 14
2.8. Proactive - Interference 15
2.9. Psychological Coping Mechanisms 15
2.10. Perspectives on Password Cracking 17
CHAPTER 3: METHODOLOGY 19
3.1. Methodology and Variables 19
3.2. Procedure 20
3.3. Description of Variables 21
3.3.1 Types of Coping Mechanisms 21
3.3.2 Entropy Measures 23
3.4. Demographics 26
3.5. Sample 27
3.6. Bias 27
3.7. Approval 28
3.8. Data Collection and Analysis 28

Page
CHAPTER 4: Results AND ANALYSIS 29
4.1. Overview 29
4.2. Policy Analysis 31
4.3. Sample Representativeness 34
4.4. Comparing Policies 35
4.5. Prevalence and Extent 36
CHAPTER 5: CONCLUSION 41
5.1. Research Question 1 41
5.2. Research Question 2 42
5.3. Research Question 3 43
5.4. Recommendations for Future Work 44
LIST OF REFERENCES 46
APPENDICES
Appendix A. Survey Questions 49
Appendix B. Institutional Review Board Application 61
Appendix C. Statistical Analysis System (SAS) Code 69

LIST OF TABLES

Table Page
Table 1: Significance in NIST Model by Policy by Iteration 31

Table 2: Average Length of passwords 32

Table 3: Significance in Post Coping Model by Policy by Iteration 33

Table 4: Sample Representativeness 2 sample t-testing 34

Table 5: Post Coping Analysis by Policy by Iteration 35

Table 6: Quantity of Coping Mechanisms by Policy by Iteration 36

LIST OF FIGURES

Figure Page
Figure 1: THC Hydra Interface 24

Figure 2: NIST Entropy Average 30

Figure 3: Post Coping Entropy Average 30

Figure 4: Prevalence of Coping Mechanisms 37

Figure 5: Extent of Coping Mechanisms in Comprehensive 8 38

Figure 6: Extent of Coping Mechanisms in Blacklist Hard 38

Figure 7: Extent of Coping Mechanisms in Basic 16 39

ABSTRACT

Curnett, Brian T. M.S., Purdue University, May 2015. Password Strength Analysis: User Coping Mechanisms in Password Selection. Major Professor: Melissa Dark.

The security that passwords provide could be seriously flawed due to the way people cope with having to memorize and recall their passwords. The National Institute of Standards and Technology (NIST) standard that is used to measure the password strength, known as entropy, is designed for a single use and does not consider that users may choose to keep parts of their password across password changes. This study shows that a portion of users keep some information from previous passwords across changes. These habits which will be called coping mechanisms that over time serve to erode the protection provided by passwords past the minimum level of security provided by the password policy which can place both individuals and enterprises into danger. This is made even more apparent with data breaches become a common phenomenon in present day life serving to expose user's password to the world. It was found that the minimum level of security can no longer be provided after one disclosure of passwords in the Comprehensive 8 password policy, and after two disclosures in passwords in the Blacklist Hard and Basic 16 policy. Coping mechanisms are most prevalent in password policies that have many requirements placed on users. The Comprehensive 8 policy showed the most coping followed by the Blacklist Hard and Basic 16 policies.

CHAPTER 1: INTRODUCTION

This chapter provides an introduction and overview of the research project. The introduction includes a background of the research, the significance of the research, the research questions, assumption, limitations, delimitations, and the definitions of terminology that will be used throughout.

1.1. Background

Even though passwords have been used for as long as people have needed authentication, password strength analysis is an ongoing subject of research. There are many viewpoints that can be taken on the security of passwords from the cryptographic approach which has resulted in the hashing and salting of passwords to the network administrators viewpoint of policies and standard requirements both becoming standard practice (Klein, 1990 ; Zviran & Erlich, 2006). There is not currently a consensus within the field of information security as to "the best" security mechanism. This leads to many arguments as to how to best protect many systems. Passwords still remain a relatively inexpensive option for authentication and appear to be here to stay (Furnell & Zekri, 2006). While passwords are a dominant authentication measure, weaknesses must be addressed in order to protect all of the users whose information depends upon the security that passwords provide. A password is only as strong as the user creates it to be. Weaknesses

need to be examined in this creation process to see how much users cope with policy restrictions by using known information, partially repeating past passwords, and using easily identifiable patterns. Then these coping mechanisms can be examined to determine how they can serve to erode the strength of password policies over time.

1.2. Significance

This study will concentrate on identifying the user habits that will be identified as coping mechanisms. Most studies in the field of password strength analysis will take a large set of passwords from a hacker's leaked list, but this represents only a snapshot in time of all users' passwords. This study is separated by the fact that it will analyze user behavior across several iterations of time. This way, the evolution of the password will be tracked to see the password's susceptibility given prior knowledge by an adversary such as a username and at least one password.

1.3. Statement of Purpose

This study will focus on identifying coping mechanisms that will remain within users' passwords across password change. These coping mechanisms will be added to the NIST model of entropy in an attempt to show the reduction of the effective key space that these coping mechanisms will cause in the password as time progresses.

1.4. Research Questions

1. What is the prevalence of the coping mechanisms to which users engage by policy and across iterations?
    - Using known information
    - Partial repetition
    - Easily identifiable patterns
2. To what extent do users engage in coping mechanisms by policy and across iterations?
3. Do user coping mechanisms decrease the strength of password policies to such a degree that the increase of security sought by the stringent policy is negated?

1.5. Assumptions

1. The participants will respond truthfully during survey questions.
2. At the start of each iteration all previous iterations passwords as well as the username will become known.

1.6. Limitations

1. Participants in this study live in the United States limiting the applicability across the world. This may be relevant if coping habits are structured differently in other languages and in other regions the primary language may not be English.
2. The sample population of Mechanical Turk users may not be representative of the average user as no demographic information is available for this population.

3. No personally identifiable information will be collected in this study. If any coping mechanisms are based in using personally identifiable information, names, addresses, birth dates, etc., these mechanisms will not be identified in this study.

1.7. Delimitations

1. Only seven iterations of password changes will be collected.
2. Passwords that are reused across accounts cannot be accounted for within the data set.

1.8. Definitions

Coping Mechanisms – Any behavior that serves to transfer information across password changes or to reduce the effective key space of a password.

Basic 16 – A password policy that asks the user to create a password of at least 16 characters in length (Kelley, et al., 2012).

Blacklist Hard – A password policy that requires the user to create a password of at least 8 characters in length. This password is then checked against a dictionary blacklist and if the password is on the dictionary blacklist it requires the user to choose a new password (Kelley, et al., 2012).

Comprehensive 8 – A password policy that requires a password to be at least 8 characters long and contain at least 1 lowercase letter, 1 capital letter, 1 number, and 1 special character (Kelley, et al., 2012).

Dictionary blacklist – A collection of commonly used words that are banned from use as passwords.

Entropy – The standard unit of measurement for password strength.

Memorability – The ease that a user can remember a password.

Password – A set of characters known by a user and a system used to authenticate the identity of a user.

Password Policies – The minimum requirements placed on the user.

Post Coping Mechanism Entropy – The entropy of a password after coping mechanisms have been analyzed for their effect on security.

CHAPTER 2: LITERATURE REVIEW

This literature review examines a collection of disciplines in order to provide the sufficient background information necessary for research into password strength analysis. This chapter includes sections on the origin of passwords, entropy, human memory, and coping mechanisms. The first section describes passwords' origins, associated terminology, and metrics. The second gives a cognitive psychology perspective on how the human mind will view passwords and the associated phenomena. The final section details the what, why, and how we use coping mechanisms.

2.1. Passwords: An Origin and why we need them after all these years

Passwords can trace their lineage back to the ancient Roman military. Then known as watchwords, passwords were used to verify the identity of troops (Eaton, 2011). The reason we need passwords hasn't changed in 2000 years. Confidentiality, Integrity, and Availability, also known as the CIA spectrum in information security, are still needed to insure that the person accessing a piece of information is indeed the person who should be accessing the information. Passwords fit well within the CIA spectrum. Passwords help keep data confidential by ensuring that only the people who know the secret, a password, have access to the data. Passwords only work if they themselves

remain confidential. By ensuring only people who have the appropriate clearance have access to the data serves to keep the data's integrity intact. Passwords function by denying availability to people who do not know the secret password. However, if a password is forgotten, it can easily deny availability to the intended user. Traditionally this has been done by using one or more of three authentication techniques; something you have, something you are, or something you know. Debates about which of these techniques is the best or which should be used together are ongoing and not the subject of this literature review. This paper solely focuses on how passwords fall into the last category of authentication, something you know, and how certain coping mechanisms erode the security that they provide.

2.2. Entropy: How to Measure Passwords

Entropy in units of bits is the measurement of the strength of a password. A simple definition of entropy provided by Shay, Komanduri, Kelley, Leon, Marzurek, Bauer, Cristin, and Cranor (2010) is that it is the "measure of the difficulty of guessing a password." Originally, entropy in Information Theory, introduced by Claude Shannon (1948) simply measured all possible passwords that could be contained within a key space. So, a password's information entropy, H, can be found by the following formula

$$H = L\frac{logN}{log2}$$

Where L, Length, is the number of characters in the password, and N is the number of possible symbols in each character's slot. The key space can be found by simply taking 2 to the power of the entropy of the password to find the total number of possible

passwords. The post coping entropy measurement used later can be converted the same way in order to determine the number of attempts needed for a password cracker to find the password. When using entropy as a measure of password strength in this method it assumes that all passwords within a key space are equally likely. When dealing with human generated passwords this assumption is broken completely. In 2004, Burr, Dodson, and Polk at the National Institute of Standards and Technology (NIST) defined entropy as "A measure of the amount of uncertainty that an attacker faces to determine the value of a secret." This publication also known as NIST Special Publication 800-63 provided a model to "correct" human generated passwords to a more accurate entropy (Burr, Dodson, & Polk, 2004).

> NIST Special Publication 800-63 of June 2004 suggests the following scheme to roughly estimate the entropy of human-generated passwords
>
> - The entropy of the first character is four bits;
> - The entropy of the next seven characters are two bits per character;
> - The ninth through the twentieth character has 1.5 bits of entropy per character;
> - Characters 21 and above have one bit of entropy per character.
> - A "bonus" of six bits is added if both upper case letters and non-alphabetic characters are used.
> - A "bonus" of six bits is added for passwords of length 1 through 19 characters following an extensive dictionary check to ensure the password is not contained within a large dictionary. Passwords of 20 characters or more do not receive this bonus because it is assumed they are pass-phrases consisting of multiple dictionary words.

This set of policies put forward by NIST serves as a foundation for most modern commercial password systems and allows a system administrator an easy metric to design the password policy for their organization. In order to create policies that will hopefully

cause strong password creation by users, organizations use this NIST entropy to determine the strength of a password. Examples, of password policies would be Comprehensive 8, Blacklist Hard, and Basic 16. (Kelley, et al., 2012) Where the minimum entropy in these policies are 24 bits, 24 bits and 30 bits respectively. Thus, the greater the entropy, the more difficult it is for a hacker to predict the value of a variable, and therefore gain access to the protected information. All of this information is taken into consideration when forming a password policy.

2.3. Threats to Passwords

There are many types of external threats to a password, social engineering, physical intrusion, password guessing, and password cracking. Each of these threats to passwords has a place in risk analysis for security. Threats such as social engineering are handled by training a workforce in best practices with passwords. Physical intrusion is mitigated by having a security policy of locking doors or depending on the organization, physical security. (Sarkar, 2004) Password guessing comes in several different forms, including brute force attacks where the attacker tries to guess the password by using all possible combinations of characters, and dictionary attacks where the attacker uses a list of words to try to guess the password, and the attacker searches for the user's personal information, such as birthdays and names, to attempt to guess the correct password, and rule based attacks where an attacker defines a list of rules toward the creation of a wordlist for example at least 8 letters long contains 1 number, etc... (EC-Council, 2009) Password cracking attempts to create a string of hashes that has the same encrypted hash as the password. The password created by the new hash may or may not be the same as the users, but because

the hash matches, the attacker is granted access. The other method of compromise is when a hacker compromises an organization's server and downloads the usernames and passwords of all users. (Albanesius, 2011) This may be detected by an organization's defenses in which case the organization will instruct users to change their passwords. But an attacker's knowledge of a previous password could still have implications on the passwords through the predictable ways that users cope in creating passwords.

A secure password should be able to withstand these attacks until the next scheduled password change. Different password cracking algorithms may be used with varying degrees of effectiveness. Kelley et al. (2012) further expanded on the research by Komanduri et al. (2011), in utilizing a leaked password list and the data collected in the Komanduri study (Kelley, et al., 2012; Komanduri, et al., 2011). The analysis reviewed both the ease at which a password could be cracked and entropy of passwords and allowed for an infinite number of guesses on each password. It was found that though NIST considers the password policies known as Basic 16 and Comprehensive 8 equivalent, Kelley et al. (2012) found that the Basic 16 was the more secure against a large number of guesses. Basic 16 requires the password to have at least sixteen characters, and Comprehensive 8 requires a password with at least eight characters, an uppercase, a lower case, a symbol, and a digit. In addition, Comprehensive 8 could not contain a dictionary word (Kelley, et al., 2012; Komanduri, et al., 2011). The work done by NIST was an enormous step forward in providing a quantitative measure for the strength of the password away from using the same formulas for randomly generated passwords. However, the work done by Kelly et al. and Komanduri et al. serve to highlight some remaining flaws with

this system for quantitative analysis of passwords. This paper will serve to further highlight flaws in the way the NIST model calculates entropy.

