5-2021

Security fatigue and its effects on perceived password strength among university students

Chase Carroll
University of Tennessee at Chattanooga, csy385@mocs.utc.edu

Security Fatigue and its Effects on Perceived Password Strength Among University Students

Chase Carroll

Departmental Honors Thesis
The University of Tennessee at Chattanooga
Computer Science: Cyber Security

Examination Date: 10/27/20

David Schwab
Professor of Computer Science and Engineering
Thesis Director

Dr. Mengjun Xie
Professor of Computer Science and Engineering
Department Examiner

Table of Contents

Abstract 1

I – Introduction 2

- Background of the Problem 3
- Problem Statement 5
- Purpose of the Study 5
- Population and Sample 6
- Significance of the Study 6
- Nature of the Study 7
- Research Questions/Hypotheses 8
- Assumptions, Limitations, and Delimitations 8
- Chapter Summary 9

II – Literature Review 12

- Title Searches and Documentation 12
- Historical Content 13
- Current Content 16
- Methodological Literature 19
- Research Design Literature 19
- Conclusions 21
- Chapter Summary 22

III – Research Methodology 24

- Research Method and Design Appropriateness 24
- Research Questions/Hypotheses 25
- Population and Sample 26
- Informed Consent and Confidentiality 27
- Instrumentation 27
- Data Analysis 28
  - Qualitative Analysis 28
  - Quantitative Analysis 30
  - Combined Analysis 31
- Summary 31

IV – Results 34

- Research Questions/Hypotheses 34

Results 35
Fatigue Results 35
Tags 37
Password Strength Results 40
Summary 42
V – Conclusions 45
Research Questions/Hypotheses 45
Discussion of Findings 46
Hypothesis 1 46
Hypothesis 2 47
What Could Be Improved 48
Recommendations to Leaders and Practitioners 49
Recommendations for Future Research 51
Summary 52
References 55
Appendix A: Informed Consent 57
Appendix B: Survey Questions 59
Section 1: Fatigue-targeting Questions 59
Section 2: Password Strength Perception Questions 59
Appendix C: List of Tables 60
Appendix D: List of Figures 60
Appendix E: IRB Approval Number 60

Abstract

This study was performed with the goal of observing the effect, if any, that security fatigue has on students' perceived strength of passwords. In doing so, it was hoped to find some correlation between the two that would help in establishing a measurable effect of the phenomenon in students. This could potentially aid organizational decision-makers, such as security policy writers and system admins, to make more informed decisions about implementing security measures. To achieve the goal of observing this fatigue and attempting to measure it, a survey was distributed to numerous students on the University of Tennessee at Chattanooga campus. The results of the final data analysis show no correlation between security fatigue and students' perceived password strength, but other findings of note did emerge. Notably, both fatigue-state groups of students showed very close mean scores for perceived password strength, with those scores indicating a higher trust in the strength of passwords than is actual. This result implies a lack of influence from security fatigue, as well as a general deficiency in students' abilities to properly judge the strength of passwords. 2-Factor Authentication is thus proposed as a primary item of interest for addressing this deficiency and meeting the needs of students with varying priorities according to their fatigue state.

I – Introduction

In today's technologically advanced society there is seldom a time to let your guard down. With computing devices being ever-present in the form of cell phones, smartphones, computers, tablets, and more, humanity is more connected than ever. However, any computer expert, or even general user, can tell you that such a vast amount of interconnectivity brings its own host of problems. For instance, Verizon's Data Breach Investigations Report for 2020 [1] showed that two cyber threats implicated in data breaches, those being phishing and the use of stolen credentials, appeared only slightly less between 2019 and 2020 while still taking the top spots for threats among those breaches; they are affectionately referred to by the authors as "our old foes." Threats like these may be well known and expected now, but their continued prevalence in real cyber-attacks indicates they are still a serious risk. When users are told to engage in a protective mindset and behavior against these threats, day in and day out, even when the threats seem distant or unlikely to affect them personally, that is when users' willingness to participate in those protective practices dwindles.

This effect has already been documented by others [2] and coined as "Security Fatigue." It is a kind of mental fatigue that pertains to the user's unwillingness to adhere to security-minded actions and behaviors, though according to Steven Furnell and Kerry-Lynn Thomson, it is more of a gradual depreciation of security compliance than outright rebellion against it. Those actions and behaviors include things such as keeping a computer system updated, being cautious of malicious emails, and changing one's password on time, to name a few. The last item, passwords, in conjunction with security fatigue, will be the focus of this paper. Specifically, this work is to find what effect- if any- security fatigue has on the perceived strength of passwords.

The study will be performed on a collection of survey responses from a university student population. The broad outline of this chapter is described thusly:

(1) The Background of the Problem, which will explain the origins and context of the problem on which this study is based. (2) The Problem Statement, which will identify the problem. (3) The Purpose of the Study, which will elaborate on the research objectives and means of achieving them. (4) The Population of the Study, which will speak about the group from which data for this thesis was collected. (5) The Significance of the Study, which will outline the ways in which this thesis is unique and contributes meaningful information to the field of cybersecurity, or wherever else it is pertinent. (6) The Nature of the Study, which will provide an overview of the methodology and justification. (7) The Research Question/Hypotheses, which is fairly self-explanatory. (8) The Assumptions, Limitations, and Delimitations, which will state any prior assumptions about the outcome/data that may have been present, any limitations to the design and/or execution of the various processes of the study, and what limits on the scope and generalizability of the project may exist. Finally, a (9) Chapter Summary will summarize key points and takeaways from the introduction.

Background of the Problem

When it comes to research studies, there are two types that are commonly used: quantitative and qualitative. The former is mainly useful for answering the “how many” and “how much” questions and has many advantages for the researcher [3]. To name just a few, data for these studies can be collected quickly, and it is both independent of the researcher and numerical, making it easily validated. For instance, identifying the number of soda cans sold across many locations can indicate where hotspots of soda drinkers are present, thus enabling the soda company to direct supply where the demand is highest. That is of course a very trivial form

of quantitative research, but it exemplifies how a qualitative method is useful. On the flipside, qualitative research provides different kinds of insights that aren't necessarily so universal, but in many ways more useful within a defined context.

Say that the same soda company wanted to improve one of their product lines after reporting a decrease in sales. A quantitative study of why people don't like the drink would not be very useful. Given that quantitative studies are based on numbers and measurable things that should be able to exist in a generalized, contextless sense- at least after analysis, then it stands to reason that applying the methods of a quantitative study to a question that is inherently based on individual context is wrong. If the study designers only intended to ask about certain measurable aspects of the product- sugar content, carbonation level, etc.- or wanted to know about all measurable aspects of it, then they could perhaps attempt a quantitative study outright, but in the former situation they would be biased in choosing aspects and may miss the root cause of the issue, and the latter situation could drive away respondents with the great length and complexity of the data collection tool- a survey, most likely.

Alternatively, they could choose to solicit open-ended responses from the customers with just one or a few questions, such as "How could we improve (insert drink here)?" While responses from such a question are not strictly measurable and require more subjective interpretations, the smaller number of questions would likely encourage higher response rates because of the shorter time investment and better respondent focus on the issues most pressing to them, their "emic" or insider viewpoint [3]. The onus is then on the researcher to properly analyze the responses for keywords, phrases, and the like, which can be turned into a quantitative study assuming appropriate interrater reliability [4] when coding or due diligence in ensuring credibility of a single coder and validity of their work [5].

All of this is to say that quantitative and qualitative studies each have their own purposes. The first provides insights through statistical analysis that are more uniformly interpretable, acceptable, and actionable, at the cost of sometimes missing important context that could help explain the human side of a problem. The second provides rich details and important context that can be lost in a general study, at the cost of being more subjective and less verifiable. Where this applies to security fatigue is in the lack of quantitative information.

Problem Statement

The problem that drives this study is a lack of empirical data on the effects of security fatigue. Studies involving security fatigue thus far have used it in a qualitative manner- for instance, identifying whether it seems to exist- without trying to make any quantitative connection between the fatigue and its alleged effects. This is somewhat reasonable given that drawing conclusions between qualitative and quantitative data can dilute the value of the latter with the subjectivity of the former, however, considering the increasingly digitized nature of everyday life in the modern world- especially following the Covid-19 pandemic, understanding more about the factors that impact cyber security efforts is paramount, especially for one that is tied closely to the implementation of those efforts.

Purpose of the Study

For this study, both qualitative and quantitative data will be collected through an online survey and analyzed in conjunction with each other to determine: Whether security fatigue may be present in a respondent, and whether it has a consistent effect on the perceived strength of passwords. Additionally, recommendations on how to interpret and use the findings of this study will be provided. This study is being performed in the Southeast United States in the city of

Chattanooga, Tennessee, locally based out of and on the University of Tennessee at Chattanooga (UTC) campus.

Population and Sample

The population from which the sample used in this study was pulled is the student body of the UTC. The sampling type used in this study is homogenous convenience sampling due to the identity of all participants as students of the university, though varied disciplines are represented across the sample. This population was selected because of the many avenues for distribution that the author possessed. After the survey had concluded and non-viable responses were removed from the data set, 135 responses were recorded.

Significance of the Study

This study takes a unique approach to security fatigue by attempting to go further than a qualitative analysis. By isolating responses that show signs of fatigue against those who do not and then examining password scores from both groups against each other, noteworthy differences may be found. The implications of this research are such that they could, in a limited sociodemographic capacity, identify a measurable correlation between security fatigue and students' perceptions of password strength. This could provide previously unexplored insights and implications about the costs of security fatigue in a quantifiable manner, assisting policy writers, systems administrators, and universities in making the right decisions to offset any expected issues from the implementation of fatigue-causing security measures. Additionally, this study may help future research with a focus on, or relation to, security fatigue by opening new avenues of inquiry, such as an experiment to see which security measures have the greatest impact on users- for better or worse.

