Privacy and Security Attitudes, Beliefs and Behaviours: Informing Future Tool Design

by

Janna-Lynn Weber

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Computer Science

Waterloo, Ontario, Canada, 2010

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Usable privacy and security has become a significant area of interest for many people in both industry and academia. A better understanding of the knowledge and motivation are important factors in the design of privacy and security tools. However, users of these tools are a heterogeneous group, and many past studies of user characteristics in the security and privacy domain have looked only at a small subset of factors to define differences between groups of users.

The goal of this research is to critically look at the difference between people, their opinions and habits when it comes to issues of privacy and security. By conducting and analyzing 32 in-depth qualitative interviews the heterogenous nature of this community.

Attitudes and actions around the principles of knowledge with tools and motivation for self-protection. To show how, together, the attributes of different groups of users can be used as a set of privacy and security personas, or prototypical privacy and security tool users. These personas are a tool for incorporating a broader understanding of the diversity of users into the design of privacy and security tools. The goal is to present the personas as a starting point for a detailed discussion about the different types of target users in the design of privacy and security technologies, in the hopes that this discussion highlights new design opportunities for usable privacy and security.

Acknowledgements

I would like to thank my supervisors, Daniel Berry and Edward Lank for all the support you've given me. For allowing me to find my own way, guiding me when I needed it and believing my work. I have learned many things from both of you and am very grateful for your guidance. I would also like to thank Michael Terry and the guys in the HCI lab, the advice and encouragement has been priceless.

I would like the thank the participants of my study for taking the time to be interviewed and helping me achieve this goal. Those who I didn't know before the interviews, I feel like I know you personally now. Some of you are my close friends and know just how much this really means to me.

Without the support and encouragement of my parents, Bill and Ginger Weber, and my siblings Julie and Jacob, I couldn't have accomplished as much as I have. Thank you for helping me find the strength to keep moving forward and reaching for my goals.

Finally, I could not have done this without the love and support of my boyfriend, Sean. You is my rock. When nothing else makes sense I know I can count on you to make me feel better and figure out my next step.

Contents

List of Tables **vi**

List of Figures **vii**

1 Introduction **1**

1.1 Motivation 1

1.2 Research Goal 4

1.3 Study 4

1.4 Findings 5

1.5 Contributions 6

1.6 Definitions 7

1.7 Organization 8

2 Background **10**

2.1 Foundations of Usable Privacy and Security 10

2.1.1 Privacy vs. Security 12

2.2 Identifying and Understanding Users 12

2.2.1 Qualitative Studies 13

2.2.2 Surveys and Questionnaires 13

2.2.3 Identifying Differences Between Users 15

2.3 Limitations of Past Research 16

2.4 History of Personas 18

3 Study 21

3.1 Overview 21

3.1.1 Qualitative Research Methods 23

3.2 Interviews 24

3.3 Analysis Methodology 26

3.3.1 Creating and Grading Concepts 26

3.3.2 Clustering Participants 26

4 Theory One: Analyzing Views of Self 30

4.1 Analyzing Views of PAS 31

4.2 A Unifying Theme 33

4.3 P-Clustering Based Upon Self Rated Motivation 35

4.4 An Analysis of Motivation-Based Clustering 37

5 Theory Two: Sharpening the Distinctions 40

5.1 Participant Cluster – The Marginally Aware 45

5.2 Participant Cluster – The Fundamentalist 47

5.3 Participant Cluster – The Struggling Amateur 49

5.4 Participant Cluster – The Technician 51

5.5 Participant Cluster – The Lazy Expert 53

5.6 Shared Traits 55

5.7 Discussion 56

6 Design Implications - Personas 58

6.1 Mark — The Marginally Aware 59

6.2 Robert — The Fundamentalist 60

6.3 Allison — The Struggling Amateur 61

6.4 Patricia — The Technician 61

6.5 Henry — The Lazy Expert 62

6.6 Summary 63

7 Applications for Requirements and Design **64**

7.1 Evaluate User Categories 64

7.1.1 Study Method 64

7.1.2 Results 65

7.2 Evaluate Design 67

7.3 Focus of Concerns 68

7.4 Explain, Understand and Predict Users 70

8 Conclusion **72**

8.1 Goals 72

8.2 Results 73

8.3 Limitations and Future Work 73

8.4 Contributions and Summary 75

APPENDICES **76**

A Terminology **77**

B Participant Profiles **78**

C Study Materials **80**

C.1 Recruitment Script 80

C.2 Interview Script 81

C.2.1 Goal 81

C.2.2 Introduction Script 81

C.2.3 Potential Question List 81

C.3 Thank You Email 85

D The Affinity Diagram **86**

References **90**

List of Tables

3.1 Discovered Concepts 27

3.2 Example of knowledge and motivation grading. Concept: Reasons for making changes 28

3.3 Clusters of participants based on the affinity diagram with knowledge and motivation scores. 29

4.1 Concepts and Subconcepts of the Affinity 31

4.2 Listing of Personal Assessments in order of concern levels 34

4.3 Distribution of participants in clusters compared to previous research. . . . 37

5.1 Dominant and Shared Traits of Each P-Cluster 42

7.1 Participants in the WiFi study. 65

7.2 WiFi participants grouped according to motivation and knowledge 66

List of Figures

1.1 Some of the ways society has generalized users. [39] 3

7.1 Microsoft's User Account Control 67

Chapter 1

Introduction

1.1 Motivation

As software requirements, usability and security share similar properties; both must be considered early in the development process, both serve the needs of many stakeholders, both are critical to a product's success [47]. However, usability and security also seem to be opposing forces. Intuitively people believe that to be secure is to make things harder to do, while to be usable a product must be easy to interact with. Both usability and security are highly desirable features of any product especially in our increasingly digital world. The idea of usable security, the merging of these opposing forces, has been of interest for many years. It is an growing area of study that blends disciplines of security research with human computer interaction. Despite concepts of usable security dating back over a decade, the research area still has many concepts to explore [48].

One of the partially explored areas is how users interact with security software, how users understand concepts like privacy and security (PAS) and what this all means to the creators of products that are both usable and secure. In PAS research it is quite common to see a user community reduced to a single dimension based on some naive observation from survey data. For example, authors state that "The average user has 6.5 passwords, each of which is shared across 3.9 different sites" [21], that "66% of respondents in a survey of 300 corporate users wrote down work-related passwords" [19]; or that, "given the right circumstances, online users easily forget about their privacy concerns and communicate even the most personal details without any compelling reason to do so" [6]. Essentially, researchers claim that there exists a generic user who creates bad passwords, is naive about privacy, or is undereducated about the security issues that affect her[1]. [21, 2, 3].

[1] To avoid both the use of plural pronouns to describe a single person and the heavy he or she and related

A significant problem exists when prospective users of technology are generalized to a single dimension. Software is designed for a generalized user who rarely exists in practice. These generalized statistics can also be misleading. Consider the statistic on users in the previous paragraph. While a user may share their 6.5 passwords across 3.9 sites each, perhaps the user has 5 or 6 passwords, one for each banking sites and email accounts. He then has one throw-away password he uses for other sites that require registration – for example for a cbc.ca account that allows the user to post comments on news stories. While two-thirds of corporate users may write down a password these cases might involve giving access to a boss or trusted colleague to ensure access to materials in the account if necessary while the user is away on vacation. The specific point to be made here is that for every statistic or generalization that seems problematic, reasonable motivations may exist for any user actions. Reducing users to a single statistic, or passing a value judgement on the action may inflate negative perceptions of the user community.

Among many other researchers, Dourish and Anderson [17] highlight the dangers of this reductionist philosophy. Of course, one challenge with analyzing any user community during design is that the goal of the analysis is, at heart, reductionist. Often when looking at a community researchers break users down into common attributes that unite the community. Designers look at practices [18] or demographics [22] or knowledge [17] to determine the characteristics of a user community and to express consolidated characteristics that they wish to address in a new design [7]. By categorizing users on broad generic levels individuality suffers. On some level society knows that different types of users exist, but seem to be still trying to come to terms with how to describe these groups. Unscientifically, Figure 1.1 represents some of the different ways that categorizations of users have formed around password authentication [39].

In 1999, Adams and Sasse [3] first promoted closer examination of security issues from a user point of view. In their paper titled “Users are not the enemy” they note that, contrary to the beliefs of many security researchers, users are not a security risk to be controlled. Instead by understanding users, they can become a participant in ensuring the security of systems they interact with. The need to examine the users of security technology has persisted. More recently Herley [27] has argued that people know what they are doing and the avoidance or manipulation of security rules is a rational rejection based on personal cost benefit analysis. Similarly, issues of privacy concern arise at a personal level where each user has a different response and opinion. Ackerman and Mainwaring wrote a summary of research of privacy issues and human computer interaction, in which they discussed Egan’s chapter on individual differences [14]. Egan notes that user have “even wider variance when considering privacy. One can see that people vary not only in their system performance

constructions, this thesis randomly varies the gender of arbitrary individuals and study participants. In the case that a quotation reveals the gender of its speaker, the revealing part of the quotation is neutered and the neuter term is enclosed in brackets.