2.4. Cognitive Psychology Aspects of Memory: Recall

When examining memory the question that is usually asked first is "How much can someone remember?" Turning to cognitive psychology the traditional and most widely accepted answer is "Seven, Plus or Minus Two". This is what is called Miller's Law, coming from George Miller's (1956) seminal publication determining "that the average human can hold 7 ± 2 objects in working memory. It is well established within psychology that this is the limit of human memory. At first glance this is easily applicable to passwords and many take this statement at face value and most passwords fall in the 8 character range. In addition, this rule is only applicable to the short term memory instead of long term memory where passwords should be stored to be memorable. The concept of "Seven, Plus or Minus Two is used in psychology for recall tasks where a researcher gives information to a participant not information that the participant generates themselves. So this will be applicable to randomly generated passwords generated by a policy and given to a user not user generated passwords. The concept of the "Seven, Plus or Minus Two" Rule will not be utilized further due to the fact that all passwords in this study will be generated by respondents as opposed to randomly generated and that passwords will be part of user's long term memory rather than their short term memory. However, there are other principles of cognitive psychology that can be used in order to extend human memory for the purposes of password generation.

2.5. Self Generation

Cognitive Psychology does not provide much information in this area as most of the memory recall studies are performed by the researcher giving a participant information to recall later rather than the participant producing the information. However, research by Mulligan et al. show that "generation enhances memory for the occurrence of items" (Muligan, Lozito, & Rosner, 2006). This research showed that memory effects were enhanced but with a few tradeoffs, which are not relevant to passwords such as text color and font. Vu, Bhargav, and Proctor eventually applied this directly to passwords using the number of login attempts as a measure of the generation. In Vu's case, the mean time for log in was 33.7 seconds with 3.77 attempts made before successful login (Vu, Bhargav, & Proctor, 2003). Vu et al go on to show that patterns involving numbers and special characters can increase the memorability of passwords but that this comes at the cost of increased predictability of password structure. This can be shown to indicate the heavy load that passwords make on human memory and Vu et al. go on to document that there is a shown tradeoff between security and memorability in most password policies. This heavy load on memory leads some users to write down their passwords especially in cases where lockouts are used (Gehringer, 2002). Gehringer goes on to suggest methods to mitigate this coping mechanism by creating passwords that are anagrams of longer phrases.

2.6. Cueing – Context Dependency

It has been shown in cognitive psychology, when context is the same while both learning a subject and testing recall, the human memory will be much more effective and accurate in recall (Godden & Baddeley, 1975). In addition, by maintaining a consistently

themed environment at every interaction not just while learning and testing, it is shown that memory performance is increased (Pessin, 1932). Adding to this, Smith and Vela showed that simply visiting "an environmental context acts as a cue for past memories related to that particular environmental context" (Smith & Vela, 2001). In the context of passwords this means that the user interface should be kept, specifically the log in, consistent throughout the study as this will affect our participant's ability to remember their passwords. It should be noted that this context dependency will in no way affect the security of a current password only the ability to remember it in the future. This will also mean that the user can bring in habits from other experiences with similarly themed designs. In order to not incur any additional coping a unique interface for our website was created from scratch. This should help to prevent any cueing from commonly used open source interfaces. To limit the loss of memorability of the password, the interface of the website will not be changed throughout the duration of the study so as to help cue the user of their password.

The implications of drawing from cognitive psychology literature for this study is that there will be a deeper understanding of why each password was made in comparison to most studies of passwords, which tend to use an overarching analysis strategy such as Markov chains in which some information, like longer patterns, may be lost (Yiannis, 2013). At the same time, much of the cognitive psychology literature can be misused as well due to the fact that many of the recall and cueing tasks researched in the cognitive psychology realm are based on information that is generated by researchers and given to participants rather than generated by participants themselves. This makes it more difficult

to draw on each field without first considering how and by whom the information in the subsequent research was generated.

2.7. Chunking

Chunking is the concept in cognitive psychology that refers to the human mind's ability to better organize, store, and recall information better by storing it as several pieces rather than a whole string of information. Using the concept of chunking allows what appears as individual characters within passwords or a whole string to become meaningful words stored in the human mind as one concept rather than by their individual characters (Miller, 1956). By gathering characters together into meaningful groupings (i.e. chunks) the tax on memory can be kept the same or reduced whilst increasing the key space and entropy of the password. By considering the user of a password policy has certain known strengths and weaknesses in memory retrieval and then catering to these strong traits and minimizing policy reliance on weaknesses in human memory, password entropy can be greatly increased without increasing the strain on user memory.

This can be shown in the structure of passwords; for example when passwords are required to contain numbers, it is shown that a significant portion of people will place a set of numbers and/or symbols at the end of the password (Komanduri, et al., 2011). While this reduces the tax on human memory that the password takes, some of the strategies people use to cope using chunking can be exploited by password crackers' substring functions. (John the Ripper Wordlist Rules Syntax, 2015)

2.8. Proactive - Interference

The concept of proactive interference states that similar past events will interfere in the recall of future similar events (Keppel & Underwood, 1962).While originally applied to short term memory it was eventually applied to long term memory (Postman & Keppel, 1977). In an everyday example of this principle, many have experienced a moment of forgetfulness where they have parked a car in one specific spot or area and one day are forced to park else where one day. Then without thinking a person may go back to the location to which they are accustomed. Applied to passwords, this concept is the psychological framework that explains how a user will have a hard time remembering a password for a short time after changing their password. It provides the framework that states that multiple tests of this information will help the subject overcome these effects (Nunes & Weinstein, 2012).This psychological principal can also go to show that people will continue to want to remain set in their ways. This "Because I've always done it this way" principle helps to explain additional reasons why and how a person would cope by maintaining similar content and/or structure in their passwords as a coping mechanism.

2.9. Psychological Coping Mechanisms

There have been a variety of patterns that have become apparent that a significant portion of people use on a daily basis with their passwords. These common coping mechanisms can be documented by information security specialists to help encourage future password creation to occur without these influences and by password crackers to take advantage of these coping mechanisms to break the security of passwords at increased speed. Florencio and Herley analyzed the password habits of 544,960 individuals over a

three-month period. They found that the average user has 6.5 passwords shared across 3.9 different websites. In addition, each user has an average of twenty-five accounts that require passwords, and on average types eight passwords per day. Florencio and Herley estimate that at least 1.5% of users forgot their password every month (Florencio & Herley, 2007). This study is notable for showing that users regularly cope by using the same password across multiple sites and have problems remembering their passwords.

One study that seems to negate many of the concerns of users' writing down passwords was completed by Shay et al. (2010). This study had 470 participants who were required to change their university's password due to a system wide change in policy. While a much smaller percentage of individuals, in this case 13%, wrote down their password, a much larger percent, 80%, reused a password, and 50% reported modifying an old password to create a new one. Each of these are considered coping mechanisms laid out within this study that participants can do in an attempt to comply with password policy requirements(Shay, et al., 2010).

Memory is a finite resource and it has been shown that people cope in a particular way in order to reduce the load that passwords place on memory (Vu, Bhargav, & Proctor, 2003). The cognitive load placed on memory is a heavy burden as the number of passwords scale up and accounts are used less frequently. The precedent exists that when people are placed in situations that require a high cognitive load they will "adapt to the aids available to them so as to maintain low overall levels of effort expenditure" (Todd & Benbasat, 1994). Thus, coping mechanisms can be framed as an attempt to provide leverage on that burden.

These coping mechanisms include variations of passwords that remain upon password changes. The definition that will be used in this paper is any mechanism that

brings information form one password to another across password changes will be called a coping mechanism. Examples of this can include utilizing the same password across multiple accounts, doubling the password, slightly modifying the password by incrementing the number or special character, or even forgetting the password, requiring a reset, which often takes time and decreases productivity (Komanduri, et al., 2011).

2.10. Perspectives on Password Cracking

Many of the means used to think about and analyze passwords come from cryptography. In this way the password can be viewed as a key is in cryptography (excluding the one time pad). The similarities being that a finite amount of time is required to figure out the key as well as a password and a string of characters are needed in order to gain access to information. Unlike some early cryptographic algorithms, information cannot be gained from attempting to try incorrect passwords as in something like a Vigenère cipher. In both cases, a relatively calculable finite lifetime exists where passwords and cryptographic keys provide security to whatever they are protecting. Unlike keys used in cryptography where best practice suggests using completely random information for the key, passwords normally do not ask this of users. This means that the passwords will not be as uncertain as random data. NIST has documented the lowering of uncertainty this causes in their model for entropy. The model NIST set forward does not go on to model any additional drops in entropy accounted for by knowing a user's previous password or passwords. While this model makes sense it assumes that passwords will be kept secret and not broken in the time given before a scheduled password change. This assumption has been broken rather publicly many times in recent years with account

information being leaked to the web with Google, Facebook, and LinkedIn all publicly confirming user information has been breached and publicly posted (Burnett, 2015);(Kleinman, 2014);(Silveira, 2012). The traditional response in all cases to breaches such as these has been to instruct or force affected users to immediately change their passwords. The sense is that the affected accounts are now secure because the passwords are now different than the ones that were leaked. That is unfortunately not the case as said before passwords created by humans are not randomly created and humans draw from our own experiences in order to create passwords. The same psychological principles discussed earlier go into the creation of passwords in most often the same way to create similar new passwords. So that knowing a previous password means that it will probably be easier to break a new password from the same person.

This is made easier by freely available software that allows anyone to attack passwords (Metasploit, 2015; John the Ripper, 2015; THC-Hydra, 2014; Hashcat, 2015). These software packages allow passwords on a variety of systems to be attacked via brute force, dictionary, and rule based attacks. This allows for rule based attacks to be performed targeting commonly identified coping mechanisms in passwords, incrementing, doubling, etc… In addition, other techniques exist to target coping mechanism by generating passwords similar to input. Techniques such as Markov chain modelling and mangling passwords via probabilistic context-free grammars can also be used to break password given input from "training sets" (Dell'Amico, Michiardi, & Yves, 2010).

CHAPTER 3: METHODOLOGY

This chapter consists of the methodology, variables, study environment, samples, instruments, biases, approvals, data collection, and analysis.

The study started with 1030 participants and dropped down at each iteration to a final of approximately 30 participants per policy for a total of 92 participants remaining throughout the whole study. Much of this loss is attributed to a respondents being paid less for iterations 2 through 7 than for iteration 1. This was determined to not affect the representativeness of the study by comparing the passwords of the respondents who completed all 7 iterations by policy of the study to all respondents who completed each individual iteration by policy.

3.1. Methodology and Variables

This is an experimental research study designed to gather 7 password iterations from 1030 individuals recruited using Amazon Mechanical Turk for a total of 2434 passwords. These 2434 passwords were studied to classify the coping mechanisms used. Then the coping mechanisms were analyzed using the NIST measure of entropy to determine the extent to which that particular coping mechanism affects the true entropy. Finally, the coping mechanisms themselves were analyzed to determine if there are patterns in types of coping mechanisms used or patterns across the population.

3.2. Procedure

This password study tracked a participants' passwords across 7 iterations of participation with a password changing once a week. Participants were recruited through Amazon Mechanical Turk. After each iteration was completed participants were invited back through an internal email service provided by Amazon so that no personally identifiable information was exchanged including email addresses. Amazon Mechanical Turk workers not already participating in the study through iteration 1 were prevented from entering the study in any of the following iterations through the use of the Amazon Mechanical Turk qualifications mechanism. The qualification mechanisms also allowed for more efficient tracking of a participant's status within the study. Each invitation that was sent to participants included a link to the Amazon Mechanical Turk Human Interaction Task (HIT) which would provide a link to the study. Using Amazon Mechanical Turk as a proxy between participants and the researcher developed website allowed for efficient payment of participants.

When participants first entered the website they were asked to register with the website. At this point participants were randomly assigned in to one of the three password policies. The participants were asked to enter their Mechanical Turk ID as a username, select a password and a 'forget your password' question and answer at this point. At each login participants entered the study to do a survey which related to information security. The first iteration the study simply asked for demographic information followed by the next five iterations asking basic understanding of information security. The seventh iteration consisted of questions relating to how passwords were chosen for this study and password practices across accounts. For iterations two through seven, participants were

asked to change their password after logging in to the website through a change your password option. Participants were prevented from filling out the survey until they had performed a password change. The specific questions that participants were asked are located in Appendix A of this thesis. The passwords were analyzed for a variety of security related metrics explained in section 3.3 of this thesis.