Nature of the Study

This will be a mixed method study using both quantitative and qualitative data to answer the presented questions and/or hypotheses. Given that- as previously elaborated on in the Background of the Problem section- qualitative studies are not ideal for presenting generalizations because of their low verifiability, and quantitative studies are not ideal for providing valuable context and human depth to data, an effort to reconcile them has been made by collecting both kinds of data alongside each other. This was deemed necessary in order to address the problem of the thesis, and according to prior research on mixed methods studies, "What is most fundamental is the research question – research methods should *follow* research questions in a way that offers the best chance to obtain useful answers. Many research questions… are best and most fully answered through mixed research solutions." [3]

Instrumentation for this study includes a two-section survey presenting both qualitative written response questions and Likert-scale quantitative questions. The former qualitative questions address the respondent about certain password practices and concepts, while the latter quantitative questions collect data on the respondent's perception of password strength. The questions used in the written portion were adapted from a study performed by researchers at NIST, who were able to unintentionally elicit responses containing signs of security fatigue during interviews about online activity and cybersecurity [6]. Passwords were chosen as the litmus test for the quantitative portion because of their ubiquity in society, ensuring every respondent would be able to have an opinion on them. Passwords are partly chosen from a password breach list by a Python script, partly randomly generated using an online password generator, and partly chosen by hand because of unique qualities about them- such as character

replacement. Data analysis is performed in two parts, first through qualitative coding by hand, then a statistical analysis of the quantitative data using Microsoft Excel.

Research Questions/Hypotheses

This section details the research questions and hypothesis of this study.

R1: Is security fatigue present in the sample?

R1.1: What is the proportion of fatigued respondents to non-fatigued respondents?

R1.2: Does the presence of security fatigue have a consistent, observable effect on the perceived strength of passwords?

H1: It is hypothesized that security fatigue will be dominantly represented in the sample, such that the proportion of fatigued vs. non-fatigued respondents will be larger by some arbitrary amount.

H2: It is hypothesized that security fatigue will have a negative impact on the perceived strength of passwords, i.e., those displaying signs of security fatigue will show a higher trust in weaker passwords overall than those who do not show fatigue.

Assumptions, Limitations, and Delimitations

It is assumed for this study that the selected population from which a sample was taken is fairly tech savvy and is knowledgeable of common cyber security concepts like threats (viruses, phishing) and security practices (password resets, lockouts, etc.). It is also assumed that the population lies between 17 and 25 years old, that being the rounded range for undergraduate students at UTC- the dominant category of students- according to 2019 demographic information [7].

Limitations of this study include the number of analysts available to look over the data, which was only the author of this thesis. This has implications for the trustworthiness and validity of the data collection and analysis, and as such it would have been preferable to have multiple people available to verify findings for such things as qualitative coding and tagging, but this was a constraint that had to be accepted given manpower availability. Additionally, being a mixed method study has implications for the generalizability of the findings. Given that quantitative data can often be statistically generalized to a population, while qualitative data is often only analytically generalized to a theory, the lesser scoped of the two (analytic) is likely to be the extent of this study's generalizability, owing to its restricted scope and the type of data collection used [8].

The research was deliberately delimitated to the geographic and demographic region of the UTC campus because of resource constraints on the author. Additionally, the variety of data collected by the survey was limited due to shared concerns between the author and thesis director, specifically regarding the sensitive topic of passwords and respondent anonymity. Open-ended questions were designed as carefully as possible in order to limit the chance of a personally identifying information (PII) leak in the responses, such as from a respondent interpreting a question about password practices to mean "Give them one of my passwords as an example." This had the side-effect of leading to a very narrowly focused data set, and any demographics or social information about respondents- barring discipline- would have to be inferred from official UTC reports.

Chapter Summary

Thus far, the concept of "security fatigue" has been introduced and defined as an unwillingness to adhere to security practices and behaviors, manifesting as a gradual

depreciation of compliance rather than outright rebellion [2]. There have been previous studies on the topic of security fatigue [2][6], but none have attempted to quantitatively analyze the phenomenon and its effects, which would be a useful extension of the literature for those in the position to be making security decisions or performing further research on that fatigue. Purely qualitative and quantitative studies have their advantages and disadvantages [3], but by combining them into a mixed methods study, greater depth and usefulness of the data can be generated. This could help bridge the gap between the concept of security fatigue and actionable, measurable data, again, given caveats about context.

To accomplish the goal of performing a mixed methods study, password strength ranking was chosen as the vehicle to drive quantitative analysis, while open-ended questions adapted for a focus on passwords were obtained from the literature [6] to perform qualitative analysis. Passwords were obtained from a variety of sources, including pseudo-random selection from a password breach list via a Python script, machine generation via a website, and a few hand-picked passwords chosen for their unique characteristics- namely weak character replacement (P@ssw0rd, etc.).

Ultimately, this study is meant to answer the proposed research questions and approve or reject the hypotheses. It is desired to know whether the sample has shown any responses with signs of fatigue, and then if so, what proportion of the sample, and does security fatigue seem to have a consistent and measurable effect on those showing signs of fatigue? The expected effect is, of course, that fatigued respondents will consistently score weak passwords as stronger than they really are, and they will do so in higher numbers than the non-fatigued group does. It is also expected that fatigue will heavily tinge the data, being the dominant category in frequency.

Whatever the outcome, the results will need to be interpreted from a geographic and sociodemographic context of undergraduate public university students in the Southeast United States, located specifically in Chattanooga, Tennessee at UTC [7]. Additional modifiers to the consideration of this research include the singular researcher performing data collection, analysis, and interpretation, the type of data collection performed (convenience), the questionable generalizability of the study beyond theory [8], and the constraints placed upon the instrumentation due to privacy concerns regarding password confidentiality and respondent anonymity.

In Chapter 2, a thorough review of the literature involved in the construction and guiding of this thesis will be performed.

II – Literature Review

Previously, it has been discussed that the focus of this thesis is on the examination of security fatigue's effects from a quantitative standpoint. It is understood that security fatigue is a gradual disillusionment from secure behavior and practices, though it is not related to an outright refusal of those two things from the very beginning. It was also discussed that perceived password strength was chosen as the mechanism by which quantitative effect would be correlated to the fatigue. This was decided on the basis that passwords are ubiquitous and most people understand topics related to them. The following sections will cover the literature picked to represent parts of this study, including the central premises of security fatigue, passwords, methodology, and research design.

Title Searches and Documentation

A specific-to-general approach was taken for searching out sources to inform the thesis-outside sources obtained from prior sources, and a variety of databases and search engines were utilized, including but not limited to: ScienceDirect, ProQuest Central, IEEE Xplore, UTC's library, Google, and Google Scholar. Altogether, there are several categories that prior research used in this thesis could be lumped into:

- Security Fatigue & Fatigue-related
    - The central premise that the study is based on, using a definition like that found in Steven Furnell's work with NIST [2]. Additionally, any other relevant sources acquired through searches of this category may have been nominally related, such as articles on other forms of fatigue.
    - Keyword examples (in order of specific-to-general)

        - "Security Fatigue," Security Fatigue, password fatigue, information technology fatigue, security burnout, information technology burnout

- Password Strength & Passwords
    - The quantitative component of the data collection instrument. This category includes works based generally around passwords, such as articles specifically on passwords or articles focused on password strength, making it a broad group.
    - Keyword examples (in order of specific-to-general)
        - Password strength, password complexity, password strength checker, passwords
- Methodological & Research Design
    - Any articles to do with methodology or research design, such as data analysis, survey design, etc. Some overlap with other categories is present here, considering that the methodology of this study was adapted from prior work in the area.
    - Keyword examples (in order of specific-to-general)
        - Survey design, qualitative research, quantitative research

Historical Content

As far as strictly historical content is concerned there is not much related to security fatigue as it is a relatively new concept. Having emerged from a study performed by researchers at NIST, "security fatigue" is defined by Steven Furnell and Kerry-Lynn Thomson as a situation wherein users "have actually been following good practice and then drift (or completely switch) into a mode in which they become tired or disillusioned with it." [2] This central idea drives their

discussion on some possible causes of the fatigue, a hypothetical method to measure its severity, and then finally how to identify and potentially treat it.

Given the exploratory nature of the study and the topic it addresses, many of the ideas presented throughout it are difficult to approach through any lens but a conceptual one. Security fatigue itself is inherently a malleable condition, varying from one person to another in severity and ease of onset, as is noted by the authors. Additionally, their proposed equation for measuring its potential in a person is based on subjective variables like Effort, Difficulty, and Importance. Of course, none of it is meant to be taken as actionable theory, but it does a good job of establishing the core of what security fatigue is, why it exists, how to think about it, and the means by which it could be addressed.

The groundwork for security fatigue is set up as a people-problem, i.e., one that cannot easily be measurably represented. Indeed, the authors confirm that the necessity for measurement of the fatigue's severity is questionable, given the difficulty of doing so and the suggested greater benefit of preventative care over waiting to judge the severity after an incident. It is for that reason that this thesis focuses more on the effects of the fatigue rather than the fatigue itself. Current work in the field has more to say on the fatigue itself, but that is a discussion for later.

Though it falls outside the current bounds of this research to measure fatigue's varying levels of severity, which is information that would likely impact the findings of this thesis if it could be expanded, it has been suggested by previous works from Furnell and Thomson that such levels could exist. Interpreting their tiers of "user acceptance of IT" as levels of security fatigue, one can surmise that there is a positive and negative end of the spectrum that users can fall under, a security vigor and security fatigue scale if you will [9]. In their work, however, the scale was based on user acceptance levels, which could be used to determine when preventative action

is needed to ensure security compliant behavior. The parallels are obvious, though with some caveats.

There are eight levels suggested by the authors for interpreting user acceptance of IT, with the first four representing positive mindsets and the last four negative mindsets: Culture, Commitment, Obedience, Awareness, Ignorance, Apathy, Resistance, and Disobedience. From first to last, they represent a gradual decrease in security-minded behavior. When viewed from the angle of fatigue, the model for user acceptance works fairly well, though it does not account for the non-fatigued starting condition set by the authors in their later work- user was resistant from the start- or gray areas between Awareness and Ignorance as noted by the authors in the earlier article [9]. However, despite that shortcoming, it is a great way to visualize security fatigue in its more natural shades of gray form as opposed to true or false, given that each person will experience fatigue at a different level.