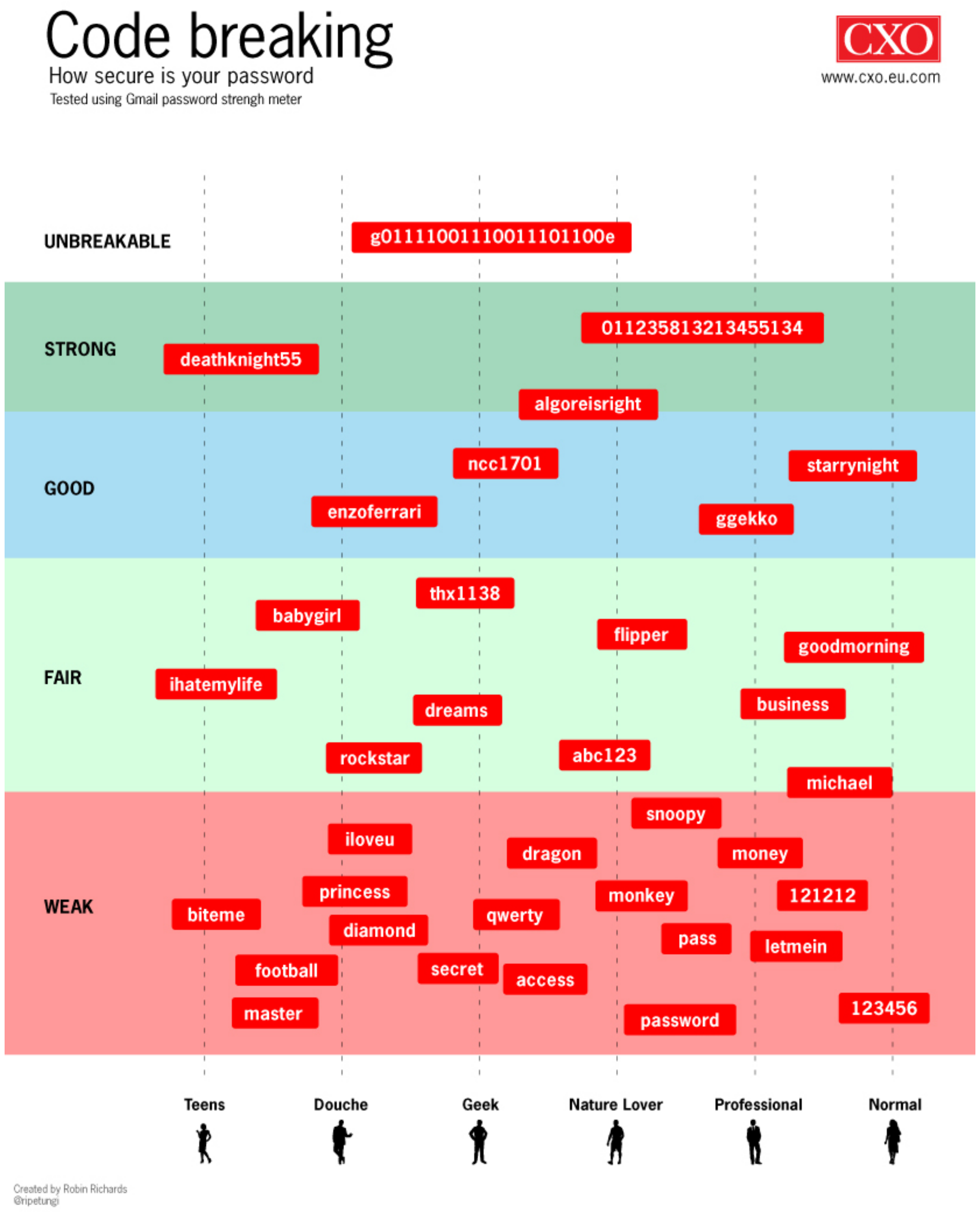


Figure 1.1: Some of the ways society has generalized users. [39]

but also in their understanding of the task and its implications for privacy." [14] However, many PAS solutions still remain one size fits all.

1.2 Research Goal

The goal of this work is to begin to address PAS attributes of users based on a view of users as a heterogeneous yet concrete community. Understanding the breadth of a user community enables better designs by tailoring any single design either to specific attributes of a subset of users, or by tailoring a design to broad attributes shared by several different classes of users of a software system.

Others have shared this goal of trying to better understand the breadth of users within PAS research. Opinion surveys have been collected to assess user awareness and concern with regards to privacy [1, 6, 41]. Based on an analysis of survey data, researchers have segmented users into three broad categories: the marginally concerned; the privacy fundamentalists; and the pragmatic majority. While survey data has proven useful in developing an initial overview of the types of users that exist within the PAS space, survey data is, by its nature, limited. Researchers cannot follow-up on answers provided, and closed questions provide a set of specific responses for a participant to choose between. Additional research is needed to develop a deeper, more nuanced picture of prospective users. Specifically this study explores the following research questions:

- How are people currently interacting and dealing with computer PAS?
- Do the current classifications of users provide an accurate view of the user community?
- How can differences between people allow for a better understanding of the community?
- What attributes allow for clustering of individuals on both a usable and relevant level?
- What design implications could be explored by better understanding and embracing the differences between people?

1.3 Study

To address these goals, this thesis describes the results of an interview study of 32 participants. Of these participants, nineteen were students, recruited locally from the university community. To diversify the user population, interviews were also conducted with thirteen non-student participants recruited online from across North America using VoIP telephony. The interviews lasted 45 minutes to an hour covering a breadth of topics. The interview structure allows for an in-depth conversation about, among other things, the participant's habits, opinions, justifications and skills with regards to PAS. This comprehensive view of the participants is critical to understanding how users deal with PAS issues.

The interviews, once completed, were transcribed and coded. An affinity diagram was constructed to cluster related data. The affinity diagram contained two types of clusters: *Q-clusters*, created from quotes; and *P-clusters* or clusters of participants. Quotes of similar opinions or habits made the basis for a Q-cluster and were eventually labeled as a *concept* of the data. No assumptions were made about concepts before beginning to build the affinity diagram. Instead, the Q-clusters were was built from the bottom upward, using the participants' quotations. While formalizing the concepts of the data, the participants sharing strong commonalities formed the basis for P-clusters that led to the eventual categorizations of users. Initial analysis led to one set of P-clusters based on each user's personal assessment. On closer inspection this theory was insufficient and after re-analyzing the data a second theory revolving around dimensions of knowledge and motivation was used to create more comprehensive P-clusters.

1.4 Findings

Based on initial P-cluster from the participants personal assessments, the interview data agreed with previous survey data. The first set of P-clusters identified two small groups of participants, the marginally concerned and the fundamentalists; and a large community, the pragmatic majority. However, upon closer inspection, the pragmatic majority became a complicated heterogeneous group of its own. Arguably, all groups claimed to be practical in all security decisions, balancing individual cost benefit ratios. The subtleties that exist between participants was masked by a uniform clustering around self-assessment. A finer grain analysis was required to analyze the individuals in the study in a meaningful way.

To address this, the data was reanalyzed. Participants were clustered based on similarities within the interview data (i.e. based on Q-clusters) and five new P-clusters of users were identified. The participants actions and attitudes identified during the analysis of interview data (using Q-clusters) was graded into two broad dimensions: the knowledge with security tools; and the motivation of the participants to protect themselves. Finally, using the gradings of knowledge and motivation five security personas, or archetypical users were created: the marginally aware, the fundamentalist, the struggling amateur, the technician and the lazy expert. While two of these P-clusters, the marginally aware and the fundamentalists, echo classifications from past research, the pragmatic majority is a richer group of participants requiring deeper descriptions. Some within this group expend little effort, but because of their sophisticated knowledge of PAS risks protect themselves efficiently. Some, despite limited knowledge, work diligently to limit the information that they share and pay close attention to security warnings, thus also protecting themselves. As a final step, the Q-clusters that typify each persona were used to create a set of traits for each persona.

1.5 Contributions

This research investigates and discusses the different ways that people think and interact with computer PAS issues through the use semi-structured interviews. The interview format provides qualitative findings across the breadth of user interaction with PAS. While others have used qualitative studies to examine PAS users, the goal of this research is unique. Rather than exploring commonalities of the community at large, the goal is to look at the differences between users, i.e. things that distinguished one user from another.

The second contribution is a categorization of users using coded results of the qualitative interviews. Several research have identified three or four categories of users, but in PAS research these segmentations of the users community have been based on survey data and have only segmented users based on one dimension (e.g. knowledge or motivation) [1, 41, 6, 10]. As well, the categorization has largely been based on the idea of closed-coding, where the distinctive feature used to segment users is identified before analyzing data. The user community is divided by evaluating responses to the pre-specified features. In contrast, this study uses open coding, where users who said similar things were linked without pre-specified features. As a result, the participant clusters are formed in a bottom-up fashion using multiple dimensions of participants' attitudes and behaviours.

In extending the analysis, five distinct P-clusters of users were observed along the dimensions of knowledge and motivation. The dominant and shared traits of these P-clusters provide deep descriptions of the groups. On their own these descriptions can be difficult to interpret or abstract to useful information during the design and development of new tools. For this reason, an additional contribution of this paper is the development of five security personas. The security personas envelop the traits of each P-cluster. Personas are common in product design [13, 37], and as a design tool, the effectiveness of personas have been of interest to HCI and RE researchers [26, 33]. Personas allow developers to personalize and visualize the targets of design, creating more realistic user communities for the artifacts being designed.

Finally, thought experiments are conducted to demonstrate the potential use and power of these personas. Thought experiments examine existing PAS tools to determine how well the tools are received by the user community. The application section also explores the deeper phenomena observed between the clusters of participants.

On a broad scale this thesis contributes to the discussion of users in PAS research. Often PAS users are misunderstood as nothing more than a variable to be controlled. As a result of the research approach, this thesis represents a significant advance in the state-of-the-art. The groups of users are new and provide a more nuanced picture of types of users than has existed in past research. Personas, while popular in design, have not been used in the security and privacy domain.

1.6 Definitions

The terms security and privacy, can be defined in a variety of different ways and often have fluid definitions that vary from context to context. For this reason it is important to define how four potentially ambiguous terms, Security, Privacy, Knowledge and Motivation, will be used in this research. A complete list of terms and definitions is available in appendix A.

Security – pertains to all actions of keeping the user and system safe. Including authentication, network security and software security. Security includes the actions that protect the user, company or system from physical, logical or economic harm. Bruce Schneier wrote "Security is a process, not a product. Products provide some protection, but the only way to effectively do business in an insecure world is to put processes in place that recognize the inherent insecurity in the products." [40] In this research the term security is used as a generalization of the area of security research in order to view the entire process.

Privacy – is both a wide and important issue, for which the definition is undecided. It is something that security tools aim to protect but inevitably often relies on trust and can be a very personal concept. Solove created a taxonomy to describe various situations where privacy can become an issue, his four main categories are: Information collection, information processing, information dissemination and invasion [42]. In this thesis, privacy, pertains to all actions of controlling personal information. However personal information, is left to the participant to describe.

While privacy and security describe different actions, this research keeps the two terms closely coupled. It is difficult to separate one from the other, particularly for tool design as both aspects are equally important. A clear distinction between security and privacy is not critical in this work as understanding the security and privacy aspect of users are interconnected and are both used in design.

For further clarity, knowledge and motivation are also defined. Throughout this thesis themes of knowledge and motivation are used as dimensions to help describe and categorize the participants.

Knowledge – is what the participant knows about PAS. In this work it encompasses many aspects including the validity and level of knowledge, the source of knowledge, the trust that the knowledge creates and the skills that are formed out of that knowledge. When a participant's knowledge is discussed it is a composite of all these things. For example, where the knowledge comes from is important as some participants are learning from university courses and others are learning from the news media. Furthermore leads to trust, not only do user trust the source but how do users determine to trust websites and applications, is that trust based on technological facts or hearsay. When these attributes are combined, a picture of both the extent of an individual's knowledge and the individual's belief in the reliability of the knowledge appears.

Motivation – is a willingness to act and how far that act would go to protect the user. Motivation in this case is about risk management as well as what causes a user to put extra effort into securing information. Questions like how much customization is being added to a security solution? Is the user seeking information to learn about security or privacy issues? This also includes aspects like triggers, when a user is prompted to update their security settings how long before he actually does it?

Some concepts contribute to both the definitions of motivation and knowledge of this thesis. Additionally how practical the user is or believes himself to be is a function of both knowledge and motivation.