3.3. Description of Variables

The variables being studied include: 1) prevalence of coping mechanisms, 2) entropy, the measure of randomness and surprise within information theory. Entropy was measured under two sets of criteria: 1) the NIST model of entropy as laid out in the NIST Special Publication 800-63 and 2) the model put forth by this research designed to take into consideration the sequential dependency of coping over time. From this model, the coping mechanisms themselves were examined for their prevalence within the password policies.

3.3.1 Types of Coping Mechanisms

Coping mechanisms have been categorized into three categories 1) using known information, 2) partial repetition, and 3) easily identifiable patterns. In the using known information category, the coping mechanisms that are expected to be seen are usernames being used as passwords and passwords being reused in entirety. The categories of partial repetition contain coping mechanisms such and simply repeating information across time.

The last category of easily identifiable patterns will contain easily recognizable patterns such as incrementing and using only one character in the entire password,

changing the case of a letter within the password, and common keyboard patterns for passwords such as ‘qwerty’. A breakdown of the entropy assignment is as follows for these coping mechanism is as follows.

For Using Known Information

Iteration 1

P	a	s	s	w	o	r	d
4	2	2	2	2	2	2	2

Iteration 2

P	a	s	s	w	o	r	d
1							

In this example I have shown how entire reuse of a password is assigned an value of 1 bit of entropy. The reasoning behind this will be explained in section 3.1.1.2

For Partial Repetition:

Iteration 1:

P	a	s	s	w	o	r	d	a
4	2	2	2	2	2	2	2	1.5

Iteration 2:

P	a	s	s	w	o	r	d	x
4								2

In this example what is seen is the data that has stayed the same over time is assigned the value of entropy for one character under the NIST Policy and any new information is evaluated as it normally would under the NIST policy. In this way the first iteration’s example password would have 19.5 bits of entropy and the second iteration’s password which is dependent on the first iteration’s information would have 6 bits of entropy.

For Easily Identifiable Patterns

Iteration 1:

P	a	s	s	w	o	r	d	1
4	2	2	2	2	2	2	2	1.5

Iteration 2:

R	a	n	d	o	m	P	W	2
4	2	2	2	2	2	2	2	

Iteration 3:

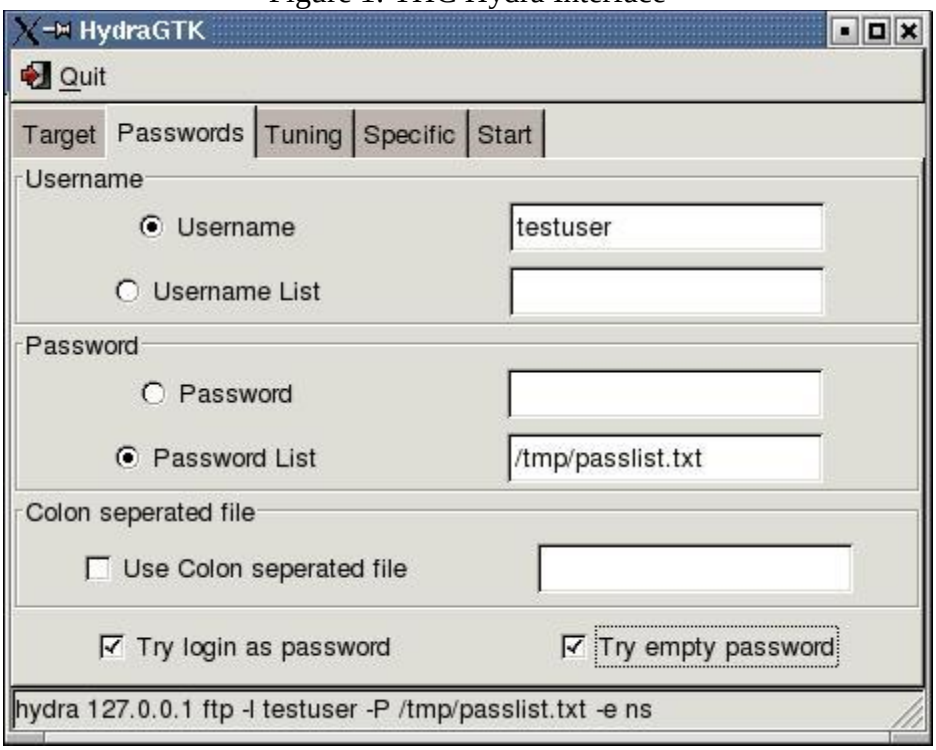
x	x	x	x	x	x	x	x	3
4	2	2	2	2	2	2	2	

In this example what is seen is incrementing over time the numbers being incremented are not assigned any value of entropy in the coping model and any new information is evaluated as it normally would under the NIST policy. Each of these categories will be analyzed for their prevalence amongst the policy.

3.3.2 Entropy Measures

In the using known information coping mechanism of using a username as a password as well as reusing old passwords these passwords will be evaluated as 1 bit of entropy. The reasoning behind this is that most password crackers have the option to use the login information (username) as the first attempt for a brute force check of the password. This is shown in the following figure where the option to “Try login as password” is listed with a checkbox. (THC Hydra Password Cracker, 2015)

Figure 1: THC Hydra Interface



Symantec published a study on zero day exploits indicating that for the average organization that is compromised it takes approximately 10 months before the intrusion is detected (Bilge & Dumitras, 2012). Most organizations have passwords expire after a set length of time usually on the scale of a few months (Scarfone & Souppaya, 2009). This means that most users have already had at least one password change since the organization was compromised or if the said organization has discovered the compromise and has implemented a mandatory password change. In most cases, the check performed on the creation of the new password will only check to make sure the password is different from the current password and conforms to the policy. If the malicious actor had compromised the passwords of the organization then it is customary for the passwords to be placed into a dictionary word list to be used to crack user's passwords in which case the previous passwords have become known quantities and are no longer secret. If the user then reverts

back to previous password it would be trivial to crack the user's account in this case and the reuse of the password should be treated as providing no security with 1 bit of entropy. The rest of the policy for assigning entropy values will rely heavily on the NIST model for entropy laid forth in the special publication 800-63. This way any new data within the password or data that cannot yet be attributed to using known information, partial repetition, or easily identifiable patterns will simply be treated as it would have in the standard NIST implementation of password entropy.

One of the common coping mechanisms within password creation is incrementing. This is done by changing one letter or number usually on the end of a password to the next letter or number in the alphanumeric sequence. The example of this would be Password 1: xxxxxxx1 to Password 2: xxxxxxx2, to Password 3: xxxxxxx3, etc.

In the case of incrementing the user will use a predictable and recognizable pattern in order to lower the load that the password places on his or her memory. Using this case above in conjunction with the NIST Model of Entropy there is one character change between password 1 and password 2, the first seven characters will be treated as a singular block and provide it with what NIST proscribes for the first character in a password 4 bits of entropy. Since the number, the 8^{th} character, in this case is new information it will also be treated as a new character under the NIST model. This means it will be treated as the second character in the NIST policy providing it 2 bits of entropy. Since the next character is still new for password two and it is a number it will still receive the 6 bit special character modifier at this second password. However, if an attacker now knows password 1 and password 2 it is likely that they will have noticed an incrementing from one to two to three this will now be a predictable pattern that can be incredibly easily exploited by password

crackers. This is exploited by the password cracker by treating the alphabetic characters that stay the same in the password as a substring allowing the remaining characters to be randomly brute forced (John the Ripper Wordlist Rules Syntax, 2015) (Rule-based Attack, 2015). So, in the third iteration two separate losses in entropy will be seen. The first is that entropy will no longer be given for the 8^{th} character because the coping mechanism of incrementing has been observed. The second loss of entropy will be from the 6 bit bonus given for using non alphabetic characters, a number. This third iteration the number that has been incremented will be treated as a "numeric variable" and the entropy gained by increasing the key space by adding in numbers to the password is now lost due to the knowledge that only numbers will be used in this character space. This is done because on the password cracking side by specifying that character as a numeric variable (John the Ripper Wordlist Rules Syntax, 2015; Rule-based Attack, 2015). This same loss of entropy for the using only numbers can be seen in other passwords. For example, if a user only uses numbers in their password and it can be predicted that they continue to do so. The entropy bonus for using numbers or special characters will be eliminated from the calculation of the entropy.

3.4. Demographics

As part of the study surveys were taken for respondent's demographics. The demographics that were collected are a respondent's age, gender, race, level of education, income, marital status, and field of employment. This information will not be correlated to password strength as predictors for the entropy strength or for the prevalence of coping mechanics as this is outside the scope of the research questions, but it is made available for

other researchers who do want to make these steps and to show the representativeness of the sample.

3.5. Sample

The original sample of participants used in this study is 1030 participants that were recruited through Amazon Mechanical Turk web services. However, many participants are expected to drop out over the course of seven iterations of the study. As for the purposes of this study as common for most applications if approximately 30 participants remain in the study per policy until iteration 7 then the sample will still remain valid (Morse, 2000). Having approximately 30 participants in each sample groups conforms to much of the literature on sample sizes in studies of memory effects in regards to passwords. (Vu, Bhargav, & Proctor, 2003) (Vu, et al., 2006) (Muligan, Lozito, & Rosner, 2006) (Nunes & Weinstein, 2012)

3.6. Bias

As this study relates to information security due to Institutional Review Board policies as well as Amazon Mechanical Turk policies, it was decided that it would be best to not form the data collection as a deception study. This means that there may be observation bias associated with the fact that the research participants know that they are involved in a research study focused on information security and may act differently in order to impress or satisfy researchers by creating stronger passwords than they would have otherwise. If significant results are observed of coping mechanism eroding the security of

a policy then it can be argued that in a normal environment where users are not being actively observed these coping mechanisms will be even stronger.

3.7. Approval

Institutional Review Board Approval was received for Protocol#1410015359 on November 11th, 2014. With approval for an amendment received on December 8th, 2014. The full application is attached in the section labeled Appendix B.

3.8. Data Collection and Analysis

Data analysis was broken down into four categories; within policy by iteration, within policy across iterations, across policy by iterations, and across policy across iterations. Research questions 1 and 2 were answered by the analysis of the coping mechanisms. The prevalence of coping mechanisms is shown within each policy by iteration and across policies (prevalence of coping mechanisms is determined by dividing the number of passwords that use coping mechanisms by the number of passwords in the policy). The prevalence of coping mechanisms was analyzed individually to show which policy are most susceptible to each coping mechanism. Entropy is given both under the NIST models as well as Post Coping Models along with Standard Deviations (SD). These entropy values were measured for significance using one sample t-testing against the minimum entropy levels allowed by the policy. This was used to determine the answer to research question 3 if the entropy fall below the minimum levels allowable by the policy and when policy the entropy levels fall significantly below the level of minimum entropy by the password policy the password policy is considered broken.

CHAPTER 4: RESULTS AND ANALYSIS

The thesis investigated the following research questions: What is the prevalence and extent of the coping mechanisms (using known information, partial repetition, and easily identifiable patterns) to which users engage by policy and across iterations? Do user coping mechanisms decrease the strength of password policies to such a degree that the increase of security sought by the stringent policy is negated?

4.1. Overview

The average entropy of all the passwords across the study is 32.51129 bits. Splitting the three policies apart and averaging across iterations with all 1030 respondents it is shown that according to the NIST measurements of entropy Basic 16 is the strongest policy followed by Blacklist Hard, and Comprehensive 8.

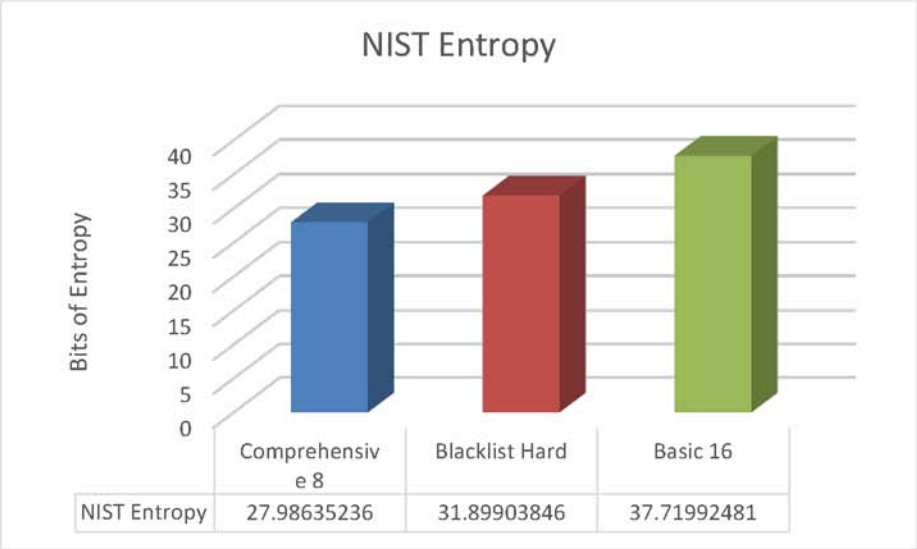


Figure 2: NIST Entropy Average

Using the Post Coping Entropy measure developed in this study the average measure of post coping entropy across the seven iterations involved in the study was 24.61474 bits of entropy. At the overview level of analysis the same strength relationship still exists of Basic 16 being the strongest policy followed by Blacklist Hard and Comprehensive 8 respectively.