Considering the knowledge about levels of security fatigue based on user acceptance, the fatigue's effect on user behavior seems to escalate quickly, but the measurements of when a user reaches those levels is still not known. Considering modern knowledge from the authors is that measuring the severity level of the fatigue itself is not necessary. They serve best as a conceptual model to base reasoning off of for further qualitative study. However, if one is to take the proposed effects of the Disobedience and Ignorance levels as true possibilities for users on the scale, then there would be a vast difference in the effect on security effort efficacy around those users. Since one is not expected to solve the problem of quantifying the fatigue's severity itself, then quantifying the effects of the fatigue is the best solution. With a large enough sample of fatigued and non-fatigued users, one could expect to find a correlation to tie to the worst level of fatigue and use as a baseline for future decision-making. Essentially, assume the worst to prepare

for the worst. The baseline found in that manner would give weight to the threat of security fatigue instead of leaving it as a vague threat.

So thus far, it has been cleared up that it would be wise to focus on quantifying the effects of the fatigue rather than quantifying the severity of it. Passwords and their application in this regard will be covered in the next chapter, Methodology.

Current Content

For current literature, there is a bit more to pull from, though the topic is still in its infancy and developing at this time. Arriving on the scene a few years after Furnell, Brian Stanton [6] from NIST proposed that security fatigue was a subset of decision fatigue. Using a previously collected dataset and coding it for fatigued responses, the researchers discovered a high number of their responses contained signs of security fatigue. Through previous work on mental models and the application of heuristics to security efforts, they were able to link security fatigue to decision fatigue as part of a whole.

This study is particularly interesting for its placement of security fatigue in the secondary focus to the more overarching theme of decision fatigue. By framing security fatigue as a situation wherein users are forced to make more decisions than they have the capacity to make, which is intrinsically what security fatigue boils down to when one considers the constantly evolving state of tech and policy, they start a discussion on not only the negative impact of security apparatus, but on the limitations and failings of the human mind that enable the negative impact of that security.

This research provides useful clarification about the driving forces behind security fatigue, which itself is truly a more domain-specific variant on decision fatigue, as noted by

Stanton. With this study, it becomes impossible to solely rest the blame for fatigue on the security mechanisms in play. The users themselves may have heuristic processes that limit their ability to make security decisions that would negatively impact their pursuit of a primary goal [10], or they may be resistant because of competence, as noted by Belanger [11], etc. This information is necessary for acquiring a full picture of the human element in the equation, potentially enriching the findings of this thesis.

On the more technical side of things, a conference held in 2017 featured proceedings from Shigeaki Tanimoto et al. for a concept modeling technique of security fatigue severity [12]. Following close behind the findings from 2016's NIST study of security fatigue, they focused on a way to visualize the fatigue, proposing that "Over the long and mid-term, it will contribute to optimizing a security policy." The study's team produced a model of the "vicious cycle" that ultimately leads to security incidents, then using that in conjunction with considerations about burnout syndrome literature, developed a matrix model for security fatigue.

The developments of this study are quite interesting and potentially impactful to the field. Firstly, they managed to create a sensible model for the perceived severity level of fatigue with respect to security countermeasures, which had not been explored before. The model consisted of not only the perceived intensity of the fatigue, but the perceived security observance level of the user. The latter factor, observance, understandably has a positive or negative correlation with the knowledge and/or awareness of the user to security items, such as policy, practice, etc. The former factor, fatigue, has an impact on their willingness to adhere to security. Both axes can be influenced via the injection of security countermeasures, such as vacation time to reduce fatigue or training to increase observance level. For the purposes of visualization it is a very efficient scale to measure with, though it naturally has some limitations, such as the model being untested,

being limited to qualitative measurements, and the fact that this work was one of the first in this area.

Of course, as mentioned, the focus of the model was on qualitative ranking. No efforts were made by the authors to apply a quantitative measure to fatigue severity beyond the categorical numbering of the matrix cells. This is not an unreasonable course of action considering prior discussion on the matter of security fatigue severity, though it seems to be an intractable issue currently. Because this thesis is focused on measuring the impact of a generalized security fatigue rather than trying to measure its effects at different intensities, the model is not useful to this study. However, knowledge gleaned from their 2017 work, as well as their 2018 continuation, in combination with whatever findings this paper may generate, could provide opportunity for overlap in future research.

Takashi Hatashima and Shigeaki Tanimoto et al. return to continue the work performed in 2017 by evaluating the effectiveness of risk assessment and security fatigue visualization for internal e-crime [13]. *Internal* e-crime was chosen to be their litmus test because of its nearly equal impact compared to *external* e-crime in the most damaging cases. After examining criminal literature to extract 33 risk factors and risk countermeasures, they were able to determine that 15 of those factors could be countered through use of their security fatigue model to identify and mitigate risk factor causes.

The continuation of their prior work on the model helped to prove its usefulness in practical scenarios, such as policy planning, by showing the ways in which it can mitigate future risk. However, much like their previous work, the authors admit that it is qualitative and indirect. However, they have proven it can be applied to other areas, which lends credibility to the model. The same consideration given to their previous work applies to this one. If some useful

quantitative findings can be made about the impact of security fatigue, perhaps future research could do a more granular exploration of it, such as in the various stages of the visualization model.

Methodological Literature

Four studies have previously been covered in relation to security fatigue [2], [6], [12], [13]. Of those four, only the two most recent ones have gone beyond the identification phase of security fatigue to try and then categorize it based on severity. The first two studies in this collection were concerned chiefly with defining, contextualizing, and examining security fatigue on its own merits [2], and contextualizing and examining security fatigue on the merit of being a subset of decision fatigue [6], respectively. So thus, the earlier two studies were concerned with developing the concept of security fatigue and the latter two were concerned with estimating its intensity based on categories. All these studies have been qualitative in nature.

The most recent two studies [12-13], both based on the visualization model of security fatigue, have attempted to show its danger qualitatively in order to guide behavior away from fatigue. Essentially, they desired to raise awareness about its presence with varied levels of threat. However, no attempts have been made to quantify the effects of security fatigue, let alone at varied levels of intensity. Lacking such information leaves too much room for interpretation and may lead to over or under-estimation of its impact on an organization, and that is why this thesis will attempt to fill the gap with a quantitative analysis of security fatigue's effects.

Research Design Literature

The primary inspiration for the data collection and processing portion of this research was the study conducted by Brian Stanton et al. [6]. Within their study, they analyzed data from

a previous study that interviewed 40 average users about their knowledge, behaviors, and emotions related to online activity and cybersecurity. Some categories of the questions used for that study were given, including questions related to online activities, computer security, security icons and tools, and security terminology. With questions like those, the responses acquired from the interviews were able to be recoded for security fatigue and produce extensive results, as noted when they said "When compiled together, there were more than eight single-spaced pages of data related to security fatigue." The question categories provided by that study were adapted to become the open-response portion of the survey, which was what would provide data for the determination of security fatigue in a respondent.

In order to show how respondents were thinking from a quantitative standpoint, a simple metric had to be chosen next, something that wouldn't need much explanation to the average user and that could be evaluated objectively from the author's side: Passwords. Literature on evaluating password strength checking software was used to determine validity of the primary tool, zxcvbn, and additional checkers were employed for redundancy and elimination of bias from one tool alone [14]. Zxcvbn 4.4.2 was chosen as the primary tool for analyzing passwords because of its open-source nature, ease of use, and documented sophistication compared to many other strength checkers. The decision to base strength estimates off a score given by zxcvbn, as opposed to a time-to-crack from a program like John the Ripper, was made primarily because of convenience and the apparent reliability of zxcvbn.

Finally, literature on qualitative and mixed-methods research was used to help guide thinking for the execution of the study and interpretation of results, given its use of both qualitative and quantitative data to draw a single conclusion [3-5], [8].

Conclusions

After review of the literature associated with security fatigue and this study, several things are clear about both. Firstly, with respect to security fatigue, it is not easy or advisable to try to quantify the severity of the fatigue [2]. Placing a measurable number to a categorical concept is not sensible or preferable to simply identifying it, however, it is plausible that the effects of the fatigue can be quantified. This is preferable, in fact, because prior research [12-13] has already done a well enough job defining categories wherein security fatigue may exist with respect to a secondary qualitative factor (security observance), and this helps organizations visualize its threat in order to take preventative actions against it. Those actions, however, may be too little or too much, running the risk of throwing fatigue-observance balance off even further. Though quantifying the effects of security fatigue at various stages of the fatigue-observance matrix is beyond the scope of this thesis, an analysis of the general effect- if any- of the fatigue will be further useful in contextualizing its threat. Future research can expand on this by combining both the effects and matrix model to look for more granular differences.

The nature of this study is also something to consider, as it uses both qualitative and quantitative measurements to draw a single conclusion. Open-response questions for the survey were adapted from literature out of NIST [6], which based their research on data obtained from another 40 person study before them, which focused on security and computer knowledge and opinions of average users. Because this study obtained such heavily fatigue-tinged data, similar questions were used for the survey of this thesis, albeit modified to pertain to passwords. On the topic of passwords, which constitute the quantitative portion of this work, they were chosen for their ubiquity and objectivity for the author. Very few if any people would have trouble

answering non-esoteric password questions, and passwords have many means of being evaluated objectively, such as strength meters/checkers [14].

Chapter Summary

In this chapter, a selection of sources was examined for their place in this work. Security fatigue literature such as [2], [6], [12], and [13] formed the backbone of the concept being tested, along with applicable insights from former work by Furnell [9]. Furnell and Stanton provided necessary fundamental looks at security fatigue, first establishing it and then expanding it into the broader area of decision fatigue, which clarified that the fatigue has as much to do with internal human mechanisms as it does with the external security mechanisms. Tanimoto and Hatashima provided first the groundwork for a model of visualizing its severity, and then a confirmation of that model's practicality in reducing risk. Furnell and Stanton constructed the frame of the problem with a stable foundation, and Tanimoto and Hatashima built out half of the structure by expanding upon the qualitative portion of security fatigue, which was the fatigue's intensity, or how much of it is present. That left the quantitative portion of the fatigue, its measurable effect, that needs to be expanded upon in order to fully- or mostly, pending future work- define its threat.