1.7 Organization

The remainder of the thesis is organized as follows. Chapter 2 highlights the research landscape and provides an overview of some of the extensive body of work seeking understanding of users in security and privacy. Initially the foundations of usable PAS and work related to classifying users of PAS then some of the limitations of previous work. Finally the chapter concludes with a brief history of personas and their added value this thesis.

In chapter 3 the study and analysis methods used in this thesis to understand PAS practices are described. The process of clustering information (Q-clustering) from the participant interviews as well as clustering the participants (P-clustering) themselves are explained in this chapter.

Next, the data from the interviews is reviewed and clusters of participants are formed. Chapters 4 and 5 explain two theories for categorization that arose from the data based on the affinity diagram. Chapter 4 looks at a categorization of users based on self-assessments, an early theory developed from the data. Chapter 5 describes a theory based on a secondary analysis of the data at a broader and more encompassing level. The second theory produced five categorizations of users.

In chapter 6, as a way to utilize the categorizations of users of PAS technologies identified in chapter 5, personas are created and introduced. The five categories of users are represented by five personas consisting of a name, a personal quote and brief profile description. In chapter 7 applications of the personas are discussed, including an analysis of an independent set of participants in comparison to the personas, how the personas can be used for analysis of existing design, how they can be used to explain complex concepts found in the data and the potential power of personas for future PAS tool design.

Finally chapter 8 concludes the thesis by summarizing the contributions. The limitations of this research and future directions for further work are also described.

Chapter 2

Background

> *"Security and usability elements can't be sprinkled on a product like magic pixie dust. We must incorporate both goals throughout the design process." [47]*

2.1 Foundations of Usable Privacy and Security

It is important to understand the drive and background behind usable PAS. Usable privacy and security research studies mechanisms to protect personal information and mechanisms to secure computer systems. While privacy and security are separate concerns, there is significant overlap in research in each area.

Concepts behind usable security began as early as 1996, when Zurko and Simon used the term *user-centered security* to refer to "security models, mechanisms, systems, and software that have usability as a primary motivation or goal." [48] Later, in 1999, Witten and Tygar defined the following set of properties for usability of security software: [46]

1. Those expected to use it are reliably made aware of the security tasks they need to perform;
2. Those expected to use it are able to figure out how to successfully perform those tasks;
3. Those expected to use it don't make dangerous errors;
4. Those expected to use it are sufficiently comfortable with the interface to continue using it.

To make security tools usable two major hurtles are often highlighted. First, many designers of security tools take a 'need to know' approach to the technology. As a result, many users do not understand and are not educated on the issues surrounding their own security practices. For example, many users are unsure why they should secure their wireless network. Unable to explain the possible attacks, some users will opt to leave their network open since it is the easiest option. Many argue that training and educating the user would increase his security. However, in their classic work, “Users are not the enemy”, Adams and Sasse examined users’ behaviours and perceptions relating to authentication systems. While education is a problem, the root they argue is a significant disconnect between workplace security departments and the users they work with.

> “Security departments typecast users as inherently insecure: at best, they are a security risk that needs to be controlled and managed, at worst, they are the enemy within. Users, on the other hand, perceive many security mechanisms as laborious and unnecessary – an overhead that gets in the way of their real work.” [3]

Not trusting the user, because of a ‘need to know’ mentality or because users are perceived as uneducated or untrustworthy hurts security. Adams and Sasse pushed for a better understanding of users and their knowledge and motivations. Essentially, Adams and Sasse argue that the need for education goes both ways. Users need better information about what they should do. At the same time, designers need a better understanding of user’s motivations and goals.

The second major hurtle for usable security is that actions to remain secure and private are often passive, ambiguous and a secondary task in the situation. For example when signing up with a new service, the primary goal is to create an account to use the service. The user’s overlooked secondary goal is protecting privacy. This highlights how motivation can impact the user’s decisions. Some designers believe that users do not care enough to protect themselves. However, in another classic paper “Why Johnny can’t encrypt: A usability evaluation of PGP 5.0”, Whitten and Tygar [46] conducted a study to determine the usability of the PGP interface. At the time many researchers believed that standard usability testing was the best way to improve security interfaces. The study showed instead that usability testing for security is radically different than usability testing on traditional products. “Standard usability evaluation methods, simplistically applied, may treat security functions as if they were primary rather than secondary goals for the user, leading to faulty conclusions.” [46] This identification of security (and, by extension, privacy) as a secondary goal has been key in highlighting the motivations of users and shaping the way testing is done.

Since 1999, more researchers have noticed the design opportunities for usable PAS. For example, most motivation behind graphical password research is to reduce the mental load for users and increase users’ ability to create unique passwords i.e. to increase the security of passwords thus protecting the privacy of information stored in the account [20]. Some have focused on creating technologies that enhance users ability to manage passwords in a user friendly manner, such as password storing services like LastPass. Others have focused on changing current technologies to protect the user in manageable way, for example by improving the design of invalid security certificate notification within browsers [45].

One result of the increased awareness of usable PAS has been an increase in research activity. In 2005 the Symposium on Usable Privacy and Security (SOUPS) was created as an annual forum for research. As well, numerous books have been written which either focus on usable PAS (e.g. "Security and Usability - Designing Secure Systems that People Can Use" [14]) or that highlight PAS flaws as one example of failed design (e.g. Why Software Sucks - and what you can do about it [36]).

2.1.1 Privacy vs. Security

One characteristic of many research forums for PAS is that they study both privacy and security in a single venue. SOUPS is a good example of this. As research topics privacy and security are so tightly coupled that the user community is the same.

Because a goal of this research is to view and understand the breadth of the user community, privacy and security are treated uniformly and studied together. A few participants did focus the majority of the interview on either privacy or security, as further explained in chapter 7. However only the participants who acted with an expert level of knowledge and advanced motivation were clearly differentiating between privacy and security. The remaining participants discuss their concerns by freely blended the concepts. This is not surprising since it is difficult to isolate one topic from the other since functions of security are designed to protect private data of one type or another. Research that looks at only security, ignores *what* is being protected. Similarly, research that focuses on only privacy will ignore *how* user are protecting it. Both how and what are important to understanding the user's opinions and behaviours and critical to future tool design.

2.2 Identifying and Understanding Users

This section examines the research focused specifically on developing an understanding of users. User studies in PAS typically make use of qualitative methods or surveys and questionnaires to develop an understanding of users.

2.2.1 Qualitative Studies

Several researchers have made use of qualitative methods to analyze the PAS concerns of users, and how users act on those concerns. Dourish et al. aimed to determine how users manage security on daily basis [18]. Their research allowed them to create a set of practices that characterize a 'typical user'.

Realizing there are differences in users' security behaviour, three studies, Friedman et al. [22] Sheehan [41] and Dourish and Anderson [17] explored factors that might affect how users perceive PAS. Friedman et al. [22] looked at 72 individuals from three different communities in hopes of seeing some differences in security practices based on location demographics. However, they observed that the high-technology participants did not always have more accurate or sophisticated conceptions of Web security than did their rural and suburban counterparts, and conclude that differences are not a result of demographic factors. In contrast Sheehan [41] conducted an email survey and found that there were demographic factors that affected privacy. For example, Sheehan notes that older users are typically either unconcerned or extremely concerned, and younger users (under 45) are more likely to be pragmatists. Dourish and Anderson [17] reviewed the literature to examine ways that social and cultural behaviours influence PAS rationale. Dourish and Anderson establish three models of current and common views of the interaction between people and privacy. The models include PAS as economic rationality, as practical action, and as discursive practice. The models demonstrate that PAS cannot be studied from a single aspect but must be viewed as a tapestry of many different concerns.

In looking at these three particular studies it is difficult to determine what factors, demographic or cultural or context, could influence users habits the most. I am reluctant to believe demographic stereotypes are the ideal way of identifying user categories in that there are too many possibilities to potentially cluster people and that PAS is a broad issue that effects everyone. Culture and context are complex issues that require closer qualitative examination i.e. the way that North Americans view privacy may be very different than people living in China. The goal of this research is to study and use a comprehensive view of the participants to inform the clustering of participants or users and discover design implications.

2.2.2 Surveys and Questionnaires

Beyond qualitative analyses of users behaviour, researchers have also explored the foci of users with respect to PAS using survey data. The following studies, while adopting similar approaches, frequently arrive at different conclusions. Some believe that PAS is unimportant to users, while find users to be concerned but unable to effectively protect themselves.

Among researchers who observe little motivation for PAS, Acquisti and Gross [2] and Conti and Sobiesk [11] used survey data to assess how comfortable participants were with information disclosure, i.e. a loss of privacy. Acquisti and Gross examined students' privacy awareness levels with respect to Facebook use to determine if intentions matched actions. They found that the majority of their sample was more concerned about controlling information than censoring information. They also showed that user goals vary in PAS protection. Conti and Sobiesk surveyed 352 university students and 25 older adults to determine privacy expectations, and found that most were not concerned about a loss of privacy because participants assumed they had nothing to hide.

In contrast to these studies, Berendt et al. [6] found that users they surveyed are much more concerned about loss of privacy. However, they also find that participants actions on an instrumented e-commerce site do not match their privacy concerns: Participants forgot privacy concerns and released a large amount of personal information to the instrumented site. These papers begin to uncover the culture and context issues surrounding privacy but have limited their scope to concerns and not protection or security methods. These studies ignore the user's justifications which continues to highlight the conflict between concern and common behaviour.

Password practices are a common target of survey-based research. Florencio and Herley [21] used survey data to analyze users password practices by embedding a monitoring application in Windows Live Toolbar. They found that passwords are frequently forgotten, are shared across sites, and are generally of poor quality. Adams and Sasse [3] conducted a mixed methods study on password practices. They noted that 50% of participants in their surveys wrote down passwords, but that the need to write down passwords was a result of having multiple passwords and a system-enforced policy requiring passwords to change regularly. They also noted a lack of knowledge about secure passwords on the part of many of their participants.

One characteristic of some password studies is an unstated assumption that violating PAS designers recommendations is a bad decision. Recently Herley argued that security researcher are unfairly judging users as lazy and weak. [27] By looking at exhaustive lists of recommended behaviours for authentication, among other topics, against the chances of the advice actually protecting the user, Herley believes that users are making rational decisions. The argument is that users are able to effectively evaluate the cost benefit ratio. Herley further argues that it is the responsibility of the security practitioners to reevaluate the recommendations and clarity of security tools.

Many researchers have also used media surveys to strengthen a security argument and demonstrate the need for a new technology. As one example of this, when promoting graphical passwords Dunphy et al. [19] noted that 2/3 of respondents in a media survey wrote down a work related password. Similarly, reports of a recent security breach of 32 million passwords at RockYou.com showed that half of all passwords were weak; the most common password was 123456 [30]. While this may be effective in strengthening the academic argument it casts the user further into enemy status by failing to capture the entire picture.