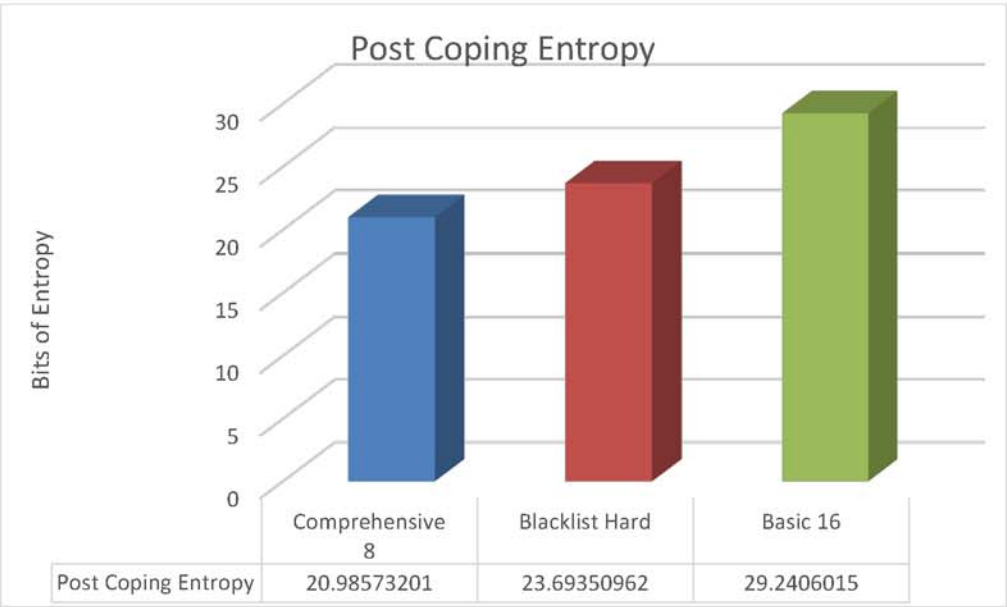


Figure 3: Post Coping Entropy Average

4.2. Policy Analysis

The study went on for 7 iterations first starting with approximately 1030 participants falling to 91 by the end. As shown in the model for NIST Entropy below the passwords created throughout the study show a relatively consistent strength growing over time.

Table 1: Significance in NIST Model by Policy by Iteration

Using NIST 800-63 Model	Comprehensive 8			Blacklist Hard			Basic 16		
	N	Entropy	STD	N	Entropy	STD	N	Entropy	STD
Password Iteration 1	335	27.8	3.5	343	31.3	4.6	350	37.0	4.4
Password Iteration 2	164	27.7	3.3	171	32.0	4.4	161	37.8	5.0
Password Iteration 3	107	28.1	3.6	107	32.2	4.1	98	38.2	5.1
Password Iteration 4	85	28.2	4.1	87	32.1	4.3	78	38.4	5.0
Password Iteration 5	47	28.5	4.0	49	33.2	5.1	44	38.9	5.1
Password Iteration 6	38	28.4	4.0	44	33.0	4.1	34	39.6	6.6
Password Iteration 7	30	29.2	5.0	30	32.8	4.9	32	38.3	4.4

The minimum entropy of the policies are 24, 24, and 30 bits for the Comprehensive 8, Blacklist Hard, and Basic 16 policies respectively. Statistical significance above the minimum entropy levels with $p < 0.05$ is shown in green.

Under the NIST policy all three of the policies examined show that their entropy is significantly above the minimum requirements for the policy. This means that from the perspective of the administrator, the organization they are trying to protect is secured above and beyond that required by the minimum requirements dictated by policy.

Table 2: Average Length of passwords

Lengths	Comprehensive 8	Blacklist Hard	Basic 16
Iteration 1	10.56	9.91	18.32
Iteration 2	10.45	10.04	18.70
Iteration 3	10.72	10.23	19.14
Iteration 4	10.82	10.21	18.91
Iteration 5	11.00	11.47	18.89
Iteration 6	10.92	10.77	19.85
Iteration 7	11.47	11.13	18.66

The thought that respondents went above and beyond the minimum requirements placed upon them by the policy is confirmed once again when looking at the lengths of the passwords that respondents created. The minimum lengths of passwords required by the policies were 8, 8, and 16 characters for the Comprehensive 8, Blacklist Hard, and Basic 16 policies. In every iteration in every policy the average length was at least one whole character above the minimum requirement for length.

This security is called into question once the coping mechanisms are taken into consideration and the passwords are examined over time. Using the methodology laid forth above coping mechanisms were examined and overtime showed to erode the security of the policy to levels significantly below the minimum standards set forth by the policy. Post coping entropy (PCE) measurements for each policy that are significantly above the minimum values for the policy with $p < 0.05$ are listed in green while Post coping entropy (PCE) measurements for each policy that are significantly below the minimum values for entropy with $p < 0.05$ are listed in red.

Table 3: Significance in Post Coping Model by Policy by Iteration

Post Coping Analysis	Comprehensive 8			Blacklist Hard			Basic 16		
	N	PCE	STD	N	PCE	STD	N	PCE	STD
Password Iteration 1	335	27.6	3.7	343	29.1	8.1	350	35.1	6.4
Password Iteration 2	164	18.8	10.7	171	24.5	12.2	161	28.9	13.5
Password Iteration 3	107	14.3	12.0	107	16.8	14.9	98	22.1	16.8
Password Iteration 4	85	15.0	13.7	87	17.1	14.7	78	24.2	16.6
Password Iteration 5	47	14.4	13.3	49	16.6	16.0	44	24.2	16.7
Password Iteration 6	38	15.1	13.3	44	17.9	15.9	34	18.2	18.7
Password Iteration 7	30	17.2	13.8	30	21.7	15.0	32	19.5	18.0

A major difference between the policies is the rate at which entropy is lost. The Comprehensive 8 policy falls below its minimum level of entropy of 24 bits with a value of 18.8 bits at its second iteration. This means that after only one iteration passwords have become public the policy is harmed by the coping mechanisms that users exhibit.

The Blacklist Hard policy falls below its minimum level of entropy of 24 bits with a value of 16.8 bits at its third iteration. The Basic 16 policy falls below its minimum level of entropy of 30 bits with a value of 22.1 bits at its third iteration. The average level of entropy for the Basic 16's post coping entropy was 28.9 at its second iteration which was below the minimum level of 30 bits. However, this did not meet the necessary levels of significance at alpha = 0.05, meaning at this point the post coping entropy was not significantly different that the minimum levels of entropy provided by the policy. In other words the policy was still secure.

This drop in entropy across iterations invalidates the security of the policy with only one disclosure of passwords for Comprehensive 8 and two disclosures of passwords in the Blacklist Hard and Basic 16 policies. This would mean that from the policy perspective, passwords of the Comprehensive 8 policy are not protecting users as well as the Blacklist Hard and Basic 16 policies.

4.3. Sample Representativeness

Table 4: Sample Representativeness 2 sample t-testing

P-values	Comprehensive 8	Blacklist Hard	Basic 16
Iteration 1	0.1042	0.8219	0.0734
Iteration 2	0.9979	0.5480	0.1723
Iteration 3	0.4431	0.3975	0.2874
Iteration 4	0.8293	0.4186	0.2319
Iteration 5	0.8607	0.3444	0.7025
Iteration 6	0.6607	0.3125	0.9068
Iteration 7	1	1	1

It could be argued that using the respondents that have not completed every iteration of the study biases results. In order to assuage these concerns, two sample t-tests were performed against respondents who completed all seven iterations against each and all respondents in the respective iteration by policy. Results show no significant difference at alpha = 0.05 between the respondents who completed all seven iterations of the study and those that did not.

4.4. Comparing Policies

Table 5: Post Coping Analysis by Policy by Iteration

Post Coping Analysis	Comprehensive 8			Blacklist Hard			Basic 16		
	N	PCE	STD	N	PCE	STD	N	PCE	STD
Password Iteration 1	335	27.6	3.7	343	29.1	8.1	350	35.1	6.4
Password Iteration 2	164	18.8	10.7	171	24.5	12.2	161	28.9	13.5
Password Iteration 3	107	14.3	12.0	107	16.8	14.9	98	22.1	16.8
Password Iteration 4	85	15.0	13.7	87	17.1	14.7	78	24.2	16.6
Password Iteration 5	47	14.4	13.3	49	16.6	16.0	44	24.2	16.7
Password Iteration 6	38	15.1	13.3	44	17.9	15.9	34	18.2	18.7
Password Iteration 7	30	17.2	13.8	30	21.7	15.0	32	19.5	18.0

The Basic 16 policy was a stronger policy from the start of the study with 30 bits of entropy and remains stronger than the Blacklist Hard policy or the Comprehensive 8 at iteration 3, after all three policies strengths have fallen below their minimum allowable values. This is measured by showing a statistically significant difference in two sample t-testing with the post coping entropy measures at alpha = 0.05 with p value of 0.0157.

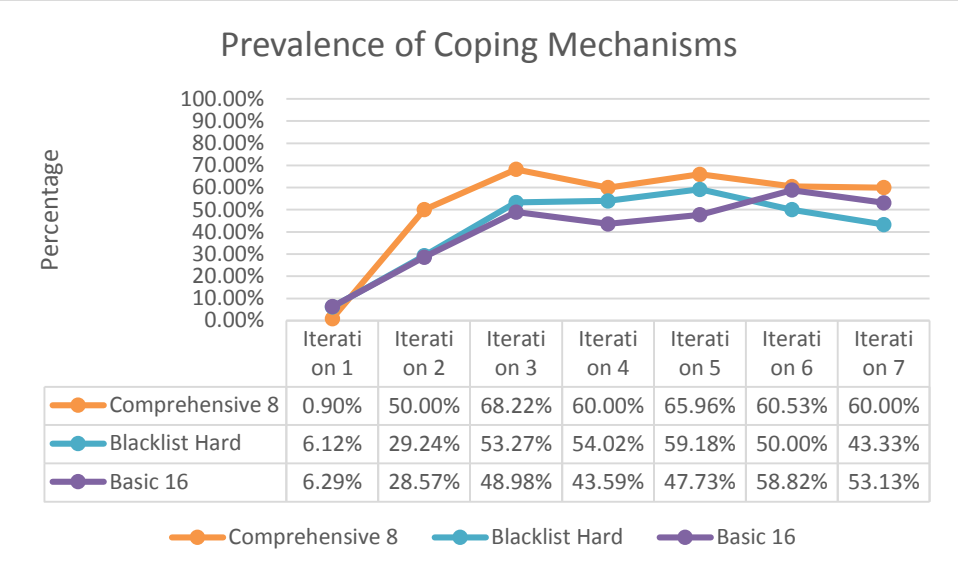
Although the mean of the Blacklist Hard policy is higher than the Comprehensive 8 policy at each iteration in the study by iteration 3 the post coping entropy values are no longer significantly different from each other at alpha = 0.05 with p value of 0.1929. In other words, in terms of protection it cannot be determined which policy (Comprehensive 8 or Blacklist Hard) provides more protection for its users at iteration 3.

4.5. Prevalence and Extent

Table 6: Quantity of Coping Mechanisms by Policy by Iteration

Identifiable Coping Mechanism Prevalence	Number of Participants who Coped In Comprehensive 8	Number of Participants who Coped in Blacklist Hard	Number of Participants who Coped in Basic 16
Iteration 1	3	21	22
Iteration 2	82	50	46
Iteration 3	73	57	48
Iteration 4	51	47	34
Iteration 5	31	29	21
Iteration 6	23	22	20
Iteration 7	18	13	17

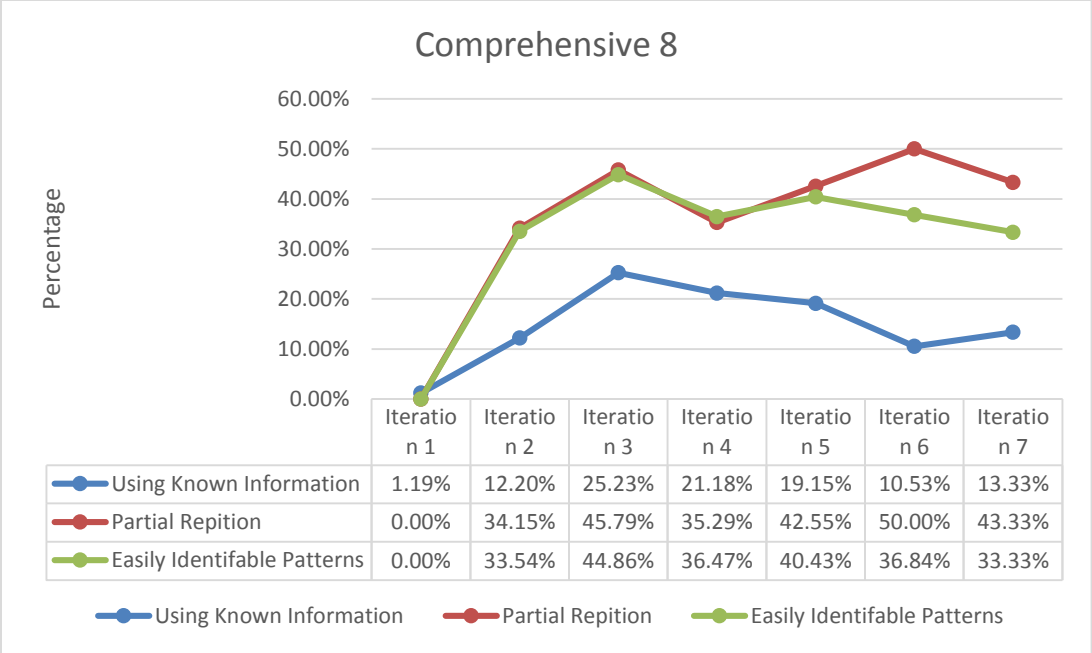
As iterations go on, more and more data can be compared to each of the previous passwords and more coping mechanisms can be identified. Two of the coping mechanisms in this study, partial repetition and easily identifiable patterns can only be identified with successive password iterations. So as passwords are subjected to more and more scrutiny as time goes on in this model as coping mechanisms become easier and easier to identify the more data becomes available to an attacker.