For this study, Stanton provided inspiration for the questions that would help determine security fatigue, based upon an older study of average users and their views on security and technology topics. Passwords were chosen independently, however, for they are ubiquitous and easy to interpret, both for the respondent and the author. Appropriate tools were chosen to guide the quantitative portion of the thesis, both during survey development, data collection, and analysis [14]. It was also determined that the nature of this research as mixed-methods will benefit its goal by providing both context (fatigued vs non-fatigued) and proof (quantitative

scoring) in the same instance. Where needed, efforts were taken to limit the disadvantages of such research- generalizability concerns, consistency issues, etc.- by using techniques and insights from past works [3-5], [8].

In the next chapter, the Methodology of this study will be reviewed.

III – Research Methodology

The methodology set forth in this chapter and its constituent parts are all used in the pursuit of quantifying security fatigue's effects. The following chapter will extensively cover the methodology of this study, which has been broken up over the following sections: (1) Research Method, (2) Research Questions/Hypotheses, (3) Population and Sample, (4) Informed Consent and Confidentiality, (5) Instrumentation, (7) Data Analysis, and (8) Summary.

Research Method and Design Appropriateness

As stated in Chapter 1, this thesis takes a mixed-method approach to the problems and hypotheses being discussed and tested. In contrast to single-method approaches that focus on purely qualitative or quantitative processes, mixed-method studies attempt to use the advantages of both approaches to enhance each other, while also assuming some of the risk associated with both designs. Quantitative studies have great potential in supporting change to existing theory or proving points, owing to their use of statistically provable results. Consequently, quantitative studies may miss important context behind data because they are chiefly concerned only with the numerical results. Qualitative studies are, naturally, the opposite. They offer more open-ended approaches to problem solving and analysis than a rigid mathematical one would, which allows them to produce more contextually relevant results. Said results can offer insight into conceptual topics and phenomena that science cannot yet fully explain with math, but of course, that means results from a qualitative study will likely be subjective to some degree, making their acceptance in less agreeable circles more difficult.

Given the topic of security fatigue, one would immediately associate it with a quality of a person, how fatigued they are about security. Like any other kind of fatigue, it will vary per

person and from situation to situation, so getting a measure on the intensity of its presence is not so easy. It is an issue that can only be directly observed qualitatively, such as by grouping people into categories of their perceived level of fatigue. Even still, however, stating that one person is "at risk" while another is "deeply fatigued" does not tell much about it. Even by associating predicted behavior patterns with those categories, such as putting off password resets until the day of expiration, the actual risk associated with each category is open to interpretation. Some may see it as normal behavior that everyone engages in, while others may think it is a critical weakness in the organization. This subjectivity inhibits a sense of urgency to rise around the fatigue.

However, though the severity of the fatigue may only realistically be qualified, the effects it leads to could be quantified. For instance- and this is the stratagem of the mixed-method approach here, the measure of how fatigued and non-fatigued people score passwords could be analyzed for a correlation. Firstly though, you would have to show that some of the people scoring were fatigued, and that is where the qualitative portion comes into play. Then it becomes possible to not only determine that fatigue is present, but to also give it more urgency with quantifiable values.

For the goals of this thesis, which are answering the questions and testing the hypotheses in the following section, this mixed-method approach is best. A qualitative analysis alone could not solve the issue of urgency, and a quantitative analysis will not work on its own. For context, the fatigue state is necessary with this topic.

Research Questions/Hypotheses

This section reiterates the previously stated research questions and hypotheses of Chapter 1.

R1: Is security fatigue present in the sample?

R1.1: What is the proportion of fatigued respondents to non-fatigued respondents?

R1.2: Does the presence of security fatigue have a consistent, observable effect on the perceived strength of passwords?

H1: It is hypothesized that security fatigue will be dominantly represented in the sample, such that the proportion of fatigued vs. non-fatigued respondents will be larger by some arbitrary amount.

H2: It is hypothesized that security fatigue will have a negative impact on the perceived strength of passwords, i.e., those displaying signs of security fatigue will show a higher trust in weaker passwords overall than those who do not show fatigue.

Population and Sample

The population for this study is the whole of the student body at the University of Tennessee at Chattanooga. This is including Undergraduates and Graduates, as well as every major department that could be included given the responses. The desired sample size is 377, which is the minimum number of responses needed to confidently make assertions about the whole population. Additionally, the sample would ideally be evenly representative of major departments so that there is not overrepresentation of certain groups. With a combined Undergraduate and Graduate population of 11651 according to 2019 data [15], the sample size for a statistically significant result at 95% confidence is 377, but the real size of the collected sample after excluding nonviable responses was 135. This gives a margin of error of approximately 8% versus the desired 5%.

Informed Consent and Confidentiality

Consent of respondents was acquired via their agreement to a consent form on the survey before any questions were presented to them. If they declined to consent to the collection of their responses and any other data associated with the survey, then they were disallowed from participating in the study. For those who did consent to participate, a minimal amount of questions that could elicit Personally Identifiable Information (PII) were presented, and the settings of the survey were set to not automatically record any respondent information such as IP address, name, etc. Emphasis was made for respondents to not include their names, and all questions in the survey were carefully designed to reduce the chance of an accidental PII inclusion. After all responses were recorded, the dataset was downloaded and cleaned of any PII as it was analyzed. The dataset was stored on the author's password-protected computer running Windows Defender and Malwarebytes scans on a regular basis. No cloud storage was used, nor were any cloud-based applications used in the analysis of the dataset. After conclusion of analysis and final results, the dataset will be erased both off the author's device and SurveyMonkey, the platform that hosted the survey.

See Appendix A for the informed consent form.

Instrumentation

Instrumentation for this thesis includes a survey collecting both open-response and Likert-scale answers from respondents. The survey was administered as a single whole, but for the purposes of discussion about it and its results, the survey is described as having two sections- the aforementioned two kinds of questions. Section 1, which contains the open-ended questions, was intended to elicit information used in the deduction of fatigue. Questions in Section 1 were inspired by the categories of questions presented in [2]. Section 2, which contains the Likert-

scale questions, was intended to measure the user's perceived strength of some given passwords. The table below shows where the sections of the survey work to answer the research questions and/or hypotheses.

RQ/H	Section 1 Questions	Section 2 Questions
R1	X	
R1.1	X	
R1.2	X	X
H1	X	
H2	X	X

Table 1: Alignment of Survey Sections to Research Questions/Hypotheses

See Appendix B for the questions of the survey.

Data Analysis

The analysis for this thesis was performed over three sets of data formed by splitting the collected information from the survey: Fatigue-related (qualitative) data, password-related (quantitative) data, and combined (both) data.

Qualitative Analysis

The fatigue-related data originated from Section 1 of the survey and elicited responses that could be analyzed for signs of security fatigue. Several categories were made for grouping responses, as well as a tags system for extracting additional information from responses that could be used to look for possible trends.

Categories and Criteria

Three overall categories were provided for the responses: Potentially Fatigued, Potentially Non-fatigued, and Inconclusive.

The former two are preceded by "Potentially" because of the difficult task of empirically proving fatigue. It may, in fact, be impossible to empirically prove fatigue given the kind of data

collected here, and so the categories are taken as estimates of a respondent's state. The choice between "Potentially Fatigued" vs. "Potentially Non-fatigued" falls to a consideration of the interpreted tone of the response, words used in the response, and nature of the question asked. For instance: The nature of the question may elicit an opinion about mandatory password changes, and the respondent may use definably negative language such as "**hate** it" or "it is a **hassle**," which may then be a potential case for labeling that question's response as "Potentially Fatigued."

The overall tone of the response may differ from small segments, however, such as if the respondent states "it is a hassle, but I understand and follow the rules," which may then be a case for labeling the question response as "Potentially Non-fatigued" since they indicated they still adhere to best practices. It could be stated that the individual still shows some signs of fatigue and may be heading toward a state of it that affects their actions as well as their beliefs, but grouping based upon severity is beyond the scope of this thesis and has already been done before [12-13]. Tags help make up for this shortcoming.

The Inconclusive category exists for catching non-relevant responses that may be off-topic or uselessly vague (single word answers), as well as instances where the count of Potentially Fatigued and Potentially Non-Fatigued answers were equal, which is taken as meaning the respondent could lean either way.

All coding for the fatigue-related questions was performed by hand. Continued quality and objectivity of coding over the dataset was ensured through reiteration of previously coded responses, but the limitations of a single coder must be considered when interpreting the final results. A team of several experts to assist in coding would have been ideal for providing group confirmation of findings, but this was an unavoidable limitation of the work. Additionally, the

analysis for this section's data was performed over the course of several weeks, often in batches of 10-20 responses at a time, which could have potentially led to minor alterations in coding performance due to changes in the author's mood and other factors.

Tags

Tags are a response to the limited number of categories used for coding. When analyzing the answers, certain keywords or phrases become shared between different respondents. Themes like "hard time remembering" and "hate" can be indicators for the tags "Difficulty Remembering" and "Frustration" respectively, which themselves have definitions that extend the value of the responses beyond the categories. Using tags, the categories are kept simple and uncluttered, and data can be examined for tags to check for trends. For instance, if by the end there was a dominant percentage of responses that had the "Difficulty Remembering" tag, then that may be an indicator of a predominant source of fatigue. Though that example is unhelpful for showing correlation between security fatigue and perceived password strength, it could allow for extension of the discussion into other areas at the end, such as what issues were faced most by the fatigued group vs. the non-fatigued group.

Tags were assigned while the questions were being coded. Performing these two duties concurrently not only saved time, but it also improved both processes. By paying close attention to the words and phrases being used in the responses, more tags could occasionally be derived from the text, which helped to make coding more systematic with the addition of new documented patterns to watch for.