Reconsider these scenarios, at work, in many situations, it may make sense to write down a password. Perhaps the building is secure and all co-workers are trustworthy; Passwords are designed for outside connectivity. Perhaps someone is leaving on vacation and leaves the password with a supervisor. In both of these scenarios, writing down a password is reasonable. RockYou.com, in contrast is a social networking application grafted onto the Facebook social networking site. RockYou.com has been known to have lax security, and contains little personal information. A user may be using an insecure password simply because he knows RockYou.com does little to protect security and because he is comfortable sharing his information with other RockYou.com users. Just because a user's RockYou.com password is weak does not imply that all of a user's passwords are weak.

2.2.3 Identifying Differences Between Users

The goal of this thesis is to develop an understanding of the different perspectives on privacy and security. A basic question this thesis explores is whether there is one prototypical security and privacy *persona*, or whether different people have different perspectives on privacy and security. Furthermore the thesis explores what the different perspectives on security and privacy are.

Both media and academic surveys have been used to identify differences across users, particularly with respect to privacy concerns. The most famous of these media surveys are a set of over 30 surveys conducted by Dr. Alan Westin (see Kumaraguru and Cranor [32] for an overview of Dr. Westin's surveys). Of particular note in Westin's surveys is a classification of users into three groups based on their privacy concerns: the minimalists or unconcerned, the pragmatists or moderates, and the fundamentalists or highly concerned.

Westin's classifications of users have frequently been used to define expected user groups. For example, Ackerman et al. [1] developed a survey to evaluate participants using Westin's scale and found that 27% of respondents were marginally concerned, 56% were pragmatists, and 17% were fundamentalists. Sheehan [41] refined Westin's user groups by conducting an email based survey that asked respondents to rate on a Likert scale their concerns about a set of specified situations (e.g. an awareness of monitoring, how others will use information you provide, etc.) and their frequency of engaging in a set of seven behaviours (reading spam, registering with a website, etc.). Using these Likert ratings, Sheehan identified a total level of concern, and arbitrarily divided the pragmatic majority into two categories. Other researchers have performed small manipulations to the categorizations [6, 10] or have found varying percentages. These follow-up studies represent only minor manipulations to the overall categories introduced by Westin.

2.3 Limitations of Past Research

As stated initially, the goal of this research is to understand the heterogeneous nature of users to enable better design of PAS tools. While past qualitative research has been valuable in identifying common themes amongst a large group of users, it has not been used to identify specific groups of users. Survey data has been a much more common tool for classifying users, a result, no doubt, of the influential survey-based research of Westin.

For this research, rather than using survey data, qualitative methods are used to analyze the diversity of users. To understand the motivations for using qualitative interview data, it is important to understand the shortcomings of survey data.

While survey studies are valuable in identifying what people do, lacking motivations for users' actions, it becomes difficult to interpret the results of surveys. To fully understand this point, consider some basic analysis of past survey data in light of participants' interview data from this research. Consider, first, the observation from surveys that many users write down passwords. From observations collected during the research towards this thesis, it was found that many participants, including the most security conscious, reported writing down a password or even sharing a password. However, there was often specific reason for needing to do so. For example, one participant, P1, notes that he shares a password, but only when there is no other option.

> [P1] At work if I need someone to take over for me when I'm gone or sick I'll write my password down, and in that case I don't really care that much ... I'm reluctant to give it unless there is no other option.

P21 also reports sharing a password, but only with specific, trusted people in her personal life:

> [P21] In one very emergent case I wanted my [partner] to check my email and then to get the attachment file and send it back to me at work.

In a similar vein, the data on passwords used by users of the social networking application provider RockYou.com are highly suspect. Many of the participants cite using different levels of passwords based on the type of site and type of information being exchanged.

> [P29] I have two, one is more in depth and that's for my email and I guess Facebook and for things like twitter and music sites like that I have a consistent password and then banking has its own password... there's different tiers of security needed so they each have different passwords.

Even among participants who could be classified as fundamentalists, a throwaway password is very common and is often as easy to type as 123456.

> [P2] then there is just things like, you want to comment on the [local newspaper] article and it makes you make an account and you know that no one will ever be trying to hack into your comments account. So you just have a throwaway password that is the least secure. Um you don't bother using capitals and numbers and those sorts of things. And it's easier to remember because you've used it so many times.

Essentially, while surveys may provide a snapshot of what people do, I argue that they are difficult to apply to design because they provide little information about why people act as they do.

Another challenge with past work in user segmentation is that researchers often focus narrowly on privacy concerns. However, it became apparent early in the analysis that, for most of the recruited participants, notions of PAS were linked into an overall view of their online strategies. Knowledgeable participants in the study were aware that technologies like Secure Socket Layer connections both secured communication (in e-Commerce) and protected privacy (by encrypting email); participants mixed these concepts freely when discussing behaviours.

Finally, past classifications of users have been shown by recent research to be of questionable value in design. Consolvo et al. [10] examined what type of location information users were willing to release, and what factors influenced that. One of the factors explored was Westin's model of classification, i.e. marginally concerned, pragmatists, and fundamentalists. Consolvo et al. found that Westins model was a poor predictor of users' behaviours. Fundamentalists released location information more freely than the marginally concerned in their study.

The last concern, that was outlined by Consolvo, is of particular concern due to the nature of past classifications. Both Ackerman and Sheehann assumed that Westin's classifications had some merit, and began their research using similar premises, i.e. that privacy concerns discriminate between users. While it is intuitively compelling to assume that privacy concerns could best explain common behaviours, Consolvo's work casts doubt on this assumption.

As a result of the failures of survey data, of the linked nature of PAS tools, and of observed shortcomings of past users models in the design process, the goal of this research is to analyze users from the perspective of PAS concerns, seeking diverse classes of users. The past analysis of classes of users via survey data might, in fact, be an accurate portrayal of users, and this research continues aware of this risk. I believed, at the outset of this study, that a validation past privacy survey data categorizations would be, in itself, a valuable contribution to research in the design of usable PAS tools: As Greenberg and Buxton note, all research results are interim hypotheses, based on a set of collected data. Cross-validating and replicating results are essential components of the scientific process [24]. However, because of the shortcomings of survey data (its closed nature, its single-factor focus on privacy), I was also aware that an analysis of user motivations might reveal a more nuanced picture of the user community.

To accomplish this revalidation of past classifications of users, an open-coding approach is used to analyze data. With open-coding, information (e.g. user quotations or observations by the researcher) is analyzed or *coded* to determine its meaning, information with similar meaning is Q-clustered together and then labeled with a descriptor. Once analysis of the data was complete, I then crafted clusters of participants and used those P-clusters to create a series of personas for design. In chapter 3, an overview of the data analysis approach is presented. This chapter concludes with a description of personas.

2.4 History of Personas

With any qualitative study of users the inevitable question is how to use the data in a significant, meaningful and manageable way. In design, one technique used to ground discussions of unique user attributes, to explore differences between users, and to highlight the fact that certain classes of users are more central to any design endeavour than other users is personas [22]. Personas are a technique drawn from design which uses broad categories of users to create snapshots, or anecdotal sketches of prototypical users. Personas are a tool used to help focus teams, and drive discussions about users. They often work well in early phases of design, but can also be used in critical analysis of products [12]. In this research, personas are created to encapsulate the traits of the discovered PAS user categories and to aid in the application of these categories to PAS design.

Personas are defined as fictional characters created to represent different users within a design process [12]. Personas were first described and used by Cooper, in his book 'The Inmates are Running the Aslyum', where he wrote:

> "Personas are not real people, but they represent them throughout the design process. They are hypothetical archetypes of actual users. Although they are

> imaginary, they are defined with significant rigor and precision. Actually, we don't so much 'make up' our personas as discover them as a byproduct of the investigation process. We do, however, make up their names and personal details"[12].

Many foundational texts regarding personas have been published, notably the works of Cooper [12], Pruitt and Adlin [37] and Mulder and Yaar [35]. Each of these books offers techniques for the creation and usage of personas in the development lifecycle. The books highlight data collection and organization methods to allow the personas to be discovered from either qualitative data or quantitative data or both. Specifically, Mulder and Yaar discuss aspects of the three approaches to data collection and which type to choose based on what data is available and who the target audience is. Mulder and Yaar advocate the creation of qualitative based personas as an initial low cost method for discovering goals or previously unknown issues about a broad community. On the other extreme, quantitative personas require a significant amount of quantitative user data, such as user survey or website log analysis, to be able to isolate and clusters of users within the community and as such narrow the focus. As previously mentioned, the limitations of quantitative studies on PAS habits has, in part, motivated this work to pursue an open ended conversation with participants. Techniques of all three books were utilized in the process.

Personas have been of interests to the human computer interaction (HCI) and requirements engineering (RE) communities for some time now, and debate has existed on the benefits and utility of personas during design [31, 8, 9, 26, 16, 33, 38, 5]. The arguments against using personas cite flawed data collection, or limited adaptability or lack of scientific rigour [9, 16]. De Voil through a critical literature review, considers personas harmful to design by insisting they encourage false belief and groupthink [16]. Chapman et al. argue that, quantitatively, the data presented in a single persona, is not representative of actual users [9]. By dividing a persona profile into its various attributes (e.g. age, location, interests) and systematically searching for those attributes among a user dataset, the authors find that any more than seven attributes and makes it difficult to find matching actual users. Chapman et al. claim "Personas need empirical evidence to substantiate claims that they present factual information about groups of people" [9].

However, recent research has shown some benefit to the use of personas [33, 8]. Specifically, Long found that, in a controlled experiment within a university classroom, students using personas produced more usable designs than a control group [33]. As well, Long argued that the personas encouraged more user-focus in design team communication. Similarly, Chang et al. observed how personas were used in the design process by comparing the results of two design teams [8]. Their study indicated "that a persona supports a design team to come to a consensus when it comes to user images" [8]. In this research, the aim is to construct personas to describe various users within the PAS domain not to specifically debate the benefits of personas.

While debate of the direct benefits continues, the research community has explored how personas can and are being applied design [26, 37, 38, 31, 5]. Pruitt and collaborators have analyzed personas from many different perspectives and highlight many applications in their work [26, 37, 38]. Interestingly Pruitt et al. note that people (developers) are very good at predicting actions based on persona descriptions and harness this in their applications. Aquino et al. explored various methods of user modelling highlighting extreme characters and personas as two techniques to effectively describe the user's model [31]. Beyond developing a technique for identifying personas, Aoyama evaluated the use of personas with scenarios to discover and analyze requirements [5]. The personas developed as a result of the qualitative interviews build upon this application work in Chapter 7 by using predictive techniques of Pruitt, modelling techniques of Aquino and scenario analysis of Aoyama.