	Iteration 1	Iteration 2	Iteration 3	Iteration 4	Iteration 5	Iteration 6	Iteration 7
Comprehensive 8	0.90%	50.00%	68.22%	60.00%	65.96%	60.53%	60.00%
Blacklist Hard	6.12%	29.24%	53.27%	54.02%	59.18%	50.00%	43.33%
Basic 16	6.29%	28.57%	48.98%	43.59%	47.73%	58.82%	53.13%

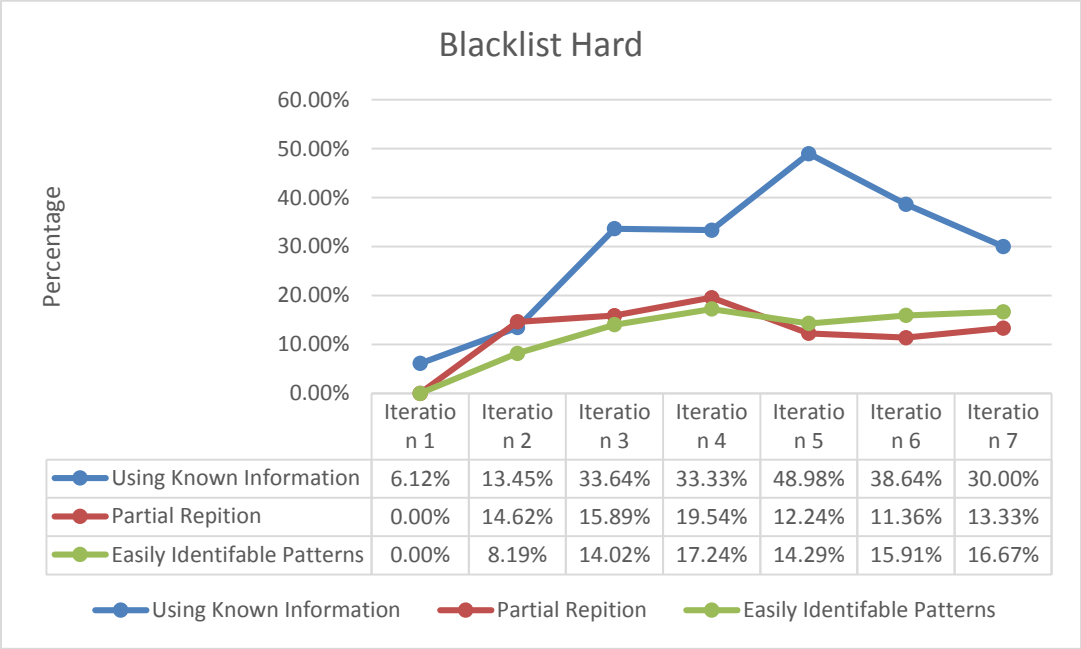
Figure 4: Prevalence of Coping Mechanisms

In terms of quantity, at iteration 1 the prevalence of respondents' coping is higher in the Blacklist Hard and Basic 16 with 21 and 22 respondents coping than in the Comprehensive 8 policy with only 3 respondents coping. The Comprehensive 8 policy shows a much greater prevalence of coping mechanisms in every iteration after iteration 1 than either the Blacklist Hard or the Basic 16 policies.



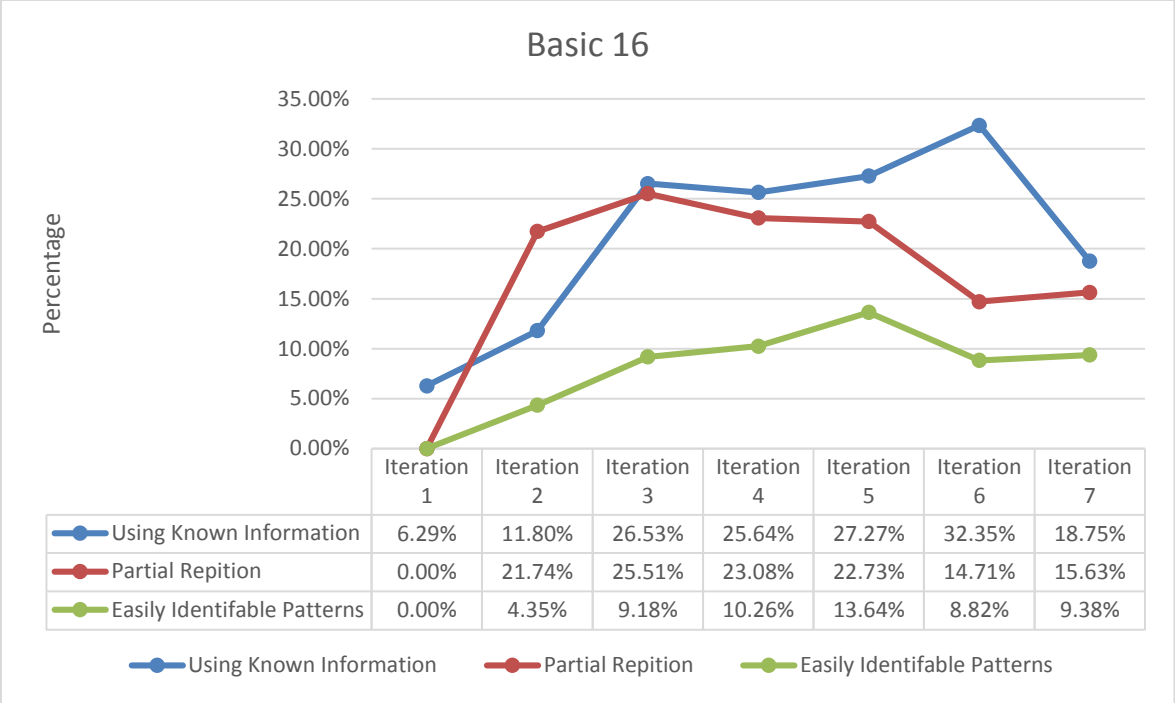
	Iteration 1	Iteration 2	Iteration 3	Iteration 4	Iteration 5	Iteration 6	Iteration 7
Using Known Information	1.19%	12.20%	25.23%	21.18%	19.15%	10.53%	13.33%
Partial Repition	0.00%	34.15%	45.79%	35.29%	42.55%	50.00%	43.33%
Easily Identifable Patterns	0.00%	33.54%	44.86%	36.47%	40.43%	36.84%	33.33%

Figure 5: Extent of Coping Mechanisms in Comprehensive 8



	Iteration 1	Iteration 2	Iteration 3	Iteration 4	Iteration 5	Iteration 6	Iteration 7
Using Known Information	6.12%	13.45%	33.64%	33.33%	48.98%	38.64%	30.00%
Partial Repition	0.00%	14.62%	15.89%	19.54%	12.24%	11.36%	13.33%
Easily Identifable Patterns	0.00%	8.19%	14.02%	17.24%	14.29%	15.91%	16.67%

Figure 6: Extent of Coping Mechanisms in Blacklist Hard



	Iteration 1	Iteration 2	Iteration 3	Iteration 4	Iteration 5	Iteration 6	Iteration 7
Using Known Information	6.29%	11.80%	26.53%	25.64%	27.27%	32.35%	18.75%
Partial Repition	0.00%	21.74%	25.51%	23.08%	22.73%	14.71%	15.63%
Easily Identifable Patterns	0.00%	4.35%	9.18%	10.26%	13.64%	8.82%	9.38%

Figure 7: Extent of Coping Mechanisms in Basic 16

As shown in the above figures the policies examined have different susceptibility to different kinds of coping mechanisms. The Blacklist Hard and Basic 16 policies are shown to be much more susceptible to using known information in the first iteration than the Comprehensive 8 policy. At the first iteration 6.12% of Blacklist Hard respondents used the coping mechanism of using known information, 6.29 % of Basic 16 respondents used known information and 1.19% of respondents for the Comprehensive 8 policy used known information. This trend fades away after the first iteration with the percentage of respondents using known information rising to 12.20%, 13.45%, and 11.80% in the second iteration for the Comprehensive 8, Blacklist Hard, and Basic 16 policies respectively.

Respondents in the Comprehensive 8 policy are much more likely to use partial repetition and easily identifiable patterns when compared to the Blacklist Hard and Basic 16 policies. For the coping mechanism of partial repetition at iteration 2, 34.15% of respondents in the Comprehensive 8 policy compared to 21.74% of users in Basic 16 and 14.62% of respondents in Blacklist Hard. For the coping mechanism of easily identifiable patterns at iteration 2, 33.54% of respondents in the Comprehensive 8 policy compared to 8.19% of users in Blacklist Hard and 4.35% of respondents in Blacklist Hard.

As each successive password became known it became easier and easier to guess the content of the respondent's passwords as users continue to maintain similar information across password changes. When taken into comparison with the previous NIST model it is discovered that entropy is increasingly lost as time goes on instead of gained. It is discovered that this loss in entropy seen at the policy level is eroding the security to the point of ineffectiveness after the first disclosure of passwords in the Comprehensive 8 policy and the second disclosure in the Blacklist Hard and Basic 16 policies. It is due to the qualities of the policies namely the more stringent requirements placed on users in the Comprehensive 8 policy that users cope more heavily. At the same time, the qualities of the Blacklist Hard and Basic 16 policy encourage respondents to cope less except in the category of using known information.

CHAPTER 5: CONCLUSION

5.1. Research Question 1

What is the prevalence of the coping mechanisms to which respondents engage by policy and across iterations?

By iteration two, 50.00% of respondents were coping in the Comprehensive 8, compared to 29.24% of respondents in the Blacklist Hard policy and 28.57% of respondents in the Basic 16 policy. By iteration 3 the prevalence rises to 68.22% of respondents coping in the Comprehensive 8 policy compared to 53.27% of respondents in the Blacklist Hard policy and 48.98% of respondents in the Basic 16 policy. Coping mechanisms are most prevalent in the Comprehensive 8 policy followed by the Blacklist Hard policy and finally the Basic 16 policy. This directly corresponds to the amount of requirements in the policies placed onto the respondents. In the Comprehensive 8 policy there were 4 requirements placed on respondents. (At least: 8 characters, 1 capital letter, 1 lowercase letter, and 1 special character) this wide variety of requirements is compared to the 2 requirements placed on users in the Blacklist Hard policy (At least 8 characters and no words in the dictionary blacklist) and only one requirement in the Basic 16 policy. This indicates that the prevalence of coping mechanisms is positively correlated with the number of requirements placed on the user by the password policy.

If coping mechanisms are treated as a function of the cognitive load that each requirement places on users then it would be best practice from a security perspective to create or choose policies with a limited number of requirements. Based on the results of this study it should be suggested that if a requirement is to be placed on a user through the password policy that the requirement should be minimum length as increasing the minimum length of the password did not serve to increase the prevalence of coping mechanisms. This may be because passwords that are longer can consist of multiple words that will be stored more efficiently in the user's mind as a concept through the process of chunking thus reducing the cognitive load on memory compared to a similar length password consisting of random information. This memory strategy provides more support towards using longer passwords with less policy requirements.

5.2. Research Question 2

To what extent do respondents engage in the following coping mechanisms?

The participants in the Comprehensive 8 policy engaged in coping mechanisms most frequently and most heavily by using by partial repetition and easily identifiable patterns. The features of this policy namely the special characters and numbers serve as the most frequent means of coping. That is the special requirements placed on users by the Comprehensive 8 policy serve as a means only to make passwords more predictable rather than increasing the entropy of the password. The Blacklist Hard and Basic 16 policies show the higher coping only in the category of using known information in a few of the iterations.