Quantitative Analysis

Responses from Section 2 of the survey contained scorings of ten passwords provided to the respondents. Scores given by each respondent could range between 1 and 5 using the zxcvbn

rating scale [14]. With ten questions, this gives each question a maximum .10 value (if rated as 5/5, or Very Strong) and a minimum 0 value (if rated 0/5, or Very Weak). The Section 2 total for each respondent was summed and then compared against the sum of the machine-calculated strength values using zxcvbn, which was .44. Any score above this threshold would indicate a higher trust in the strength of the passwords than is actual, and any score below it would indicate a lack of trust.

To give a short example, the sum score of a respondent may be .54. This value is the sum of the ten password scores given by the respondent, as each password may be scored anywhere between 0/5 to 5/5, or 0 to .10, considering that there are ten passwords in total. Compared against the true sum score of the passwords according to zxcvbn, which is .44, the respondent shows higher trust in the strength of the passwords than is actually true.

Combined Analysis

After both sections were analyzed, the respondent's Section 1 determination could then be compared with their Section 2 result to achieve any one of the following states:

1. **Fatigued** with **high trust**
2. **Fatigued** with **low trust**
3. **Non-Fatigued** with **high trust**
4. **Non-Fatigued** with **low trust**
5. **Inconclusive** (considering Section 1 only)

Summary

In this chapter, it was further discussed that the chosen research design of unifying qualitative and quantitative data was appropriate. This is because singularly qualitative or

quantitative designs would not elicit a great enough amount of urgency or context, respectively, to provide useful insights about security fatigue and its effects. To generate more practically applicable findings, it is thus necessary to approach this topic as a mixed-methods one.

This design is epitomized by the primary data collection tool, a survey, which was composed of two sections. Section 1 was concerned with gathering open-ended answers to questions that had, in previous studies [2], been found to elicit responses showing security fatigue. Section 2 was composed of ten passwords that had been chosen at random and scored with zxcvbn prior to survey distribution. Respondents would rate the passwords' strength on a Likert scale going from 1 to 5, or in other words, Very Weak to Very Strong, which was the same scale used by zxcvbn. The survey was distributed to the UTC campus student population, which according to an official 2019 report totaled 11651 students, a combined total of Undergraduates and Graduates [15]. An attempt was made to distribute as widely as possible to reach every department the author could, with the ultimate goal of obtaining 377 responses, which would equate to a statistically significant sample (5% error, 95% confidence) for the aforementioned population size. However, only 135 responses were received, putting the margin of error at around 8% instead of 5%.

Once the survey closed and the responses were collected, the analysis process followed as such: Three categories- Potentially Fatigued, Potentially Non-Fatigued, and Inconclusive- were created for grouping responses during coding of Section 1, and a tag system was developed to help offset the limited number of categories by providing additional depth and analytical potential to each response. For Section 2, the sum score given for the passwords in that section would be calculated as a fractional part from 0 to 1. The sum of each respondent's ten password scores would then be compared against the true zxcvbn calculated sum for the ten passwords.

With both sections finished, they could be compared to begin drawing conclusions. Individually, each response could be analyzed to fall into one of five categories corresponding to one of the five combinations possible with the data, those being: **Fatigued** with **high** trust, **Fatigued** with **low trust**, **Non-Fatigued** with **high trust**, **Non-Fatigued** with **low trust**, and **Inconclusive** (considering Section 1 only). Collectively, the dataset's results could be analyzed to answer the research questions and hypotheses.

In the next chapter, the results of the data analysis will be reviewed.

IV – Results

This chapter shall review the results of the thesis. As a reminder, this thesis is being conducted for the purpose of finding any correlation between the presence of security fatigue in a person and their perceptions of password strength. This chapter will be laid out as follows: (1) Research Questions/Hypotheses, (2) Results, and (3) Summary.

Research Questions/Hypotheses

Below are the research questions and hypotheses to be discussed in this section. Additionally, a breakdown of the hypotheses into their Null/Alternate form is provided.

R1: Is security fatigue present in the sample?

R1.1: What is the proportion of fatigued respondents to non-fatigued respondents?

R1.2: Does the presence of security fatigue have a consistent, observable effect on the perceived strength of passwords?

H1: It is hypothesized that security fatigue will be dominantly represented in the sample, such that the proportion of fatigued vs. non-fatigued respondents will be larger by some arbitrary amount.

H_0: $\sum_{PF} <= \sum_{PNF}$

H_1: $\sum_{PF} > \sum_{PNF}$

H2: It is hypothesized that security fatigue will have a negative impact on the perceived strength of passwords, i.e., those displaying signs of security fatigue will show a higher trust in weaker passwords overall than those who do not show fatigue.

H_0: $\bar{X}_{\text{FATIGUED PASSWORD STRENGTH}} <= \bar{X}_{\text{NON-FATIGUED PASSWORD STRENGTH}}$

H_1: $\bar{X}_{\text{FATIGUED PASSWORD STRENGTH}} > \bar{X}_{\text{NON-FATIGUED PASSWORD STRENGTH}}$

Results

The results in this section are separated out into two sub-sections: Fatigue results & Tags, and password strength results. The research questions and hypotheses of this thesis will be covered in the sections that most pertain to them.

In total, 135 viable responses were acquired during the data collection period of this study.

Fatigue Results

The distribution of potentially fatigued (PF) vs. potentially non-fatigued (PNF) responses in the collected dataset were heavily skewed in favor of the potentially non-fatigued group. The proportion of PNF to PF responses was roughly 2:1 (63.70% vs. 28.89% or 86 vs 39). The INC responses accounted for about a small 7.41% (10) of responses. Of the PF responses, none were purely fatigued across all five questions used to elicit the fatigue state. Of the PNF responses, 24 were purely non-fatigued, meaning all five of their answers showed no signs of fatigue. Of the INC responses, only 1 was deemed inconclusive because each answer was inconclusive. All other 9 INC responses were balanced out in their PF and PNF answers, meaning they each had 1 answer marked as INC.

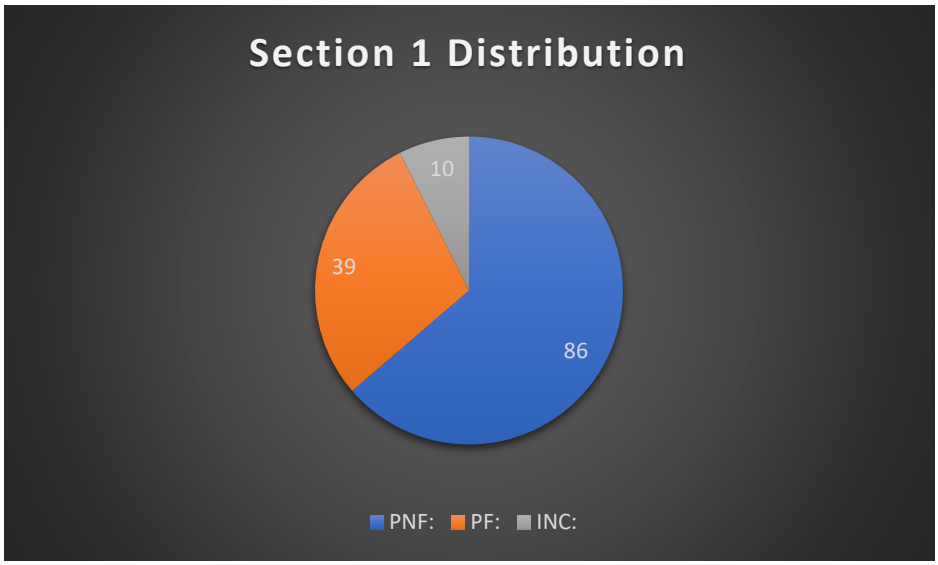

Fig 1: Fatigue Distribution

These results imply that contrary to the author's initial assumption about the prevalence of fatigue among these respondents, most of them are not dominantly showing signs of fatigue. This supports the null hypothesis of H1, and it also supplies answers to R1 and R1.1. This finding is, of course, based on the assumption that having a majority of PNF answers to the five questions over a minority of PF answers makes the respondent PNF overall. However, if a laxer interpretation of the results is taken, such that any sign of fatigue marks the respondent as fatigued, the results change dramatically.

Based on the observation that only 24 of the PNF responses were purely non-fatigued, and using the interpretation mentioned above, the number of PF responses would jump to 110, including the 9 INC responses that balanced out. That would raise the proportion of PF vs PNF responses to about 81.48% vs 17.80% or 110 vs 24. This tremendous change would then support the alternative hypothesis of H1. Prior research, as was covered in previous chapters, could support this laxer interpretation. It is to the belief of the author, however, that if given the benefit of the doubt, most people may harbor fatigue to a degree but still engage in good practices out of

either necessity or reluctant acceptance. For this thesis, the benefit of the doubt is given, but others should consider their own thoughts on the matter when interpreting the leaning of the results.

H1	R1	R1.1
Null hypothesis accepted. Majority of dataset is not fatigued.	Dataset contains fatigue.	Accepted Interpretation ~63.70% PNF vs. 28.89% PF or 86 vs 39 Alt. Interpretation ~81.48% PF vs 17.80% PNF or 110 vs 24

Table 2: H1, R1, & R1.1 Findings

Tags

Tags are descriptors added to each response to provide extra insight that may otherwise be lost with a simple category assignment. For the purposes of proving the research questions or hypotheses of this thesis, tags are not relevant, but the author wished to include this tagging system in order to document common themes in the answers. The results of the tags analysis will only be covered for the PF and PNF groups.