In the next chapters, the study design and the data analysis are described. The study follows qualitative methods practices, interviewing subjects about their concrete security behaviours, Q-clustering data around affinities, and looking for patterns within the affinities. The initial theory of data does, in fact, validate the results of survey data, showing how participants self-assessments can define the three common categories of users: minimalists, pragmatists, and fundamentalists. However, when looking more deeply at the interview data the categories appear to be more complex and multi-faceted than past researchers [1, 43] would have us believe. By reanalyzing the data in the affinity diagram, another theory using a different organization of concepts emerges, ones that better captures the different categories of users in the study. The discovered categories provide the basis of traits for the resulting PAS personas.

Chapter 3

Study

3.1 Overview

It is easy for a reader to get lost in the chronology and concepts described in this thesis. Therefore, this section gives an overview that establishes (1) the chronology of three categorizations of users and (2) the vocabulary for the concepts used to carry out two categorizations of users, which are the subject of this thesis. This overview allows the reader to read material through which there is no strictly sequential path.

Past work and the work reported in this thesis identified *categories* of users based on the strength of their PAS concerns and the strength of the actions they took. All such work produces a set of categories that is intended to partition the set of all users. One example set of categories, from past work is: (1) the minimalists, (2) the maximalists, and (3) the pragmatists.

To begin research to *categorize* users with respect to their PAS concerns, the interviews of 32 *participants* were transcribed and *quotations* were highlighted representing participants' views on PAS. An *affinity diagram* was built of the participants' quotations. To make the diagram, each quotation was coded for its meaning. All quotations of the same meaning were *Q-clustered*. When the *Q-clusters* were sufficiently large, each were labeled with a *concept* that abstracted its meaning.

In the end, there were 24 concepts on the affinity diagram. The analysis showed that the division into concepts was not fine enough, that there were indeed *subconcepts* into which many concepts could be divided. There were 85 subconcepts contained within the 24 concepts. For example, the "Corporate surveillance of their employees" concept, with which we examined participants' attitudes towards corporations' monitoring their employees workstations, has two subconcepts: "Corporate surveillance of their employees is OK" and "Corporate surveillance of their employees is not OK".

The concepts and subconcepts on the affinity diagram were used to create two different theories for categorizations of the participants.

For the first categorization, one concept in the affinity diagram was significantly large: "self-assessment of a participant's PAS behaviours". Each participant had at least one quotation in this self-assessment concept. Each participant's self-assessment quotation was examined to determine the importance of PAS to him or her along a continuum from low concern to high concern. The first categorization was the partitioning of the participants into three categories based on the participant's assessment.

Flaws in the partitioning of the participants into these three specific categories became apparent. To build a new, improved categorization avoiding these flaws, the same quotations had to be grouped differently. Regrouped participants were based on the participants' similarity. Two participants' having said quotations in the same Q-cluster means that the participants were part of the concept or subconcept that labels the Q-cluster. Therefore, when two participants had quotations with the same meaning, the participants became *linked* together. A *P-cluster* is a set of participants that share many links with each other and only a few with participants not in the P-cluster. These P-clusters ended up being the second categorization of the participants. Just as participants shared quotations, participants shared also subconcepts and concepts that these quotations were about.

The set of quotations, subconcepts, and concepts that a participant said or talked about are the *traits* of the participant. The most frequently appearing traits in the bag of traits of the participants of a cluster are the traits that characterize the P-cluster and that characterize the user category that *names* or *identifies* the P-cluster.

While the 85 subconcepts on the affinity diagram are useful in defining the set of traits shared by a category of participants, there is a benefit to having more general attributes. Examining again the concepts and subconcepts in the affinity diagram two *dimensions* were noticed running through participants and their traits: (1) *knowledge* and (2) *motivation.* These dimensions capture properties of any users's approach to PAS concerns, and the strength by which the user has any such property is reflected in the strength of his or her PAS concerns and the strength of the actions he or she takes. Thus, a participant or trait can have any dimension in an amount or strength that differs from the amount that another participant or trait has it. A 2-point scale of knowledge and a 3-point scale of motivation, (1) high, (2) medium, and (3) low, *grades*, was used to measure the amount or strength by which any participant or trait has any dimension.

Grading is the process by which a set of participants or a set of traits is assigned a grade for any dimension based on the grades for the dimension of the elements of the set. Specifically, the grade of a set of participants or traits for a dimension is some variation of a weighted arithmetic average of the grades for the dimension of the elements in the set.

3.1.1 Qualitative Research Methods

A unique goal at the outset of this research was to apply qualitative methods to better understand the *differences* between the users in the PAS user community. In finding the best application of qualitative methods it is important to acknowledge that qualitative inquiry is not a strict set of methods for each researcher to follow precisely. In fact, even researchers have difficultly agreeing on classifications of qualitative methods and the number can vary anywhere from four to 28 categories. Creswell wrote:

> The definitions of qualitative research vary, but I see it as an approach to inquiry that begins with assumptions, worldviews, possibly a theoretical lens, and the study of research problems exploring the meaning individuals or groups ascribe to a social or human problem. Researchers collect data in natural settings with sensitivities to the people under study, and they analyze their data inductively to establish patterns or themes. The final report provides for the voices of participants, a reflexivity of the researchers, a complex description and interpretation of the problem, and a study that adds to the literature or provides a call to action [15].

I adopt Creswell's view of qualitative inquiry which describes five categories of approaches [15]:

Narrative Research Studying in depth one or two individuals by gathering data through their stories and reporting the meanings.

Phenomenology Studying multiple individual's lived experiences surrounding a phenomena or concept and reporting the meanings.

Grounded Theory Studying views of multiple individuals to generate an explanation or theory that has been shaped from, or 'grounded' in, data from those views.

Ethnography Studying an entire cultural group to describe and interpret, shared and learned patterns of values or behaviours through immersed observation.

Case Studies Studying an issue explored in one or more cases within a setting or context.

Each stand alone approach is described in detail however it is noted numerous times that researchers may use more than one or integrate multiple aspects depending on the nature of the study.

The research methodology of this thesis was inspired by the grounded theory approach. By using interpreting the data using open and axial coding to formulate theories. However, the participant clustering is a unique use of data, similar to approaches described in contextual design. Because of the intention to understand differences between users the methodology needed to be adopted in a way that allowed the participants to be clustered and the differences to surface. Borrowing from quantitative methodologies Miaskiewicz et al. developed a method to segment users using latent semantic analysis (LSA) of users' responses [34]. The method, while interesting and effective, requires a rigidly structured survey process to perform the analysis including the mathematical comparisons. The method of this thesis does resemble this LSA method in the clustering of participants and the aggregation of traits but the open ended and informal structure of the data collection and coding allows a nuanced vision of the user community to be discovered. The approach to user categorization of this thesis is unique and allows for a blend of methods.

3.2 Interviews

To initiate a better understanding of users' PAS concerns, qualitative interviews with 32 participants (17 male, mean age 26.3, standard deviation 5.9) were conducted over a four-month period. The participants were recruited from two separate samples. Nineteen of the participants were recruited locally for convenience and the remaining thirteen were recruited online from across North America. A brief description of the participants can be found in Appendix B.

A concern with any accessible group of participants is that a convenience sample might introduce confounds in the data. These confounds might result from recruiting participants on a university campus, or recruiting university-aged participants. Confounds of this nature are a particular concern to this research of understanding the differences within a user community, rather than the common actions or attitudes of that community. To partially control for potential confounds due to a homogeneous population, two different groups of participants were interviewed. Nineteen participants were recruited from the student population at the University of Waterloo. These participants represent a convenience sample which are easily accessible and vary widely in knowledge of PAS. Next, thirteen interviews were conducted with non-students, recruited on-line from across North America. While less convenient than local participants, they serve to counter-balance the geographic and demographic homogeneity of the on-campus participants. In reporting the data, note that P1 – P13 are the on-line, non-student participants, and P14 – P32 are the local, student participants. Data from both local and remote participants was combined and treated uniformly during analysis. Collectively, the gathered interview data provides a rich picture of the prospective user community for PAS tools. In particular, it highlights the heterogeneous nature of the user community with respect to the design of PAS tools.

Interviews lasted approximately one hour for the in-person participants, and approximately 45 minutes for the remote participants. The interviews format was semi-structured, exploring:

Participants' knowledge of PAS concepts
Technical questions to assess the participant's knowledge.
What is a security certificate?

Participants' specific security practices
An exploration of the participant's password schemes or software used for protection or detection.
How have you protected your home computer?

Participants' privacy protection practices
Analysis of their use of social networking sites, or their attitude toward electronic monitoring and eavesdropping.
What information is part of your social network profile?

Participants' concerns about PAS
What concernes participants, how they protect themselves, why they protected the specific information they protected.
When we think about computers and the internet, who are the bad guys?

Participants' assessment of their own vulnerability or invulnerability
The security of the participant's computers and data and/or privacy from his or her point of view.
Is there anyone you know that is less (or more) secure than you are? How so?

Participants' rationales for both practices and personal assessments.
How the participant explained their action or inactions.
How do you decide to trust a website when purchasing from them?

While many questions examined participants' specific actions, some questions focused on scenarios that involved how participants would change behaviour based on concrete hypothetical examples. For example, "If you knew there was a 50% chance, that at any given time, someone was watching everything you did on your computer, would you change your behaviours? Why or why not?"

To analyze the data, interviews were transcribed manually and an initial open coding was performed. By reviewing the full transcripts, over 500 interesting and relevant quotes were extracted and coded gathering detail and breadth of the participants. As a visualization for this data, traditional affinity diagramming techniques were used.

3.3 Analysis Methodology

The data analysis explored a various of aspects of user actions, including:

- Where they learn about privacy and security (media, friends, experience).
- The use of social media and networking sites (information posted, frequency of use).
- Views of other's security habits and the tendency to help others with privacy or security.
- Attitudes toward companies that track your activities, personal or corporate.
- Personal assessment of PAS (how the user thinks and feels about their security solutions).
- Practices with respect to passwords.
- Knowledge of PAS technologies.

3.3.1 Creating and Grading Concepts

Working from the data up by grouping like-minded quotes together patterns began to emerge. As quotes were added to the affinity diagram, small groups of tightly related quotes developed. These small groups make up a total of 85 subconcepts, which were further grouped into 24 concepts. The concepts are as follows in table 3.1

Each concept consists of two or more subconcepts. For example, global concerns is divided into quotes of concern about censoring information (i.e., government intervention and firewalls) and quotes of concern about the tracking of personal information (i.e., the data collection practices of Facebook). After low-level analysis of the quotes using open coding, the concepts and subconcepts were analyzed for relationships to restructure and refine the data.