These coping mechanisms of using known information can be easily stopped at password changes by securely storing past passwords in an encrypted database and using

them as a blacklist against reuse for future passwords. Another check to ensure that the user's username is not used as the password should also be performed upon password change. Implementing these two steps could stop the coping mechanisms identified as using known information without noticeable increase in wait time from the perspective of the user at the password change. It could be argued that by preventing users from coping in this manner would cause them to move toward other coping methods but they will not be as easy to identify from the attackers perspective and would require further research to identify this movement. Additionally, if coping serves to reduce the cognitive load that passwords place on human memory then by banning some coping mechanisms it could be argued that this would reduce the memorability of passwords. This should be investigated in further research as to the extent of this possibility. However, as discussed before this could be mitigated in longer passwords consisting of multiple words if they exhibit the psychological principle of chunking.

5.3. Research Question 3

Do user coping mechanisms decrease the strength of password policies to such a degree that the increase of security sought by the stringent policy is negated?

Coping mechanisms do serve to erode password security to the point that the password will not serve as an effective security measure in all policies. The difference between policies is at which iteration this happens. After only one iteration of passwords is disclosed the Comprehensive 8 policy can no longer be expected to provide a minimum standard of security of 24 bits of entropy.

This has wide implications for organizations that have experienced password leakages as the common practice in response to this disclosure is an immediate password change. This research indicates that this response will not be enough to provide the expected minimum standard of security after user's passwords have become known once. The Blacklist Hard and Basic 16 policy required 2 disclosures of the passwords before the policy itself fell below its minimum standards of entropy. Meaning from the organizational perspective the policy can still provide adequate protection after one disclosure of passwords and one iteration of changes. However, after two iterations of password disclosure the Blacklist Hard and Basic 16 policies can no longer provide the minimum standard of security. Taking this overview of the policies examined in the study it should be stated that the Basic 16 policy is the most secure policy due to its robustness against coping mechanisms compared to the Blacklist Hard and Comprehensive 8 policies as well as its resistance to loss of entropy through multiple disclosures of passwords.

5.4. Recommendations for Future Work

The coping mechanism using known information consisted of using a username as a password and reusing old passwords in their entirety. These coping mechanisms can be easily prevented by blacklisting user's usernames for their specific accounts and all previous passwords from being reused. This should be the first coping mechanism eliminated due to the relative ease of implementation and low processing power required to maintain security. Blacklisting passwords against reuse is not necessarily a standard implementation for password authentication and not every developer considers security when developing websites. So, developing a standard library that can be used to implement

password authentication could serve to help secure many sites being developed by individuals without extensive training in security.

Other more advanced detection mechanisms should be researched as well such as the Levenshtein distance algorithm to try to prevent users from coping at password changes. Other server side predictive mechanisms can be researched to prevent users from falling into coping patterns by warning users to not use a password similar to their previous passwords. But a balance must be found between processing requirements, wait time on password change, and user memorability.

Further research should be done to develop a model that accurately represents the cognitive load that passwords place on the human memory. This would allow policy creators to more accurately estimate how much they are willing to stress or frustrate their own user base with a certain policy. This will allow policy requirements to be chosen that cater toward memorability and do not encourage coping mechanisms.

In the field of risk analysis, future work can be done to better gauge the long term impact of mass password leakages and incorporate this into the risk framework for organizations. This can help to serve to increase the number of options for system administrators who currently only reset users' passwords in response to mass leakages.

LIST OF REFERENCES

LIST OF REFERENCES

Albanesius, C. (2011). Team Poison Hacks UN, Leaks Usernames, Passwords. *PC Magazine Online*.

Bilge, L., & Dumitras, T. (2012). *Before We Knew It: An Emerical Study of Zero-Day Attacks In The Real World.* Symantec Research Labs.

Burnett, M. (2015, February 9). *Today I Am Releasing Ten Million Passwords*. Retrieved from Xato: https://xato.net/passwords/ten-million-passwords/

Burr, W. E., Dodson, D. F., & Polk, W. T. (2004). *Special Publication 800-63: Electronic Authentication Guideline.* US Department of Commerce, National Institute of Standards and Technology.

Dell'Amico, M., Michiardi, P., & Yves, R. (2010). Password Strength: An Emperical Analysis. *IEEE INFOCOM Proceedings.* IEEE.

Eaton, J. (2011). The Political Significance of the Imperial Watchword in the earl Empire. *Greece and Rome*, 48-63.

EC-Council. (2009). Four Types of Password Attacks. In EC-Council, *Ethical Hacking and Countermeasures: Attack Phases* (pp. 5-1-5-80). Cengage.

Florencio, D., & Herley, C. (2007). A large-scale study of web password habits. *World Wide Web*, (pp. 657-666).

Furnell, S., & Zekri, L. (2006). Replacing passwords: In search of the secret remedy. *Network Security*, 4-8.

Gehringer, E. (2002). Choosing Passwords: Security and Human Factors. IEEE.

Godden, D., & Baddeley, A. (1975). Context Dependent Memory in Two Natural Environments. *British Journal of Psychology*, 325-331.

Hashcat. (2015, January 5). Retrieved from Hashcat: Advanced Password Recovery: http://hashcat.net/hashcat/

John the Ripper. (2015, March 3). Retrieved from Openwall: http://www.openwall.com/john/

John the Ripper Wordlist Rules Syntax. (2015, March 5). Retrieved from Openwall: http://www.openwall.com/john/doc/RULES.shtml

Kelley, P., Komanduri, S., Mazurek, M., Shay, R., Vidas, T., Bauer, L., . . . Lopez, J. (2012). Guess Again (and again and again): Measuring password strength by simulating password-cracking algorithms. *IEEE Symposium on Security and Privacy* (pp. 523-537). IEEE.

Keppel, G., & Underwood, B. J. (1962). Proactive Inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior*, 153-161.

Klein, D. V. (1990). Foiling the cracker: A survey of, and improvements to, password security. *2nd USENIX Security Workshop.*

Kleinman, A. (2014, September 11). *5 Million Gmail Usernames And Associated Passwords Leaked*. Retrieved from Huffington Post: http://www.huffingtonpost.com/2014/09/11/gmail-passwords-hacked_n_5805104.html

Komanduri, S., Shay, R., Kelley, P., Mazurek, M., Bauer, L., Cristin, N., . . . Egelman, S. (2011). *Of Passwords and People, Measuring the Effect of Password Composition Policies.* Vancouver: CHI 2011.

Metasploit. (2015, March 3). Retrieved from Rapid7: http://www.rapid7.com/products/metasploit/download.jsp

Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *Psychological Review*, 81-97.

Morse, J. M. (2000). Determining Sample Size. *Qualitative Health Research*, 3-5.

Muligan, N. W., Lozito, J. P., & Rosner, Z. A. (2006). Generation and Context Memory. *Journal of Experimental Psychology*, 836-846.

Nunes, L., & Weinstein, Y. (2012). Testing Improves True Recall and Protects against the Build-up of Proactive Interference without Increasing False Recall. *Memory*, 138-154.

Pessin, J. (1932). The Effect of Similar and Dissimilar Conditions upon Learning and Relearning. *Journal of Experimental Psychology*, 427-435.

Postman, L., & Keppel, G. (1977). Conditions of Cumulative Proactive Inhibition. *Journal of Experimental Psychology*, 376-403.

Rule-based Attack. (2015, March 5). Retrieved from Hashcat: Advanced Password Recovery: https://hashcat.net/wiki/doku.php?id=rule_based_attack

Sarkar, K. (2004). Two converging worlds: Cyber and physical security. *Federal Computer Week*, 56-57.

Scarfone, K., & Souppaya, M. (2009). *Guide to Enterprise Password Management.* National Institute of Standards and Technology.

Shannon, C. E. (1948). A Mathematical Theory of Communication.

Shay, R., Komanduri, S., Kelley, P., Leon, P., Marzurek, M., Bauer, L., . . . Cranor, L. (2010). Encountering Stronger Password Policies. *Symposium on Usable Privacy and Security (SOUPS).* Redmond.

Silveira, V. (2012, June 6). *An Update on LinkedIn Member Passwords Compromised*. Retrieved from LinkedIn: http://blog.linkedin.com/2012/06/06/linkedin-member-passwords-compromised/

Smith, S., & Vela, E. (2001). Environmental Context-Dependent Memory: A review and meta-analysis. *Psychonomic Bulletin & Review*, 203-220.

THC Hydra Password Cracker. (2015, March). Retrieved from THC-Hydra: https://www.thc.org/thc-hydra/hydra_pass.jpg

THC-Hydra. (2014, December 8). Retrieved from THC-Hydra: https://www.thc.org/thc-hydra/

Todd, A. P., & Benbasat, I. (1994). The Influence of Decision Aids on Choice Strategies Under Conditions of High Cogniive Load. *IEEE Transactions on Systems, Man, and Cybernetics*, 537-547.

Vu, K., Bhargav, A., & Proctor, R. W. (2003). Imposing Password Restrictions for Multiple Accounts: Impacts on Generation and Recall of Passwords. *Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 1331-1335). Santa Monica, CA: Human Factors and Ergonomics Society.

Vu, K.-P. L., Proctor, R. W., Bhargav-Spantzel, A., Tai, B.-L., Cook, J., & Schultz, E. E. (2006). Improving Password Security and Memorability to Protect Personal and Organizational Information. *International Journal of Human-Computer Studies*, 747.

Yiannis, C. (2013). Modern Password Cracking: A hands-on approach to creating an optimised and versatile attack. *Royal Holloway University*.

Zviran, M., & Erlich, Z. (2006). Identification and Authentication: Technology and Implementation Issues. *Communications of the Association for Information Systems*.

APPENDICES

Appendix A. Survey Questions

Demographic Survey (Log In #1)

1. What is your gender?
    - Male
    - Female

2. What is your age?
    - 18 – 24
    - 25 – 34
    - 35 – 44
    - 45 – 54
    - 55 – 64
    - 65+

3. Where were you born?
    - North America
    - Europe
    - Asia
    - South America
    - Australia
    - Africa
    - Antarctica

4. What is your race?
    - Caucasian
    - African-American
    - Asian
    - American Indiana
    - Hawaiian/Pacific Islander
    - Hispanic
    - Other

5. What is your marital status?
    - Single
    - Married
    - Widowed
    - Divorced
    - Separated

6. What is your Ethnicity?

- Hispanic/Latino
- Non-Hispanic/Latino

7. What is the highest level of education you have completed?
    - High School (or equivalent)
    - Technical/trade school completion
    - Some college
    - Undergraduate Degree
    - Some Post Graduate Work
    - Graduate Degree
    - PhD

8. What is your primary occupation?
    - Management
    - Business and Financial Operations
    - Computer and Mathematical
    - Architecture and Engineering
    - Life, Physical, and Social Science
    - Community and Social Service
    - Legal
    - Education, Training, and Library
    - Arts, Design, Entertainment, Sports, and Media
    - Healthcare Practitioner
    - Healthcare Support
    - Protective Service
    - Food Service
    - Building and Grounds, Cleaning and Maintenance
    - Personal Care and Service
    - Sales
    - Office and Administrative Support
    - Farming, Fishing, and Forestry
    - Construction and Extraction
    - Installation, Maintenance, and Repair
    - Transportation and Material Moving

9. What is your income level?
    - $0 - $15,000
    - $15,001 - $25,000
    - $25,001 - $40,000
    - $40,001 - $60,000
    - $60,001 - $80,000
    - $80,001+

Survey Log In #2

1. Were you affected by the Home Depot Breach?
   - Yes
   - No

2. Have you ever received a Data Breach Disclosure Notice?
   - Yes
   - No

3. Do you subscribe to Wired Magazine?
   - Yes
   - No

4. Do you read terms of service policies?
   - Yes
   - No

Survey Log In #3

1. Do you regularly back up your computer system? (Regularly means a minimum of once per month)
   - Yes
   - No

2. If yes, please answer the following. If No, please proceed to number 3.
   - Please indicate what you use for backing up your computer system.
     - Cloud
     - External Hard drive
     - Other

3. Are you able to recognize phishing emails?
   - Yes
   - No

4. I am more concerned about my financial data then my health data?
   - True
   - False

5. Are you familiar with Stuxnet?
   - Yes
   - No

Survey Log In #4

1. What computer operating system do you use?
   - Windows
   - Mac
   - Linux
   - Other

2. Do you use public WiFi to access secure accounts?
   - Yes
   - No

3. Are you concerned about cyber crime?
   - Yes
   - No

4. Were you affected by the Target breach?
   - Yes
   - No

Survey Log In #5

1. Which of the following do you have in your home? (please check all that apply)
   - Smart cellular telephone
   - Laptop computer
   - Desktop computer
   - Tablet
   - Other

2. Are you concerned about identity theft?
   - Yes
   - No

3. Are you able to recognize spam emails?
   - Yes
   - No

4. Have you ever been a victim of a social engineering attack?
   - Yes
   - No
   - Don’t know

Survey Log In #6

1. Do you reach computer acceptable use policies?
   - Yes
   - No

2. Have you ever received a phishing email?
   - Yes
   - No

3. Have you heard of Stop. Think. Connect?
   - Yes
   - No

4. Have you heard of Stop, Drop, and Roll?
   - Yes
   - No

Exit Survey with Password Specific Questions:

1. Did you utilize the same password in this exercise that you used on another account?
   - Yes
   - No

   If yes, please answer the following. If no, please skip to number 2.