Overall, 11 tags were developed in accordance with themes in the dataset. Only 10 will be listed, as one tag only appeared once throughout the entire dataset. Additionally, only the analysis of the top three will be covered. The tags, their descriptions, their frequencies, and their approximate percentage of responses that they appear in, proceed as follows:

Tag	Description	Frequency	Approx. Perc. Freq.
Frustration	Respondent indicates anger/frustration.	72	53%
Forced Adherence	Respondent indicates that they are forced to adhere to certain rules, guidelines, or practices.	55	41%

Difficulty Remembering	Respondent indicates difficulty in remembering certain things, like passwords, rules, etc.	49	36%
Lax	Respondent indicates that they can be lax in regard to something, such as password security or guidelines adherence. Differentiated from Frustration by a lack of animosity.	31	23%
Too Much	Respondent indicates that there is an aspect of something which overwhelms them.	24	18%
Less Concern Locus	Uneven distribution of security mindfulness.	19	14%
Not Target	Respondent does not believe they are a target for cyber threats	11	8%
Pointless	Respondent feels that something security-related is pointless or of little importance	9	7%
Need More Security	Respondent thinks that a security aspect is lacking in some way and should be improved.	8	6%
Victim	Respondent has been a victim of a cybercrime before.	7	5%

Table 3: Tags List

Of the top three tags, it can be seen that frustration is dominant. This was typically expressed in negative words synonymous with anger- or with the actual words “frustrated” or “frustration.” It was also usually expressed with regards to questions 2 or 3, which elicited feelings about password complexity requirements and mandatory password resets, respectively.

Forced adherence was typically indicated in response to question 5, which asked whether the respondent always adhered to password complexity guidelines. The question may have been misunderstood as mandatory requirements set by organizations instead of recommended guidelines, such as those published by NIST or other standards bodies.

Difficulty remembering was often associated with questions 2 or 4, which dealt with feelings about password complexity requirements and whether password strength/complexity was a concern for the respondent, respectively.

Breaking down the top three tags based on their distribution within both the PF and PNF groups, these are the results:

Tag	Potentially Fatigued	Potentially Non-Fatigued
Frustration	24 (~61.5% of responses)	44 (~51.2% of responses)
Forced Adherence	15 (~38.5% of responses)	36 (~41.9% of responses)
Difficulty Remembering	23 (~59% of responses)	23 (~26.7% of responses)

Table 4: Top 3 Tags Breakdown

Note that all percentages are based off the accepted interpretation of the fatigue results.

As can be seen for the PF group, the Frustration and Difficulty Remembering tags were very close in frequency. However, cases where they both appeared in a response in the PF group were rarer (11, or ~28.2% of PF responses), indicating they may not be connected factors. Forced Adherence showed up the least of the top three tags in the PF group.

By contrast, Difficulty Remembering was the lowest of the top three tags in the PNF group. This makes sense considering that fatigued individuals have some tolerance for password management, or at the least employ strategies to reduce the workload of it, such as using password managers. Forced Adherence was marginally greater in its overall distribution compared to the PF group, but the two were close, which again may be related to wording of

question 5 in the fatigue section of the survey. The most common tag of the top three in the PNF group was, again, Frustration. Compared to the overall distribution of the PF group, it was over 10% less, but it still tinged over 50% of the PNF responses.

Password Strength Results

For judging the password strength results, a baseline score of the ten provided passwords was determined to be 22/50, or .44. The distribution for this section for Potentially Fatigued, Potentially Non-Fatigued, and Inconclusive groups is shown in the chart below.

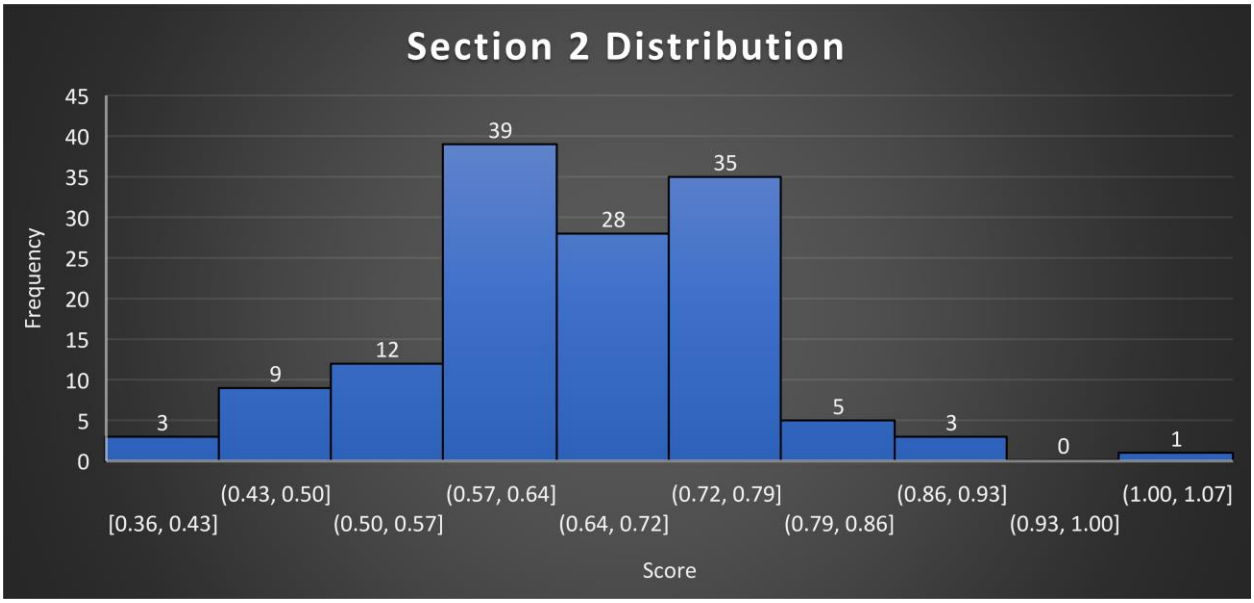

Fig 2: Password Scoring Distribution

As is clear from the chart, most responses fell above the true score threshold of .44. The mean, standard deviation, median, mode, and variance rounded to three decimal places of all groups, all groups sans INC, the PF group, the PNF group, and the INC group are listed below.

Group	Mean	Stan. Dev.	Median	Mode	Variance
All groups	.655	.104	.660	.660	.011
All groups (no INC)	.652	.107	.660	.640	.011
PF	.654	.111	.640	.620	.012

PNF	.651	.106	.660	.760	.011
INC	.698	.048	.690	.660	.002

Table 5: Password Scoring Group Results

The numbers for the INC group are, of course, vastly different from the other groups, owing to the small number of responses deemed INC (only 10). For that reason, and because this thesis is primarily concerned with the PF and PNF groups, only the results of the combined scores without INC will be covered in future discussion. The differences between the combined scores with and without INC are fairly small, but for stats like the mean it is an important distinction. The conclusions of this thesis are not predicted to change regardless of whether the outlier INC group is excluded from final consideration or not.

The main takeaway from this data is that is that the mean of the combined groups (.652) is significantly higher than the true mean of the given passwords (.44). This implies that respondents (students) generally scored the strength of passwords higher than they really were. This is also true for each subgroup (PF, PNF, INC), showing that the overestimation of strength is not tied to fatigue state.

This is further reinforced by t-tests with equal variances of the PF and PNF groups. Compared against an alpha of .05, the one-tailed test reports a P score of about .429, and the two-tailed test reports a P score of about .857. There is no statistical difference between the two means, and thus, H2 must accept the null hypothesis as well, that potentially fatigued respondents do not show insecure perceptions of password strength to a greater degree than non-fatigued respondents do, at least by any significant margin. Additionally, this answers R1.2, that there is no consistent observable effect of security fatigue on perceived password strength.

H2	R1.2
Null hypothesis accepted No statistical difference between means of Potentially Fatigued and Potentially Non-Fatigued groups.	There is no consistent observable effect on perceived password strength as a result of security fatigue.

Table 6: H2 & R1.2 Findings

Summary

In this chapter, the results on analysis of 135 responses was discussed. Two hypotheses and three research questions were tested and answered, respectively.

During discussion of the fatigue data results, it was reported that the proportion of PNF to PF responses was roughly 2:1 (63.70% vs. 28.89% or 86 vs 39). This made it so that the null hypothesis of H1 was accepted, and it also answered research questions R1 and R1.1. Musings on the number of purely non-fatigued vs. purely fatigued responses brought up an interesting conflict of interpretations, however. By adopting a laxer stance that says any fatigue in a set of five questions causes the whole response to be fatigued, the proportion of PF to PNF responses changes to about 81.48% vs 17.80% or 110 vs 24, including the 9 non-pure INC responses as part of PF. This change would have a dramatic impact on the conclusions of the thesis, but it was decided to continue using the strict interpretation of the results, where a majority PF or PNF answers in the set of five questions made the response PF or PNF respectively. This was decided based on the grounds that most people may harbor fatigue to a degree, but they still engage in good practices out of either necessity or reluctant acceptance, thus their practice has not necessarily suffered from their beliefs yet.

The tags system, a support for the fatigue questions, was also reviewed. The top three tags based on frequency were: Frustration (FR), Forced Adherence (FA), and Difficulty Remembering (DR), in that order. FR was described as “Respondent indicates anger/frustration,”

which was dominantly linked to survey questions 2 and 3. FA was described as “Respondent indicates that they are forced to adhere to certain rules, guidelines, or practices,” which was dominantly linked to survey question 5. Finally, DR was described as “Respondent indicates difficulty in remembering certain things, like passwords, rules, etc,” which was dominantly linked to survey questions 2 and 4.

These three tags were then broken down by their presence in the PF and PNF groups specifically. FR and DR were more common in the PF group than the PNF group, but FA was almost proportionally equal between them. Notably, DR was significantly less present in the PNF group than the PF group, potentially owing to higher tolerance for password management or the use of software to reduce the workload of it.

Finally, in conjunction with the fatigue results, the password strength results were reviewed. The true mean score of the ten provided passwords on the survey was .44, and the mean score of PF+PNF groups was .652. This implied that respondents (students) generally scored the strength of passwords higher than they really were. This was also true for all subgroups, including INC, when analyzed on their own, with each reporting above .651, up to .698 with the outlying INC group. This showed that perceived password strength was not tied to the fatigue state. This result is then further supported by t-tests with equal variance between the PF and PNF groups. Compared against an alpha of .05, the one-tailed test reported a P score of about .429, and the two-tailed test reported a P score of about .857.