3.3.2 Clustering Participants

The initial theory for clustering participants developed early from the concept of personal assessment. In describing their behaviours, participants would indicate varying levels of motivation. For example, some participants indicated that they always verified credit card purchases and limited their online shopping specific trusted sites indicating higher motivation for protection. Other participants felt they were protected by obscurity and very

Global concerns - concerns for society	Website settings and trust
General concerns - Nondescript “hackers”	Network settings and trust
Security versus convenience	Determining trust of online sites
Privacy versus social	Personal assessments
Software protection methods	Learning source (media, friends, experience)
Motivations	Social network usage
Motivation indicators	Those that claim to censor information
Knowledge indicators	Reasons for making changes
Password layer schemes	Opinions of fallback authentication
Number of Passwords	Opinions of monitoring
Sharing passwords - frequency and situations	Opinions of others’ PAS habits
Writing down passwords (opinions/habits)	Helping others’ with PAS

Table 3.1: Discovered Concepts

little concern for protecting themselves demonstrating lower motivation. These judgements led to three classifications that echo the previous classification systems:

- The marginally concerned - P7, P9, P12, P16, P28
- The fundamentalist - P14, P31
- The pragmatic majority - P1, P2, P3, P4, P5, P6, P8, P10, P11, P13, P15, P17, P18, P19, P20, P21, P22, P23 P24, P25, P26, P27, P29, P30, P32

The full description of these P-clusters and observations are in Chapter 4.

However, based on the initial coding of participants, shortcomings of a one-dimensional understanding of the user community soon became apparent. To address these shortcomings, participants were analyzed using subconcepts from the affinity diagram. If participants expressed common attitudes on the affinity diagram at the subconcept level, the participants were linked. More connections created stronger links. A P-cluster is a group of a participants that share a lot of links with each other and only a few with participants not in the P-cluster. Five P-clusters were initially formed here, however not all participants easily fell into these groups. In order to determine if the remaining participants belonged to one of these P-clusters or new P-cluster of their own, and to determine the defining characteristics of these P-clusters, further analysis was needed.

As previously discussed, PAS user research often revolves around themes of knowledge or motivation. Both motivation and knowledge dimensions were clearly observed within the data The 24 concepts were classified further into dimensions of knowledge or motivation.

Subconcept Name	Grade	Example Quote
Based on unfounded advice	Knowledge Low	[P5] "Then I met [my partner] and just started doing strings of numbers [as passwords], which wasn't the best thing but [my partner] was like, 'you know what? No one is ever going look at it', and so convinced me that it was fine and honestly I've had no problems.
Based on reasonable advice	Knowledge High	[P23] "I got a message through facebook from a friend. It explained how you can have a much more secure facebook page, and, at that time, I changed the security because I didn't want everything to be visible."
Only after being attacked	Motivation High	[P1] "I stopped blogging after my ex-[partner]'s mother read one of my entries and misinterpreted it. She went off the handle, and I thought it just wasn't worth the hassle."

Table 3.2: Example of knowledge and motivation grading. Concept: Reasons for making changes

In some cases, both knowledge and motivation were indicated by the quote. Based on the quotes, each subconcept was assigned a level of knowledge or motivation. The quotes, i.e., coded behaviours, that revealed aspects of knowledge were graded on a scale of low or high. Quotes that described aspects of motivation were graded on a three-level scale, low, medium and high. For example, in the concept 'Reasons for making changes', three subconcepts were graded based on the quotes within, as seen in Table 3.2.

Returning to the participants quotes and the affinity diagram the participants were assigned knowledge and motivation grades. Using the quotes in the previously graded subconcepts, each participant was given a tally for each possible knowledge or motivation grade (knowledge low, knowledge high, motivation low, motivation med, motivation high). For example, if a participant had three quotes, in different subconcepts graded knowledge-high, she would then have grade of three for knowledge-high. Because not all grades types were represented equally in the affinity diagram, i.e. more subconcepts focused on motivation low than motivation high, the grades were normalized to be out of ten to give a simple way to compare and assess the participant. This process of developing P-clusters is very similar to the cosine matrices used during the latent semantic analysis for user categorization [34]. Most importantly, this process allowed characteristics of all P-clusters to be accurately described within the dimensions of knowledge and motivation.

From this coded analysis of participant similarity, a set of five clusters of participants was identified, and correspond to unique user-types. Combining both the subconcept Q-clustering and the knowledge and motivation grades for the participants of the various P-clusters the characteristics were also identified. Table 3.3 defines the participants in each P-cluster, and their knowledge and motivation scores. Chapter 5 further explains the traits of each of these P-clusters.

Cluster Participants	Knowledge	Motivation
P5, P7, P9, P12, P15, P16	Low	Low
P1, P4, P8, P22, P25, P29	High	High
P10, P14, P17, P18, P19, P21, P23	Low-High	Low-Medium
P2, P6, P20, P26, P27, P30, P32	Low-High	High-Medium
P3, P11, P13, P24, P28, P31	High	Low

Table 3.3: Clusters of participants based on the affinity diagram with knowledge and motivation scores.

Chapter 4

Theory One: Analyzing Views of Self

After examining the participant's PAS opinions and habits, the next step is to characterize the users into distinct P-clusters. Much of the previous work with this goal analyzed survey data around the dimension of motivation. By looking at various levels of enthusiasm or concerns previous researchers were able to develop a set of three or four user categories. An initial goal of this thesis was a replication study, using qualitative data, to develop categories as the initial step for persona creation.

To do this, the concepts and subconcepts of the affinity diagram were analyzed and some common themes were observed. Like previous researchers, one very prevalent theme found across many subconcepts was the issue of motivation. The participants expressed different levels of motivation when reviewing their PAS concerns or habits. Similar to the self reported data of surveys each participant reviewed and justified her motivations and actions from her point of view. The subconcepts were organized based on these levels of motivations and perspectives.

Based on this organization three categories of users emerged: The marginally concerned, the fundamentalist and the pragmatists. The marginally concerned users expressed limited motivations and identified themselves as being technologically insecure but unmotivated to change their ways. On the other extreme the fundamentalist were highly motivated to protect themselves and believe they are doing everything within their power to do so. Finally, the category of pragmatists, those that have some concerns, represents the middle ground between the other two categories. The pragmatists category can be further divided into those felt that they personally are not a target for various reasons and those who felt they did something, the minimal amount, to protect themselves, demonstrating some motivation.

This chapter reviews an initial coding of the interview data. First, an overview of the subconcepts found within the affinity diagram. Next, the subconcepts are organized according to differing self rated motivations levels, resulting in clusters of participants. Finally, an analysis of the effectiveness of this coding is provided.

4.1 Analyzing Views of PAS

As previously described, the interviews covered many aspects of user interaction. The user quotes were divided in to 24 concepts and 85 subconcepts based on similarities in meaning or context. Some concepts focused around common elements from every interview, such as the knowledge indication question or number of passwords, while other concepts flowed from phenomena of the discussion such as the concept of security verses convenience or personal assessments. To gather an idea of the breadth of the affinity diagram the table 4.1 lists all of the subconcepts and their respective concept.

Table 4.1: Concepts and Subconcepts of the Affinity

Concept	Subconcept
Censoring Information	Not Signing up, Censorship
	Not Storing Info or Limiting Info, Control
Determining Trust Online	Security Claims (low trust)
	Look and Feel (non-sketchy)
	What the site says
	Reputation
	Popularity or Size
Fallback Authentication	Negative
	Causes them not to log in.
	Positive
	Neutral
General Concerns	Depends on Point of View
	Criminals/Hackers/ID Theft
Global Concerns	Tracking Info (Google and Facebook)
	Censors of Information, Government
Help or no Help	Helpers
	Non-Helpers
Home Network and Trust	Maintains and trusts
	Little or No Trust and Maintains it
	Trusts but does NOT maintain

continued on next page

continued from previous page

Concept	Subconcept
	Trusts all public networks
Knowledge Indicators	Security Cert. Understand
	Cookies Understand
	Security Cert. Does not understand
	Cookies Does not understand
Learning Source	Security Advanced
	General (Word of Mouth / Trail and Error)
	General News
	TV Shows
Made changes to security habits	Based on Advice - Most founded
	Based on Advice - Least Founded
	Only as Attack Response
Monitoring	Not Okay with Monitoring
	Okay with Monitoring
Motivation Inidicators	Put Things Off
	Small changes to password habits
	Not so lazy, in some specific cases
Motivations	Trusting in Bank Safeguards and Being refunded
	Physical Security Concern and Response
Number of Passwords	Mostly Unique - Level 7
	Level 6
	Level 5
	Level 4
	Level 3
	Level 2
	Least Unique - Level 1
Password Layers Schemes	Most
	Mid-Level
	Least
Personal Assessments	No worries or concerns
	I'm not very secure
	I don't know how / I haven't bothered to look it up
	I'm not a target (companies are)
	I don't Matter (Honest Man)
	You can't find me (Obscurity)

continued on next page

continued from previous page

Concept	Subconcept
	I know it when I see it
	I used to be very insecure
	I take extra care to protect my Bank
	I treat it like its public
	I monitor very closely
Security verses Convenience	Security
	Convenience
Social Network Usage	Low Usage
Social verses Privacy	Privacy
	Social
Software Protection Methods	Distrust of Software
	Advanced Software Manipulation
	Reformatting
	Additional External Habits
	Locking Comp (All Times)
	Anti-Virus and Something
	Only Anti-Virus
	Nothing because of O/S
Website Settings and Trust	Use but unsure (limited Trust)
	Yes (trust)
View of Others	Uneducated/insecure Specific
	Uneducated/insecure Generalization
Writing Down Passwords	Previously or 1-Off Cases
	Written Currently Safest
	Written Currently Most Exposed

4.2 A Unifying Theme

To cluster participants, subconcepts were examined for a unifying theme. One theme that became apparent early in many of the user quotes involved self-assessment of their own PAS practices. During the course of the interviews, many different attributes contributed to participants' self-assessments of security habits. The participants frequently came to a revelation of their level of concern and the amount of motivation that they feel they put towards protecting themselves. Some participants came to these conclusions while comparing themselves to others, or when recounting their security habits and justifications or when trying to describe perceived threats. Some examples include:

> [P14] And my parents its so funny are like 'You do online shopping? Your going

> to get caught with fraud, people are going to steal your identity' and I'm like, I'll worry about it when it happens, I haven't had a problem.

> [P17] Well I sort of consider Facebook as public. I mean anybody could get to you, probably, so whatever I don't want known I don't put there. I don't consider the login terribly secure.

> [P6] I don't know how weird I am but I have no fear when it comes to hackers and all that crazy stuff. There's no reason for them to go after me, the way I see it. I'm not a target.

The way that the participant described themselves or their actions led to various subconcepts of self-assessment. Each subconcept of self-assessment was reviewed on a continuum from self-assessed low or no concern to self-assessed high concern. Table 4.2 shows the movement from low to high concern.

Low/No Concern	No worries or concerns
	I'm not very secure
	I don't know how / I haven't bothered to look it up
	I'm not a target (companies are)
	I don't Matter (Honest Man)
	You can't find me (Obscurity)
	I know it when I see it
	I used to be very insecure
	I take extra care to protect my Bank
	I treat it like its public
High Concern	I monitor very closely

Table 4.2: Listing of Personal Assessments in order of concern levels

In fact, many of the concepts of the affinity diagram can be arranged in a continuum from low to high based on a unifying theme such as motivation. Refer back to table 4.1 where the subconcepts have been arranged in their own continuum for each concept.