   At which point did you start using this similar password?
   - Week 1
   - Week 2
   - Week 3
   - Week 4
   - Week 5
   - Week 6
   - Week 7

2. Did you utilize a similar password in this exercise that you used on another account? Similar means using the same root word, or branch of the same word, with a different prefix or suffix, different placement of capital letters, different numbers, or different special characters.
   - Yes
   - No

   If yes, please answer the following. If no, please skip to number 3.

At which point did you start using this similar password?

- Week 1
- Week 2
- Week 3
- Week 4
- Week 5
- Week 6
- Week 7

3. Did you write down a password during this exercise?
    - Yes
    - No

   If yes, please answer the following. If no, please skip to number 4.

   Only for this exercise, why did you write down your password (please check all that apply):
    - Too difficult to remember due to the policy requirements
    - Too difficult to remember over multiple changes
    - Too difficult to remember due to too many passwords
    - Regularly write down my passwords
    - Other _____

   When did you first write down your password in this exercise? [Purdue students]
    - Week 1
    - Week 3
    - Week 5
    - Week 7
    - Week 9
    - Week 11
    - Week 13

   When did you first write down your password in this exercise? [Mechanical Turk]
    - Week 1
    - Week 2
    - Week 3
    - Week 4
    - Week 5
    - Week 6
    - Week 7

4. Did you utilize any personal information to assist you in creating a password in this exercise?
    - Yes
    - No

If yes, please answer the following. If no, please skip to number 5.

What type of personal information did you use?

- Pet's name
- Street name where you grew up
- Date of birth
- Name of person with whom you are close
- Favorite sports team
- School where you attended
- Organization in which you are a member
- Vehicle you have or wish to have
- Other _____

5. Did you ever become frustrated by the password policy in this study?
    - Yes
    - No

6. Did you ever become frustrated by the frequency of password changes required in this study?
    - Yes
    - No

7. What type of device did you use to log in to this study?
    - Laptop computer
    - Desktop computer
    - Mobile phone
    - Tablet
    - Other _____

8. In your previous experiences with passwords, have you ever been frustrated by a password policy?
    - Yes
    - No

    If yes, please answer the following. If no, please skip to number 9.

    Was the policy that caused you frustration similar to any of the following (please check all that apply)?
    - Minimum of eight characters with a requirement for special characters
    - Minimum of eight characters with a requirement for capital letters
    - Minimum of eight characters with a requirement for numbers
    - Minimum of eight characters with a requirement for numbers, capital letters, and special characters
    - Requirement that you not use any words found in a dictionary
    - Other _____

9. Have you ever been frustrated by the frequency of password changes required by any site?
    - Yes
    - No

    If yes, please answer the following. If no, please skip to number 10.

    How often were you required to change the password?
    - Every 30 days
    - Every 60 days
    - Every 90 days
    - Every 120 days
    - Every 180 days
    - Every 365 days

10. How many accounts do you use that have passwords.
    - 0 – 5
    - 6 – 10
    - 11 – 15
    - 16 – 20
    - 21 – 25
    - 25+

11. Have you ever slightly modified a password to comply with a change in password policy within the same account?
    - Yes
    - No

12. Have you ever written down a password for any reason?
    - Yes
    - No

13. Have you ever used the same password to access multiple accounts?
    - Yes
    - No

Exit Survey with Password Specific Questions:

14. Did you utilize the same password in this exercise that you used on another account?
    - Yes
    - No

    If yes, please answer the following. If no, please skip to number 2.

    At which point did you start using this similar password?
    - Week 1

- Week 2
- Week 3
- Week 4
- Week 5
- Week 6
- Week 7

15. Did you utilize a similar password in this exercise that you used on another account? Similar means using the same root word, or branch of the same word, with a different prefix or suffix, different placement of capital letters, different numbers, or different special characters.
    - Yes
    - No

    If yes, please answer the following. If no, please skip to number 3.

    At which point did you start using this similar password?
    - Week 1
    - Week 2
    - Week 3
    - Week 4
    - Week 5
    - Week 6
    - Week 7

16. Did you write down a password during this exercise?
    - Yes
    - No

    If yes, please answer the following. If no, please skip to number 4.

    Only for this exercise, why did you write down your password (please check all that apply):
    - Too difficult to remember due to the policy requirements
    - Too difficult to remember over multiple changes
    - Too difficult to remember due to too many passwords
    - Regularly write down my passwords
    - Other _____

    When did you first write down your password in this exercise? [Purdue students]
    - Week 1
    - Week 3
    - Week 5
    - Week 7
    - Week 9
    - Week 11
    - Week 13

When did you first write down your password in this exercise? [Mechanical Turk]

- Week 1
- Week 2
- Week 3
- Week 4
- Week 5
- Week 6
- Week 7

17. Did you utilize any personal information to assist you in creating a password in this exercise?
    - Yes
    - No

    If yes, please answer the following. If no, please skip to number 5.

    What type of personal information did you use?

    - Pet's name
    - Street name where you grew up
    - Date of birth
    - Name of person with whom you are close
    - Favorite sports team
    - School where you attended
    - Organization in which you are a member
    - Vehicle you have or wish to have
    - Other _____

18. Did you ever become frustrated by the password policy in this study?
    - Yes
    - No

19. Did you ever become frustrated by the frequency of password changes required in this study?
    - Yes
    - No

20. What type of device did you use to log in to this study?
    - Laptop computer
    - Desktop computer
    - Mobile phone
    - Tablet
    - Other _____

21. In your previous experiences with passwords, have you ever been frustrated by a password policy?
    - Yes
    - No

    If yes, please answer the following. If no, please skip to number 9.

    Was the policy that caused you frustration similar to any of the following (please check all that apply)?
    - Minimum of eight characters with a requirement for special characters
    - Minimum of eight characters with a requirement for capital letters
    - Minimum of eight characters with a requirement for numbers
    - Minimum of eight characters with a requirement for numbers, capital letters, and special characters
    - Requirement that you not use any words found in a dictionary
    - Other _____

22. Have you ever been frustrated by the frequency of password changes required by any site?
    - Yes
    - No

    If yes, please answer the following. If no, please skip to number 10.

    How often were you required to change the password?
    - Every 30 days
    - Every 60 days
    - Every 90 days
    - Every 120 days
    - Every 180 days
    - Every 365 days

23. How many accounts do you use that have passwords?
    - 0 – 5
    - 6 – 10
    - 11 – 15
    - 16 – 20
    - 21 – 25
    - 25+

24. Have you ever slightly modified a password to comply with a change in password policy within the same account?
    - Yes
    - No

25. Have you ever written down a password for any reason?

- Yes
- No

26. Have you ever used the same password to access multiple accounts?
    - Yes
    - No

Appendix B. Institutional Review Board Application

APPLICATION TO USE HUMAN RESEARCH SUBJECTS

Purdue University

Institutional Review Board

---

1. Project Title: Coping Mechanisms in Password Selection

2. Full Review ☐ Expedited Review X

3. Anticipated Funding Source: INSuRE Grant

4. Principal Investigator *[See Policy on Eligibility to serve as a Principal Investigator for Research Involving Human Subjects]*:

Name and Title: Melissa J. Dark, Professor	Technology Department, Knoy, (765) 494-7661, (765) 496-1212, dark@purdue.edu

5. Co-investigators and key personnel *[See Education Policy for Conducting Human Subjects Research]*:

Name and Title: Christopher Foreman, Assistant Professor	College of Technology, Knoy, (765) 494-2558, (765) 494-6219, foremanj@purdue.edu
Brian Curnett, Master's Student	Technology, Recitation, bcurnett@purdue.edu
Teri Flory, Master's Student	Communication, Recitation, tflory@purdue.edu

Com

Tec

6. Consultants *[See Education Policy for Conducting Human Subjects Research]*:

Jeffrey Karpicke, Associate Professor Psychological Sciences, Psychology,

(765) 494-0273,

karpicke@purdue.edu

7. The principal investigator agrees to carry out the proposed project as stated in the application and to promptly report to the Institutional Review Board any proposed changes and/or unanticipated problems involving risks to subjects or others participating in the approved project in accordance with the HRPP Guideline 207 Researcher Responsibilities, Purdue Research Foundation-Purdue University Statement of Principles and the Confidentiality Statement. The principal investigator has received a copy of the Federal-Wide Assurance (FWA) and has access to copies of 45 CFR 46 and the Belmont Report. The principal investigator agrees to inform the Institutional Review Board and complete all necessary reports should the principal investigator terminate University association.

_____

Principal Investigator Signature Date

8. The Department Head (or authorized agent) has read and approved the application. S/he affirms that the use of human subjects in this project is relevant to answer the research question being asked and has scientific or scholarly merit. Additionally s/he agrees to maintain research records in accordance with the IRB's research records retention requirement should the principal investigator terminate association with the University.

_____

Department Head *(printed)* Department Name

_____

Department Head Signature Date

APPLICATION TO USE HUMAN RESEARCH SUBJECTS

9. This project will be conducted at the following location(s): (please indicate city & state)

- [X] Purdue West Lafayette Campus
- [ ] Purdue Regional Campus (Specify):
- [ ] Other (Specify):

10. If this project will involve potentially vulnerable subject populations, please check all that apply.

- [ ] Minors under age 18

☐ Pregnant Women

☐ Fetus/fetal tissue

☐ Prisoners Or Incarcerated Individuals

X University Students (PSYC Dept. subject pool ___)

☐ Elderly Persons

☐ Economically/Educationally Disadvantaged Persons

☐ Mentally/Emotionally/Developmentally Disabled Persons

☐ Minority Groups and/or Non-English Speakers

☐ Intervention(s) that include medical or psychological treatment

11. Indicate the anticipated maximum number of subjects to be enrolled in this protocol as justified by the hypothesis and study procedures: 2,000

12. This project involves the use of an **Investigational New Drug** (IND) or an **Approved Drug For An Unapproved Use**.

☐ YES X NO

Drug name, IND number and company:

13. This project involves the use of an **Investigational Medical Device** or an **Approved Medical Device For An Unapproved Use**.

☐ YES X NO

Device name, IDE number and company:

14. The project involves the use of Radiation or Radioisotopes:

☐ YES X NO

15. Does this project call for: (check-mark all that apply to this study)

☐ Use of Voice, Video, Digital, or Image Recordings?

X Subject Compensation? Please indicate the maximum payment amount to subjects. $2.00

Purdue's Human Subjects Payment Policy Participant Payment Disclosure Form

☐ VO2 Max Exercise?

☐ More Than Minimal Risk?

☐ Waiver of Informed Consent?

☐ Extra Costs To Subjects?

☐ The Use of Blood? Total Amount of Blood _____

Over Time Period (days) _____

☐ The Use of rDNA or Biohazardous materials?

☐ The Use of Human Tissue or Cell Lines?

☐ The Use of Other Fluids that Could Mask the Presence of Blood (Including Urine and Feces)?

☐ The Use of Protected Health Information (Obtained from Healthcare Practitioners or Institutions)?

☐ The Use of academic records?

16. Does investigator or key personnel have a potential financial or other conflict of interest in this study?

☐ YES X NO

APPLICATION NARRATIVE

A. PROPOSED RESEARCH RATIONALE

- Password policies have become so stringent and cumbersome that users may stop creating strong passwords or enlist coping mechanisms, such as writing down their passwords or using them across multiple websites, to facilitate remembering them. These coping methods have the potential to undermine and weaken the effectiveness of these password policies. Password policies, like those that require eight digits with at least one being a capital letter, one being a number, and one being a special character, are designed as authentication methods for access into a specific program, website, or other secure area. We are interested in analyzing users' behavior when creating passwords within these policies. It is our theory that it is possible to identify a policy that eliminates or severely decreases coping methods while retaining significant strength. Additionally, we intend to collect fairly comprehensive demographic and user behavior data as we anticipate future analysis by other investigators.

B. SPECIFIC PROCEDURES TO BE FOLLOWED

- Each participant will be required to log in to a website. At the initial log in, each user will be assigned one password policy to create a password within. In addition, the participant will be required to answer a survey indicating demographic information. During six of the subsequent log ins, the user will be required to change his or her password and answer a short survey of questions on information security. At the time of the last log in, the user will be requested to participate in a survey that will inquire into any coping mechanisms that were utilized throughout the study or frustrations that the user experienced.

- The data collected will be the demographic information of the participants, the specific passwords

that were created, and the answers to the surveys.