With no statistical difference between the means of the two groups, the null hypothesis of H2 was accepted, that potentially fatigued respondents did not show insecure perceptions of password strength to a higher degree than their non-fatigued counterparts. R1.2 was also

answered by this result. Below is a table of research questions and hypotheses and their corresponding results.

RQ/H	Result
R1	Dataset contains fatigue.
R1.1	Accepted Interpretation ~63.70% PNF vs. 28.89% PF or 86 vs 39 Alt. Interpretation ~81.48% PF vs 17.80% PNF or 110 vs 24
R1.2	There is no consistent observable effect on perceived password strength as a result of security fatigue.
H1	Null hypothesis accepted. Majority of dataset is not fatigued.
H2	Null hypothesis accepted No statistical difference between means of Potentially Fatigued and Potentially Non-Fatigued groups.

Table 7: Research Questions & Hypotheses Combined Findings

In the next chapter, the conclusions to this paper and its associated topics will be covered.

V – Conclusions

This chapter shall conclude the findings of this thesis, discuss some shortcomings and limitations of the work presented herein, and provide recommendations for further work on this subject. The outline is as follows: (1) Research Questions/Hypotheses, (2) Discussion of Findings, (3) Limitations, (4) Recommendations to Leader and Practitioners, (5) Recommendations for Future Research, and the (6) Summary.

Research Questions/Hypotheses

Below are the research questions and hypotheses to be discussed in this section. Additionally, a breakdown of the hypotheses into their Null/Alternate form is provided.

R1: Is security fatigue present in the sample?

R1.1: What is the proportion of fatigued respondents to non-fatigued respondents?

R1.2: Does the presence of security fatigue have a consistent, observable effect on the perceived strength of passwords?

H1: It is hypothesized that security fatigue will be dominantly represented in the sample, such that the proportion of fatigued vs. non-fatigued respondents will be larger by some arbitrary amount.

H_0: $\sum_{PF} <= \sum_{PNF}$

H_1: $\sum_{PF} > \sum_{PNF}$

H2: It is hypothesized that security fatigue will have a negative impact on the perceived strength of passwords, i.e., those displaying signs of security fatigue will show a higher trust in weaker passwords overall than those who do not show fatigue.

H_0: $\bar{X}_{\text{FATIGUED PASSWORD STRENGTH}} <= \bar{X}_{\text{NON-FATIGUED PASSWORD STRENGTH}}$

H_1: $\bar{X}_{\text{FATIGUED PASSWORD STRENGTH}} > \bar{X}_{\text{NON-FATIGUED PASSWORD STRENGTH}}$

Discussion of Findings

The final results for the two hypotheses of this thesis will be discussed in this section. Discussion will primarily take place through answering the research questions which tie into the hypotheses and then addressing the relevant hypothesis itself. Exact measurements will not be given as such would be redundant with the previous chapter, but the conclusions of the findings will be reiterated.

Hypothesis 1

For R1, it was evident in the results that the collected dataset did contain responses showing security fatigue. Such was clear from the markedly negative language directed toward security topics or objects, such as the expression of frustration toward password reset practices.

For R1.1, the results were a bit more unexpected. The ratio of fatigued to non-fatigued responses was notably skewed in favor of non-fatigued, which was not the assumed reality of the group being sampled from. The assumption of greater fatigue in the sample was made primarily from the author's own experiences with these topics and conversations with other students in the past.

This, of course, meant that H1 could not find support for the alternative hypothesis. This could, however, be different if a less strict categorization technique was employed. As was covered in Chapter 4, out of the five questions used to determine the fatigue state of the respondent, the majority category represented in those five questions became the overall category. This meant responses with 2 non-fatigued and 3 fatigued answers were categorized as fatigued, but responses with 3 non-fatigued and 2 fatigued answers were deemed non-fatigued. With a less stringent approach that equates any fatigued answer to the overall state of being fatigued, H1 would accept the alternative hypothesis by a large margin. Interpretation, as with most qualitative analyses, is partially up to the interpreter.

Hypothesis 2

For the last research question, R3, the presence of security fatigue did not seem to have any effect on the perceived strength of passwords among students. Rather, there was an almost similar mean perceived strength between the fatigued and non-fatigued groups. That mean was also well above the true mean strength score of the passwords given for scoring. A possible cause for this similarity between groups may, at first, be attributed to the presence of some fatigue in a large portion of the non-fatigued responses, but as will be covered soon, the mean of the purely non-fatigued group does not actually differ much from the mean of the loosely interpreted fatigue group (the vast majority). Thus, it can be assumed that there is something else causing the results to clump together. Password managers could be one cause, as they reduce the burden of remembering passwords and often offer to generate new ones, so a user of such a tool may not have a realistic outlook on password strength after becoming comfortable with a manager doing all the heavy lifting. Both fatigued and non-fatigued individuals have reasons to use such tools as well, with the former choosing it because it reduces the memory burden or

provides convenience, and the latter choosing it because it possesses security functions that they deem valuable.

While the results of analysis could not show any effect of security fatigue, it did show consistently inaccurate perceptions of password strength among the sampled students. Explanations for this could range anywhere from lack of education on the topic of password security to apathy or cultural influences, but it is unlikely that any of the respondents scored them while being unfamiliar with passwords, as that is an extremely unlikely scenario for this sample.

The implications for H2 are, like H1, that there was no evidence to support the alternative hypothesis. This would not change even in the event of the looser interpretation of categorization being used, as the mean score for the purely non-fatigued group (.631) is only slightly lower than the score for the loosely interpreted fatigued group (.657).

What Could Be Improved

Some things about this thesis, specifically the data collection portion of it, could have used some extra attention.

Firstly, the survey needed more work, such as a pilot study, to be air-tight in the quality of its questions. Some of them, namely questions 1 and 5, elicited responses that were of dubious usefulness. Question 1 overall elicited many identical responses from students across the board, barring a few outliers. The responses given read very much like the usual guidelines one would hear or read for passwords, such as using an assortment of different keyboard characters, making the password long, and making it something only the creator would know. While such responses are not necessarily useless, in that they establish that the majority of respondents know what

precautions to take for password creation, they nonetheless bias the dataset toward the non-fatigued camp.

Question 5 was poorly worded, asking respondents if they always "adhere to password complexity guidelines?" This was poorly worded because the distinction between "guidelines" and "requirements" may not always be clear, and many respondents seemed to show this. The Forced Adherence tag was greatly prevalent throughout the dataset as a result of respondents interpreting "guidelines," which can be pieces of advice or rules of thumb that are not necessarily mandatory, as "requirements," which would be the mandatory criteria enforced during something such as password creation. As a result of this confusion, question 5's usefulness in determining security fatigue was reduced, as the true meaning behind each student's answer was often obscured behind the issue of Forced Adherence.

Recommendations to Leaders and Practitioners

If the results of this study are to be taken for practical application, it should be done so under the following context and in the following way, or similar ones as determined by the user.

These results are primarily going to be applicable to public universities, especially those in the Southeast United States, that have a wide range of degree programs across many disciplines. These results are also most useful to said universities that do not have, or have not widely adopted, a 2-Factor Authentication (2FA) system to supplement the use of passwords. 2FA will reduce the risk posed by passwords, thus reducing the risk posed by those with insecure password perceptions which may bleed into their practices.

Security fatigue, whether interpreted strictly on a majority *x out of y* scale or loosely on a *x in y* scale, is almost certainly present in any given student population. Pressures from

coursework, technology knowledge gaps, and students' personal lives may not permit them the constant vigilance or patience needed to put up with increasing security requirements and the like. However, the effect this has on their idea of what constitutes a strong password seems negligible. According to the data, most of the students, fatigued or not, are going to think about password strength in similar ways to each other. That is to say, they will generally view a weak password as stronger than it is, and thus they may also create such insecure passwords. This is a threat to both the student and university, as a compromised student account could lead to further complications elsewhere on the network.

Other research surrounding security fatigue proposes means to reduce the fatigue in a person, but how useful this would be for changing their perception of password strength is unclear, as the data suggests it remains, on average, quite steady regardless of fatigue state. Considering that, the focus should not be on trying to reduce security fatigue as much as it should be providing password education, or even better, authentication alternatives. 2FA is a powerful tool to augment the meek password, and if such a system has not yet been implemented at the university, or simply hasn't been advertised, then more effort and funding should go toward doing so. The various costs both in time and money to setup such a system may be scrutinized, but the overhead from its installation will be paid back in the form of hardened security and, if implemented in a user-friendly manner, greater client satisfaction from both camps, fatigued and non-fatigued. 2FA would bring enhanced peace of mind without adding as significant a user obligation as increasing the length of their password, which would appeal to those seeking easy solutions as much as it would to those who value their security at any cost. Furthermore, while password strength perceptions in this sample overall were less secure than ideal, they were not excessively so, indicating at the very least that these students make an effort

to recognize bad passwords. Such individuals may be more receptive to alternative authentication schemes, given that it would reduce the effort they need in order to be compliant and secure.

In the interim period between having 2FA and not having it, educational resources about password security should be made readily available and distributed to students in an easily digestible format. Additionally, password managers and other helpful tools to reduce the self-security workload should be pushed as well.

All the advice given here has been centered around students, but of course, any sweeping change like implementing 2FA or more actively spreading educational resources are going to benefit all other groups as well, such as faculty.

Recommendations for Future Research

It is worth considering that the study which inspired this one was using data from in-person interviews, where the richness of human interaction can reveal far more detail and nuance than a mere survey can. Given time and resource constraints, this thesis had to make a sacrifice in that regard to acquire enough data for a thorough analysis. However, if a researcher or team were to recreate this study using interviews and improve upon the noted shortcomings of it, then even more insightful discoveries could be made.

Additionally, there are some avenues that exist that may branch off from this work and the works of others referenced within it. Most importantly, confirmation of the findings of this research would be of great interest, whether that is done through the aforementioned interviews or some other means. Others may also seek to find if varying stages of security fatigue, such as those used in the model made by Tanimoto and Hatashima [12-13], show different password

scoring results. The findings of such a study could reveal more about the influence of security fatigue beyond what this one or any other study has done before.