Using the personal assessments, based on the quotes, the participants were placed along the continuum and P-clusters emerged. Those that show no concern; those that said they did a significant amount; those that cared, but thought they were not a target for practical reasons (e.g. obscurity); those that did a minimal amount to protect themselves.

4.3 P-Clustering Based Upon Self Rated Motivation

From the participant's self-assessments, three broad categories of users were identified: The marginally concerned, the fundamentalist and pragmatists. First, a set of participants, the marginally concerned, claimed they were insecure or that they had no significant security concerns.

> [P12] I have to say that I often don't keep things that secure, because I don't leave my computers on, I usually turn them off. So when I'm not using them they are off. So I'm secure that way.

> [P16] I know there are precautions to take and to a certain extent I take them but at no point in my day do I worry about them. I have a thought of 'oh I hope this is safe' but at no point after that do I worry about it.

Participants P7, P9, P12, P28, P16 all fell into this group. On the other extreme there exist the fundamentalist, a rare group consisting of only P14 and P31. These participants claim they are safe because they monitor very closely.

> [P31] I think more so I am very meticulous about keeping all of my receipts and checking them to my accounts when the statement comes in. So I really don't think twice about using my credit card cause I know that I keep tabs on it.

Based on self-assessments the remainder of the participants belong as members of the pragmatic majority. Within this group of participants, two sub-categories were identified: those that do little and rationalize their inaction; and those that take some basic steps to protect themselves.

Within the first sub-category, a set of varied rationalizations for inaction existed. These ranged from "I'm not a target" or "I don't matter" to "They can't find me." A group of participants also trusted themselves to recognize unsafe practices, stating "know danger when I see it." While participants expressed these sentiments in various ways, the various feelings of safety were consistently used as a rationalization for inaction.

> [P13] I'm sure [hackers] are, I just don't think they are as much a concern for the average person as they are to a major corporation or something where they actually have more to gain.

> [P5] But I mean I'm giving it too much credit to hackers, but I don't know what anyone would want to hack into my Facebook account. I've got nothing in there.

> [P6] Well because I don't think I'm ranked high enough on a Google search that I don't think I would be found very easily it would have to take one heck of a crawler for somebody not to know that I'm there to stumble upon me.

> [P30] I figure that if it's between me and my [partner], who is a CS major I'm pretty sure [my partner] won't click on the screen saver ad. So we just go to safe sites like Yahoo or Google or whatever, so it should be safe.

Participants P2, P3, P4, P5, P6, P15, P17, P19, P20 and P21 all expressed rationalizations for their passive approach to PAS.

Finally, a second sub-category of participants also self-identified within the pragmatic majority because of a set of minor steps they took to protect themselves. These participants would restrict the information posted to online sites like social networking sites, or limit their financial activities based on context.

> [P32] I try to avoid the Facebook applications cause I don't like the fact that they require you to essentially sign away all of the privacy stuff.

> [P1] Yeah, you know some sites they ask if you want to store your credit card number and I always say no.

> [P25] *Is there anything you don't do on public wifi?* Um I would do everything except online banking because it can't be trusted.

Participants P1, P8, P10, P11, P13, P18, P22, P23, P24, P25, P26, P27, P29 P30 and P32 all took small steps to protect themselves. It should be noted that these participants also rationalized their inaction or limited actions, as did participants in the other sub-category of pragmatists. The difference is that these participants also took specific actions to protect themselves.

These user categories based on motivations and self-assessments echo and help validate the previous research. Three categories were identified with habits similar to the categories based off the Westin survey model. In fact, the distribution of participants into the clusters resembles previous research, as table 4.3 demonstrates.

Table 4.3: Distribution of participants in clusters compared to previous research.

	This study	Sheehan 2002	Consolvo 2005	Ackerman 1999
Marginally Concerned	16%	16%	19%	27%
Fundamentalist	6%	3%	12%	17%
Pragmatic	78%	81%	69%	56%

The similarities are not so surprising as the methods used to identify user clusters are similar. When doing a survey the users are asked to identify and grade their actions and concerns. This projection of themselves was also observed in the interview data as the participant's self assessments. In both cases the participants are identifying their motivation levels or concerns through their habits or justifications. In this way it is easy to identify the users on the extreme ends of the spectrum as unconcerned and highly concerned however the majority of the participants continue to fall into a middle ground.

4.4 An Analysis of Motivation-Based Clustering

In analyzing the relationships between categories, however, the relationship of participants to others within categories became very tenuous. Consider, as a simple example of this, P4, P6, and P30 in contrast to P22. All of these participants belong to the broad category of pragmatists, those that rationalize inaction, and all claim to be safe because they can't be targeted. In the case of P4, P6, and P30, this sense of security is a result of owning an Mac OS X based computer rather than a Windows based computer.

> [P6] Ever since I've had my Mac, which is almost 5 years old, I haven't worried about it. When I was on the PC I was way way more worried about that stuff.

> [P30] I really like my Mac I find I can do whatever I want and not have to worry about viruses so much and I don't have to worry about things coming on my computer and trying to trash around with it.

In contrast, P22 notes that:

> [P22] I mean the type of tracking and the amount of data they can accumulate based on the different searches you do and the websites you visit because they can track referrals if you click on different links in the Google search then okay because I mean I've registered with Google. But I have blocked Google analytics with the noScripts extension on Firefox.

Essentially, while all of these participants viewed themselves as obscure, not a target, the sophistication of the techniques that masked their online vulnerabilities was very profound. P4, P6, and P30 believed that, because Macs were rarely targeted in the past, they were safe, whereas P22 was much closer to a fundamentalist. P28 is another unusual participant. Like P22, P28 used the noScripts option to protect their privacy. P28 also has a relatively strong password scheme, among the strongest of the participants, with nearly unique passwords across most of her accounts. However, beyond setting the noScripts option and creating good password, P28 does little about security. P30, one of the Mac owners, also feels protected by being part of a larger group.

> [P30] So on my Facebook I might put my cell phone number, I put the city that I live in but not my address. I put that I go to Waterloo but since there is 22,000 it doesn't matter. I only put information that could mean that I'm part of a really large group I try not to put anything very specific so you can narrow it down and catch me.

However, P30 is aware of privacy issues and this has an effect on her behaviour. P30's awareness of being watched causes her to censor her online activities.

> [P30] I don't really change my status that often [on Facebook] because I know that people are watching, that people are commenting, I know that some people are watching intently cause I've been called on it.

P30 also goes to extra lengths to keep her home network secure.

> [P30] Yes I have MAC filtering and WEP. I live in an apartment building so it's pretty important for me to have it secure because I've definitely stolen other people's wifi before.

Essentially, while self-assessment gives a sense of a broad middle-ground in PAS practices, the pragmatic majority is still a very complex category, so complicated that it was difficult to assign attributes to the category. A pragmatist is equally likely to have strong and weak passwords. He sometimes helps others with PAS issues and at other times does not. Some accept unfounded advice, while others learn from their experiences and the experiences of others how best to protect themselves. Each has a different view on the acceptability of monitoring. Some learn about security on TV, some by experience, and some from others. Finally, their knowledge of PAS issues – from Google Analytics to cookies to potential vulnerabilities of WPA/2 – varies from no knowledge to expert level knowledge.

Additionally, based on the nature of the study and the format of the interview that allowed concepts of categorizations to flow up from the data, participants were not directly asked to assess themselves. Because of this placing some participants within the groups became difficult. For example P27, has a fragmented self assessment where she doesn't seem to worry as much as she did in the past and becomes apathetic about maintaining her accounts, by not bothering to learn how to delete a friendster profile.

> [P27] *Do you do online shopping?* Yes, not in the past it started, I would say three years ago. Previously most people were worried about the security problems but then um I think that more and more of my friend would go for it, but right now I'm still going for big, brand name stores because they are safer but even I do online banking so I guess the bank is safer, like they have better policy.

> [P27] I don't really use friendster anymore. When it was the very first time when I got into the social network it was friendster but later on facebook comes up and I joined facebook and now I don't really use friendster any more. I would cancel it but I don't know how to cancel it so I just leave it there.

These self-assessments could make P27 a candidate for the minimalist category however, later she goes on to explain her concerns about social network information and about downloading viruses. P27 likes to limit information on Facebook by never listing contact information beyond her city and closely monitoring her privacy settings. To ensure P27's computer is clean she frequently formats the system.

> [P27] I'm the person who used to reformat my laptop every term. Just incase... reformatting for health or sometimes just I have too many files or too many programs or it's too slow and I would reformat it too.

These self-assessments and actions makes P27 a stronger candidate for the large pragmatic majority, but she is clearly a complicated individual unsure of the strengths and weaknesses of her PAS behaviours. P27 was not the only individual that was difficult to place into one of these three categorizations. The qualitative nature of the interviews show how brittle the self-assessment categorization can be and why it has been routinely applied using mainly survey data.

With these limitations in mind, the data is reanalyzed to broaden the views on the participants. To truly examine the differences between people, deeper analysis of the data needed to be performed to find strong connections between like-minded users.

Chapter 5

Theory Two: Sharpening the Distinctions

One problem with a three-category view of participants is that the pragmatic majority has relatively few unique characteristics beyond a propensity to rationalize their actions. In analyzing the various concepts of the affinity diagram, it seemed that a large number of the pragmatists could either be classified as minimally concerned or fundamentalists, and that the pragmatic majority was actually a small minority of participants who took some modest steps toward self-preservation.

Participants varied in their behaviours across each of the 24 previously explained concepts. For example, many participants had admitted to writing down passwords, but motivation for writing down a password indicated aspects of the participant's motivation to protect his PAS. As rationale for writing down passwords, responses ranged from those that had written down passwords previously or in one-off cases:

> [P29] Probably, when I was trying to figure them out for school and because of such stringent things, you just kinda need to see what it looks like.

To those who currently write them down insecurely,

> [P9] Yes. I have the password for my computer written on a sticky note underneath the little tray in my desk because there is another worker that comes in and has access to my office and I put it there so she could have access to my computer if she needs to.

To those who write them down securely,

> [P22] I do store my passwords in an encrypted container in my computer and I have a master password to unlock that if for some reason I needed to access it. As well I have them written down on a piece of paper that I have hidden in my room because I assume that if somebody has access to my room and can actually spend the hour [to] find it I've got much worse problems than that.

To understand the similarities and differences between participants, quotes were thoroughly examined. As explained in section 3.3.2, by analyzing the quote set, these connections between participants produced a set of five categories of tightly linked participants. Graded dimensions were applied to each of these categories based on the participants within, the final categories appear as follows:

- The Marginally Aware - Knowledge Low, Motivation Low
- The Fundamentalist - Knowledge High, Motivation High
- The Struggling Amateur - Knowledge Low-High, Motivation Low-Medium
- The Technician - Knowledge Low-High, Motivation High-Medium
- The Lazy Expert - Knowledge High, Motivation Low

Linked concepts, subconcepts and quotes of the participants within a P-cluster define a shared dominate set of traits. As well, there exist a set of traits that are shared between P-clusters. Table 5.1 summarizes the characteristics of each of these clusters of participants. The next five sections list the dominate and shared traits for each P-cluster, including example quotes.