C. SUBJECTS TO BE INCLUDED

Describe:

- The subjects will be students at Purdue University. There will be no additional requirements for inclusion.
- There is no exclusion criteria for subjects.
- The maximum number of subjects we seek approval to enroll is 1,000. We anticipate only a Thirty Percent participation rate, so our final subject pool will likely be approximately Three Hundred per collection method.

D. RECRUITMENT OF SUBJECTS AND OBTAINING INFORMED CONSENT

- Certain courses at Purdue will be participating in the project, and each student will be advised that for attendance purposes, he or she will be required to log in to a website after each class to answer one question from lecture that day
- The students will be advised that a password study is also being completed, and will be given the opportunity to participate
- Each student will log in each day he or she is scheduled to attend that course.
- Upon the initial log in, the student will be randomly assigned to a password policy, and will be required to create a password that satisfies that policy.
- If the student is participating in the study, his or her password will be collected and he or she will be requested to participate in an initial survey to collect demographic data.
- Every two weeks, over a fourteen week period, every student will be required to change his or her password for the website (for a total of seven passwords to be created)
- The passwords of the students who are participating will be collected and analyzed
- At the end of the data collection, the students will be invited to participate in a survey regarding any coping mechanisms used throughout the study

E. PROCEDURES FOR PAYMENT OF SUBJECTS

- There will be no financial compensation for participation in the study.
- Purdue University Professors will be able to award extra credit to participating students, so long as the non-participating students are also awarded a non-related extra credit opportunity for the same amount of points that takes a similar amount of effort as participation in the study.

F. CONFIDENTIALITY

- The investigators do not anticipate having or keeping any names, telephone numbers, or email addresses of the participants. Each user will be given a random identification number. The investigators will not have access to the participants' names, only this assigned number. The initial list of names that connects to the user identification number for Purdue students will be held by Dr. Melissa Dark and Dr. Christopher Foreman in a secure location.
- Attached to this application is the Data Management Plan.

G. POTENTIAL RISKS TO SUBJECTS

- The potential risk is minimal, in that no personal identifying information will be utilized by the

investigators.

- It is likely that the participants will use a password throughout the process that might be used in

another aspect of their lives. The participants will be advised of this before data is utilized, and it will be

suggested that if the password used was not unique, that he or she may wish to change the other passwords

that are similar. The participants will also be granted the opportunity to not have their passwords analyzed

once the data collection has been complete and the full study explained.

H. BENEFITS TO BE GAINED BY THE INDIVIDUAL AND/OR SOCIETY

- Each of the subjects will have the opportunity to reflect on his or her own password selection, and how it might be advantageous to utilize stronger passwords throughout their lives.
- Society is very dependent upon authentication measures such as passwords, but we

believe our

study will indicate that passwords are a very weak type of such measures. Society as a whole will hopefully

observe the results and begin seriously looking at alternative or supplemental authentication measures.

I. INVESTIGATOR'S EVALUATION OF THE RISK-BENEFIT RATIO

- The risk involved is no greater than that involved in every day life.

J. WRITTEN INFORMED CONSENT FORM

- The Informed Consent Forms are attached to this Application.

K. WAIVER OF INFORMED CONSENT OR SIGNED CONSENT

If requesting either a waiver of consent or a waiver of signed consent, please address the following:

1. We are not requesting a waiver of consent.

L. INTERNATIONAL RESEARCH

Our research will not be international

M. SUPPORTING DOCUMENTS

- Attached are copies of the surveys to be completed at the beginning and end of the study

Appendix C. Statistical Analysis System (SAS) Code

```
options ls=72;

goptions colors=(none);


data NISTcomprehsive8iteration1;

input NISTEntropy1;

datalines;

;

data NISTcomprehsive8iteration2;

input NISTEntropy2;

datalines;

;

data NISTcomprehsive8iteration3;

input NISTEntropy3;

datalines;

;

data NISTcomprehsive8iteration4;

input NISTEntropy4;

datalines;

;

data NISTcomprehsive8iteration5;

input NISTEntropy5;

datalines;
```

```
;
data NISTcomprehsive8iteration6;
input NISTEntropy6;
datalines;
;
data NISTcomprehsive8iteration7;
input NISTEntropy7;
datalines;
;

data NISTblacklistHarditeration1;
input NISTEntropy1;
datalines;
;
data NISTblacklistHarditeration2;
input NISTEntropy2;
datalines;
;
data NISTblacklistHarditeration3;
input NISTEntropy3;
datalines;
;
data NISTblacklistHarditeration4;
```

```
input NISTEntropy4;
datalines;
;
data NISTblacklistHarditeration5;
input NISTEntropy5;
datalines;
;
data NISTblacklistHarditeration6;
input NISTEntropy6;
datalines;
;
data NISTblacklistHarditeration7;
input NISTEntropy7;
datalines;
;

data NISTbasic16iteration1;
input NISTEntropy1;
datalines;
;
data NISTbasic16iteration2;
input NISTEntropy2;
datalines;
```

```
;
data NISTbasic16iteration3;
input NISTEntropy3;
datalines;
;
data NISTbasic16iteration4;
input NISTEntropy4;
datalines;
;
data NISTbasic16iteration5;
input NISTEntropy5;
datalines;
;
data NISTbasic16iteration6;
input NISTEntropy6;
datalines;
;
data NISTbasic16iteration7;
input NISTEntropy7;
datalines;
;

data comprehensive8iteration1;
```

```
input PostCopingEntropyAndNIST1;
datalines;
;
data comprehensive8iteration2;
input PostCopingEntropyAndNIST2;
datalines;
;
data comprehensive8iteration3;
input PostCopingEntropyAndNIST3;
datalines;
;
data comprehensive8iteration4;
input PostCopingEntropyAndNIST4;
datalines;
;
data comprehensive8iteration5;
input PostCopingEntropyAndNIST5;
datalines;
;
data comprehensive8iteration6;
input PostCopingEntropyAndNIST6;
datalines;
;
```

```
data comprehensive8iteration7;
input PostCopingEntropyAndNIST7;
datalines;
;

data BlacklistHarditeration1;
input PostCopingEntropyAndNIST1;
datalines;
;
data BlacklistHarditeration2;
input PostCopingEntropyAndNIST2;
datalines;
;
data BlacklistHarditeration3;
input PostCopingEntropyAndNIST3;
datalines;
;
data BlacklistHarditeration4;
input PostCopingEntropyAndNIST4;
datalines;
;
data BlacklistHarditeration5;
input PostCopingEntropyAndNIST5;
```

```
datalines;
;
data BlacklistHarditeration6;
input PostCopingEntropyAndNIST6;
datalines;
;
data BlacklistHarditeration7;
input PostCopingEntropyAndNIST7;
datalines;
;

data basic16iteration1;
input PostCopingEntropyAndNIST1;
datalines;
;
data basic16iteration2;
input PostCopingEntropyAndNIST2;
datalines;
;
data basic16iteration3;
input PostCopingEntropyAndNIST3;
datalines;
;
```

```
data basic16iteration4;
input PostCopingEntropyAndNIST4;
datalines;
;
data basic16iteration5;
input PostCopingEntropyAndNIST5;
datalines;
;
data basic16iteration6;
input PostCopingEntropyAndNIST6;
datalines;
;
data basic16iteration7;
input PostCopingEntropyAndNIST7;
datalines;
;



proc ttest H0=24 alpha=0.05 data=NISTcomprehsive8iteration1;
   var NISTEntropy1;
run;
proc ttest H0=24 alpha=0.05 data=NISTcomprehsive8iteration2;
```

```
  var NISTEntropy2;
run;
proc ttest H0=24 alpha=0.05 data=NISTcomprehsive8iteration3;
  var NISTEntropy3;
run;
proc ttest H0=24 alpha=0.05 data=NISTcomprehsive8iteration4;
  var NISTEntropy4;
run;
proc ttest H0=24 alpha=0.05 data=NISTcomprehsive8iteration5;
  var NISTEntropy5;
run;
proc ttest H0=24 alpha=0.05 data=NISTcomprehsive8iteration6;
  var NISTEntropy6;
run;
proc ttest H0=24 alpha=0.05 data=NISTcomprehsive8iteration7;
  var NISTEntropy7;
run;

proc ttest H0=24 alpha=0.05 data=NISTblacklistHarditeration1;
  var NISTEntropy1;
run;
proc ttest H0=24 alpha=0.05 data=NISTblacklistHarditeration2;
  var NISTEntropy2;
```

```
run;
proc ttest H0=24 alpha=0.05 data=NISTblacklistHarditeration3;
  var NISTEntropy3;
run;
proc ttest H0=24 alpha=0.05 data=NISTblacklistHarditeration4;
  var NISTEntropy4;
run;
proc ttest H0=24 alpha=0.05 data=NISTblacklistHarditeration5;
  var NISTEntropy5;
run;
proc ttest H0=24 alpha=0.05 data=NISTblacklistHarditeration6;
  var NISTEntropy6;
run;
proc ttest H0=24 alpha=0.05 data=NISTblacklistHarditeration7;
  var NISTEntropy7;
run;

proc ttest H0=30 alpha=0.05 data=NISTbasic16iteration1;
  var NISTEntropy1;
run;
proc ttest H0=30 alpha=0.05 data=NISTbasic16iteration2;
  var NISTEntropy2;
run;
```

```
proc ttest H0=30 alpha=0.05 data=NISTbasic16iteration3;
  var NISTEntropy3;
run;
proc ttest H0=30 alpha=0.05 data=NISTbasic16iteration4;
  var NISTEntropy4;
run;
proc ttest H0=30 alpha=0.05 data=NISTbasic16iteration5;
  var NISTEntropy5;
run;
proc ttest H0=30 alpha=0.05 data=NISTbasic16iteration6;
  var NISTEntropy6;
run;
proc ttest H0=30 alpha=0.05 data=NISTbasic16iteration7;
  var NISTEntropy7;
run;


proc ttest H0=24 alpha=0.05 data=comprehensive8iteration1;
  var PostCopingEntropyAndNIST1;
run;
proc ttest H0=24 alpha=0.05 data=comprehensive8iteration2;
  var PostCopingEntropyAndNIST2;
run;
proc ttest H0=24 alpha=0.05 data=comprehensive8iteration3;
```

```
  var PostCopingEntropyAndNIST3;
run;
proc ttest H0=24 alpha=0.05 data=comprehensive8iteration4;
  var PostCopingEntropyAndNIST4;
run;
proc ttest H0=24 alpha=0.05 data=comprehensive8iteration5;
  var PostCopingEntropyAndNIST5;
run;
proc ttest H0=24 alpha=0.05 data=comprehensive8iteration6;
  var PostCopingEntropyAndNIST6;
run;
proc ttest H0=24 alpha=0.05 data=comprehensive8iteration7;
  var PostCopingEntropyAndNIST7;
run;

proc ttest H0=24 alpha=0.05 data=BlacklistHarditeration1;
  var PostCopingEntropyAndNIST1;
run;
proc ttest H0=24 alpha=0.05 data=BlacklistHarditeration2;
  var PostCopingEntropyAndNIST2;
run;
proc ttest H0=24 alpha=0.05 data=BlacklistHarditeration3;
  var PostCopingEntropyAndNIST3;
```

```
run;
proc ttest H0=24 alpha=0.05 data=BlacklistHarditeration4;
  var PostCopingEntropyAndNIST4;
run;
proc ttest H0=24 alpha=0.05 data=BlacklistHarditeration5;
  var PostCopingEntropyAndNIST5;
run;
proc ttest H0=24 alpha=0.05 data=BlacklistHarditeration6;
  var PostCopingEntropyAndNIST6;
run;
proc ttest H0=24 alpha=0.05 data=BlacklistHarditeration7;
  var PostCopingEntropyAndNIST7;
run;

proc ttest H0=30 alpha=0.05 data=basic16iteration1;
  var PostCopingEntropyAndNIST1;
run;
proc ttest H0=30 alpha=0.05 data=basic16iteration2;
  var PostCopingEntropyAndNIST2;
run;
proc ttest H0=30 alpha=0.05 data=basic16iteration3;
  var PostCopingEntropyAndNIST3;
run;
```

```
proc ttest H0=30 alpha=0.05 data=basic16iteration4;
   var PostCopingEntropyAndNIST4;
run;
proc ttest H0=30 alpha=0.05 data=basic16iteration5;
   var PostCopingEntropyAndNIST5;
run;
proc ttest H0=30 alpha=0.05 data=basic16iteration6;
   var PostCopingEntropyAndNIST6;
run;
proc ttest H0=30 alpha=0.05 data=basic16iteration7;
   var PostCopingEntropyAndNIST7;
run;



data comprehensiv8blacklisthard;
   input group entropy;
   datalines;
;
run;
data blacklisthardBasic16;
   input group entropy;
   datalines;
;
```

```
run;
/* tests if the two groups are equal, we want the line that says “Diff” for the Confidence
Interval and the line that says “Satterthwaite” for the test */
proc ttest H0=0 alpha=0.05 data=comprehensiv8blacklisthard;
   class group;
   var entropy;
run;
proc ttest H0=0 alpha=0.05 data=blacklisthardBasic16;
   class group;
   var entropy;
run;
```