Summary

In this chapter, and this work as a whole, three research questions and two hypotheses were answered and tested respectively, producing conclusions that ran contrary to what the author initially believed would be the case.

Hypothesis 1, which posited that the proportion of fatigued to non-fatigued respondents in the sample would be greater, was not supported by the facts. The null hypothesis was accepted, given that the non-fatigued group was significantly larger than the fatigued group. This conclusion was influenced primarily by two factors in the data collection and analysis of the work: Issues with a few survey questions and a rigid stance on results interpretation. In the former, two questions of the survey suffered from distinct problems, with one biasing the dataset and the other being misinterpreted due to wording. In the latter, a rigid interpretation of the results was used when categorizing responses, which ended up creating a situation wherein the results could swing between very non-fatigued or very fatigued depending on one's choice of interpretation. In the end, the rigid variant was kept, as the author felt it was appropriate to give respondents the benefit of the doubt about their fatigue state. Research questions 1 and 1.1 were answered alongside hypothesis 1, as there was fatigue present in the sample and the proportion of the groups was analyzed.

Hypothesis 2, which guessed that the sample would show security fatigue having a negative impact on perceived password strength, was also unable to support itself. The null hypothesis, that there was no negative effect, was taken instead. However, this is not to say that the sample was primarily leaning toward secure scoring of passwords, far from it in fact. Both

groups, fatigued and non-fatigued, showed very similar mean scores, both of which were well above the true strength score of the given passwords. Even in the case where the loose interpretation of results was taken, the differences between the two groups were still negligible. This conclusion also satisfied research question 1.2, which desired to see whether a consistent and observable effect on perceived password strength was caused by security fatigue, which it was not, as no meaningful difference could be found between the groups.

With these answers in mind, the problem of how to deal with security fatigue's effect on perceived password strength among students became clear: Do not focus on the fatigue. As far as the data suggests, there is no real difference in the average fatigued student's mind vs a non-fatigued one when they are asked to rate the strength of a password. Thus, when they are told to pick a password, the ones they choose are likely to be less secure than they think they are. This threat can be alleviated with more dissemination of educational security materials in a simple and engaging format, as well as the recommendation for students to use things like password managers, which can reduce self-security workload. The ultimate goal of any university trying to reign in insecure password use, however, is to implement 2-Factor Authentication, or to push it more aggressively if it hasn't been widely adopted yet.

The time of the password is nigh, or rather, has been. Yet, the world cannot shake the curse that is passwords, and they've only become more aggravating as the standards for them increase year after year. As a byproduct of this and other security efforts, a fatigue has set in and threatened to uproot many of the teachings of security professionals. Or perhaps not? Perhaps, at least in the case of determining the strength of passwords, fatigued and non-fatigued students alike have become disillusioned with the password's supposed simplicity. This thesis has determined that, among a student population, security fatigue has no clear impact on their

perceptions of password strength. Rather, students from both sides of the aisle give credit where it is not due, thinking highly of passwords that should never be used in modern times. This may thus show that, in a limited context that should be carefully applied elsewhere, security fatigue may not be an impactful factor in why people choose insecure passwords.

References

[1] G. Basset et al, “2020 Verizon Data Breach Investigations Report,” Verizon, vol. 2020/issue 6, June 2020. Accessed: August 2020. [Online]. Available: https://doi.org/10.1016/S1361-3723(20)30059-2.

[2] S. Furnell and K-L. Thomson, “Recognizing and addressing ‘security fatigue’,” *Computer Fraud & Security*, vol. 2009, iss. 11, pp. *7–11*, Nov, 2009, doi: *https://doi.org/10.1016/S1361-3723(09)70139-3*.

[3] R. B. Johnson, & A. J. Onwuegbuzie. “Mixed Methods Research: A Research Paradigm Whose Time Has Come,” *Educational Researcher*, vol. *37*, iss. *7*, pp. *14–26*, Oct, 2004, doi: *https://doi.org/10.3102/0013189X033007014.*

[4] M. L. McHugh, “Interrater reliability: the kappa statistic,” in *Biochemia medica*, Oct 2012. [Online] Available: https://pubmed.ncbi.nlm.nih.gov/23092060/

[5] M. Q. Patton, “Enhancing the quality and credibility of qualitative analysis,” in *Health Services Research*, Dec 1999. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/10591279/

[6] B. Stanton, M. F. Theofanos, S. S. Prettyman and S. Furman, "Security Fatigue," in IT Professional, vol. 18, no. 5, pp. 26-32, Sept.-Oct. 2016, doi: 10.1109/MITP.2016.84.

[7] “Student Age Fall 2015 through Fall 2019,” University of Tennessee at Chattanooga, May 2020. Accessed: August 2020. [Online]. Available: https://new.utc.edu/sites/default/files/2020-07/student-age-fall15-fall19.pdf.

[8] A. J. Onwuegbuzie and K. M. T. Collins, “A Typology of Mixed Methods Sampling Designs in Social Science Research,” in *Qualitative Report*, Jun 2017. [Online]. Available: https://eric.ed.gov/?id=EJ800183.

[9] S. Furnell and K-L. Thomson, “From culture to disobedience: Recognising the varying user acceptance of IT security,” *Computer Fraud & Security*, vol. 2009, no. 2, pp. 5-10, Feb, 2009. Accessed: August, 2020, doi: https://doi.org/10.1016/S1361-3723(09)70019-3. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1361372309700193.

[10] A. Sasse and I Flechais, “Usable Security: Why Do We Need It? How Do We Get It?,” *ResearchGate*, Jan, 2005. Accessed: August, 2020, [Online]. Available: https://www.researchgate.net/publication/316236669_Usable_Security_Why_Do_We_Need_It_How_Do_We_Get_It.

[11] F. Belanger, “When users resist,” *Pamplin College of Business Magazine*, Fall, 2011. Accessed: August, 2020, [Online]. Available: https://web.archive.org/web/20190314184452if_/https://www.magazine.pamplin.vt.edu/fall11/passwordsecurity.html. Author Note: Original webpage has disappeared. Article cannot be located elsewhere.

[12] S. Tanimoto, K. Nagai, K. Hata, T. Hatashima, Y. Sakamoto and A. Kanai, "A Concept Proposal on Modeling of Security Fatigue Level," *2017 5th Intl Conf on Applied Computing and Information Technology/4th Intl Conf on Computational Science/Intelligence and Applied Informatics/2nd Intl Conf on Big Data, Cloud Computing, Data Science (ACIT-CSII-BCD)*, Hamamatsu, Japan, 2017, pp. 29-34, doi: 10.1109/ACIT-CSII-BCD.2017.30.

[13] T. Hatashima et al., "Evaluation of the Effectiveness of Risk Assessment and Security Fatigue Visualization Model for Internal E-Crime," *2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)*, Tokyo, 2018, pp. 707-712, doi: 10.1109/COMPSAC.2018.10323.

[14] X. de C. de Carnavalet and M. Mannan. (22 Feb 2014). From Very Weak to Very Strong: Analyzing Password-Strength Meters. Presented at NDSS Symposium 2014. [Online]. Available: 10.14722/ndss.2014.23268.

[15] “Enrollment by Classification and College,” University of Tennessee at Chattanooga, May 2020. Accessed: September 2020. [Online]. Available: https://new.utc.edu/sites/default/files/2020-07/enroll_classification_college.pdf.

Appendix A: Informed Consent

INFORMED CONSENT

Security Fatigue and its Effects on Perceived Password Strength Among University Students

You are being invited to participate in a research study about the effects of the "security fatigue" phenomenon on how students view the strength of passwords. This study is being conducted by Chase Carroll (csy385@mocs.utc.edu, (731)468-4818) with support from thesis director David Schwab (David-schwab@utc.edu) at the University of Tennessee at Chattanooga.

The questionnaire(s) will take about 6 minutes to complete.

This survey is anonymous. Do not indicate your name on the survey. No identifiable information will be gathered through the questions in this survey or automatically by QuestionPro, and all responses will be kept secure in a digital format according to relevant data storage standards. No one will be able to identify you or your answers, and no one will know whether or not you participated in the study.

Your participation in this study is voluntary. By selecting "Yes" below you are voluntarily agreeing to participate and you are acknowledging that you are **18 years of age or older**. You are free to stop answering questions at any time or to decline to answer any particular question you do not wish to answer for any reason. If you are younger than 18, do not proceed.

Research at the University of Tennessee at Chattanooga involving human participants is carried out under the oversight of the Institutional Review Board. Address questions or problems regarding these

activities to Dr. Susan Davidson, UTC IRB Chair, email: susan-davidson@utc.edu; phone: (423) 425-5568.

Appendix B: Survey Questions

Section 1: Fatigue-targeting Questions

What makes a strong password? Do NOT provide examples, only give your opinion.

What is your opinion on password complexity requirements?

What are your feelings about mandated password changes?

Is password strength/complexity a concern for you? Why or why not?

Do you always adhere to password complexity guidelines? Why or why not?

Section 2: Password Strength Perception Questions

	Very Weak	Weak	Average	Strong	Very Strong
asdfghjkl1					
YAgjecc826					
babygurl1					
Francesco					
1v7Upjw3nT					
P@ssw0rd					
vC3qqeA1					
234dak61					
pE3^&zSx"DP					
>7ncZm					

Appendix C: List of Tables

Table #	Table Name	Pg #
1	Alignment of Survey Sections to Research Questions/Hypotheses	31
2	H1, R1, & R1.1 Findings	39
3	Tags List	39
4	Top 3 Tags Breakdown	41
5	Password Scoring Group Results	42
6	H2 & R1.2 Findings	44
7	Research Questions & Hypotheses Combined Findings	46

Appendix D: List of Figures

Figure #	Figure Name	Pg #
1	Fatigue Distribution	38
2	Password Scoring Distribution	42

Appendix E: IRB Approval Number

UTC IRB # 20-006