Table 5.1: Dominant and Shared Traits of Each P-Cluster

P-Cluster	Traits
The Marginally Aware Knowledge Low Motivation Low Participants: P5, P7, P9, P12, P15, P16	- Learning Source: TV shows and Word of Mouth or Friends - When determining what is a safe site, trust what the site says. - Doesn't grasp more basic technical terms (Cookies) - In favour of using fallback authentication questions - Only identified software protection is anti-virus scanner - Small changes from triggers. Changes password 123456 to May0584 - Knows there exists threats but doesn't worry about them. - Often have a small set of passwords but one is heavily favoured. - "I'm not very secure." Recognizes personal insecure habits. - Occasional distrust of software (e.g., Norton) *Shared Traits* - Generalization of Concerns (hackers or id thieves) - Does not know what a security certificate is. - More often and currently has passwords written down insecurely - "I dont know or haven't bothered to look it up" - Believes monitoring or being watched would be okay. - "I don't matter" (Honest man) - Trust in the bank safeguards and believes all money would be refunded no problem.
The Fundamentalist Knowledge High Motivation High Participants: P1, P4, P8, P22, P25, P29	- Non or reluctant helpers - Little or no trust of home network (WPA2 is questionable) - Looks for security claims on websites (https, padlock etc) - Sometimes refuses to sign up or participate - Monitoring or watching is not okay - Maintains global concerns (censorship or tracking) - Multi-layer passwords, important passwords unique - May extend protection beyond computer - Views the general public as uneducated and insecure *Shared Traits* - Advanced security, independent learning - Hate Fallback authentication - Knows about more technical terms, Security certificate

continued on next page

continued from previous page

P-Cluster	Traits
	- Some advanced software manipulation knowledge - Identified software protection as anti-virus scanner and something extra (Firewall or malware scanner) - Chooses security over convenience
The Struggling Amateur Knowledge Low-High Motivation Low-Medium Participants: P10, P14, P17, P18, P19, P21, P23	- Understand basic technical terms (cookies) - Stuck in their set up. - Make changes based on weak or inaccurate advice. - Trust but does not maintain their usual wireless network. - Like to limit the information that is given out. - Has specific or inaccurate views of other people's PAS - Odd cases of sharing passwords - One stronger password or mid-level of layered password schemes. - Passwords are unique and personal "to me" - "I monitor very closely" - Are not lazy in every aspect of PAS - Basic trust of all wireless networks. *Shared Traits* - Generalization of Concerns (hackers or id thieves) - Does not know what a security certificate is. - More often and currently has passwords written down insecurely - "I don't know or haven't bothered to look it up" - "You can't Find me" (Obscurity) - Trust in the bank safeguards and believes all money would be refunded no problem.
The Technician Knowledge Low-High Motivation High-Medium	- Learning source includes news and blogs - Limited trust of privacy settings on website like Facebook - Trust in look and feel of website, looks for non-sketchy websites - "I used to worry about those things" - Made changes based on sound advice - Choose privacy over being social - One off cases of writing down passwords - Passwords all based off one thing - Passive user of social networking - Physical concerns

continued on next page

continued from previous page

P-Cluster	Traits
	- Trust sites based on reputation or rating systems - “I know it when I see it’
Participants: P2, P6, P20, P26, P27, P30, P32	*Shared Traits* - Knows about more technical terms, Security certificate - Some advanced software manipulation knowledge - Identified software protection as anti-virus scanner and something extra (Firewall or malware scanner) - Generalization of Concerns (hackers or id thieves) - Chooses security over convenience - “You can’t Find me” (Obscurity)
The Lazy Expert Knowledge High Motivation Low	- Helping others with PAS - Participants in this group also choose convenience over security - These participants also rationalize a lower level of concern. - They have mostly unique passwords - They treat the web like its public domain - They share passwords, only rarely with trusted people - They typically choose being social over privacy - They don’t believe they are personally a target - They trust and maintain their own home network - Careful with passwords
Participants: P3, P11, P13, P24, P28, P31	*Shared Traits* - Advanced security, independent learning - Hate Fallback authentication - Knows about more technical terms, Security certificate - Some advanced software manipulation knowledge - Identified software protection as anti-virus scanner and something extra (Firewall or malware scanner) - Believes monitoring or being watched would be okay. - “I don’t matter” (Honest man) - Trust in the bank safeguards and believes all money would be refunded no problem.

5.1 Participant Cluster – The Marginally Aware

Contributing Participants: P5 P7 P9 P12 P15 P16

Dominate Traits:

Participants in this cluster identify an anti-virus scanner as their only software protection method.

> [P9] It is whatever came with my computer. I was actually talking to someone about this the other day that I may or may not have a real program on there to protect my computer. I know I get security updates when I check it... it's Windows Defender.

The participant makes small changes after a trigger or being forced. For example, would change password from 123456 to May0584 when the system requires it.

> [P5] Recently because most place have been requiring more and more you need to have at least one number and 8 characters I've been using variations of my birthday, so I'll either use a number for the month or the day and I'll type out either the month or the day or whatever the other one is. I have a couple different variations of it so if one doesn't work I can try the other one. And it's pretty simple.

Knows there exists threats but does not worry about them enough to take action.

> [P16] I know there are precautions to take and to a certain extent I take them but at no point in my day do I worry about them. I have a thought of 'Oh I hope this is safe', but at no point after that do I worry about it.

Limited understanding of basic technical terms, for example, cookies.

> [P12] I think a cookie is information that was used on your computer and it's all saved in a cookie folder and basically if you delete your cookie folder you are cleaning it out, it's almost like cleaning out your RAM I think.

Occasionally questions the effectiveness of anti-virus software.

> [P5] I don't think those norton things are doing anything. I don't think they are helping. I had no luck with them I just get a bug anyways.

Often have a smaller set of passwords but one is heavily favoured.

> [P12] I have three passwords, I have my one that is my ‘go-to’ that I use most of the time and my other two are cause most times on my work site every month we are asked to change it so often.

When determining what is a safe site, trust what the site advertises. For example has won award for most secure site, or ‘we promise not to sell your information.’

> [P12] When you first go into it they show all the awards and everything they have won for security and they could be totally fictional and made up but it still makes me feel better to know they have won awards for that.

The participants recognize that they aren’t very secure.

> [P15] *Why did you say you are careless?* Because I think that if someone has my facebook password or something or my friends on messenger and I don’t lock things, and I think it doesn’t matter and I always put the secret thing in my diary.

The participant’s learning source includes media sources such as TV shows (like CSI) as well as word of mouth or friends.

> [P7] Just because with doing different accounts and those sorts of things they sort of cox you into making the proper password especially within the bank you have to read all this information and they tell you what is a good password, um just because it is so crucial with what we are dealing with, so they just want to make sure that it is secure.

In favour of using fallback authentication questions as backup to log into online accounts.

> [P9] Sometimes I don’t pass them so I feel like they are probably pretty secure. Sometimes a capital letter or something. I feel that they are pretty secure cause it’s stuff like, one of them was what was your first pet’s name like there’s very few people who would know that or even like what was your confirmation name, so like my grade 7 teacher and my parents probably remember that so.

The marginally aware cluster seem to represent the naive or lazy users that are stereotypically depicted as the enemy to PAS. While this group has many areas for improvement, the fact that the participants are able to identify that they are insecure offers hope.

5.2 Participant Cluster – The Fundamentalist

Contributing Participants: P1 P4 P8 P22 P25 P29

Dominate Traits:

These participants maintain global concerns for example concerns regarding censorship or tracking.

> [P22] *What about Google and Facebook?* It kinda freaks me out to say the least, I mean the type of tracking and the amount of data they can accumulate based on the different searches you do and the websites you visit because they can track referrals if you click on different links in the google search they're okay because I mean I've registered with google I have a gmail account so obviously ... and they also know who I associate with and they know who I send emails to and facebook knows my group of friends and roughly where I live and who I talk to on a friendly basis which is kinda unnerving but as I've said before you have to weigh off that compared to being shunned socially.

Have refused to sign up for a new service or participate on social networks on occasion.

> [P29] I've kinda stopped using chrome cause I'm kinda paranoid toward google right now.

This cluster motivated to implement and maintain advanced software protection and may even extend protection beyond the computer system.

> [P1] Every blue moon or so I do a scan of my computer just in case. And yes I'm becoming concerned that Mac OS X is going to be exploited by viruses. Not as safe as it used to be.

> [P22] I check to make sure that SSL is on I mean there are still MITM attacks and stuff like that... I make sure that my checking account has relatively small amounts of cash in at all times and I probably try not to do it over an unsecured connection I would only do it over a connection I trust. Still I mean, there is always a risk that something might happen.

All *important* passwords are unique and as such maintains a mid or low multi-layer password systems.

> [P8] *How many passwords?* 20 maybe *And thats how many online accounts you have?* Yeah.

> [P22] I have a couple different layers of passwords with different levels. Maybe 5 different levels depending on how important the site is. I've got some speciality passwords for sites like my banking or whatever but in general I have one level for email, a higher level for school, a lower level for just forums and a last level for random sites.

> [P4] Some of them share the same password, some of them share other passwords. I generally have a structure where I have three passwords. One of them being a high security password, one of them being a normal security password and one of them being an I don't care password. and depending on how important the site is will depend on what password I'll use.

This cluster is highly skeptical and has little or no trust of their home network even though they maintain it.

> [P22] In some sense trust it I guess... well I mean if you just go to some random network, the probably is extremely low but there is a possibly that someone is spoofing pages for a MITM attack and my home network is over wifi but its got WPA2 so I assume is relatively safe and from that it's connected to some random ISP I think it's Rogers so it's relatively well known I mean, I'm not going a lot, there is no real hard data on wether it's secured or not but I'm assuming it is.

Looks for technical security claims on websites such as https or the padlock symbol before establishing trust.

> [P1] You can only know so much right, and a lot of it is just their claims. So I look for instance, when I log in is it over https? Do they make a claim of what sort of encryption they are using? um and through any of my dealings with this site do they ever email my password in plaintext? If so, not going to get [my information].

Monitoring or surveillance of any kind is not tolerated.

> [P8] I would probably do a lot less on my computer. I would use it less or use it for less important things. I probably wouldn't do my banking online anymore.

Views the general public as uneducated and therefore insecure leading to a reluctantness to help or educate other people with PAS issues.

> [P29] I guess it's all coming together with all the information that people are putting on there and you hear the news stories about these kids who put nonsense and entirely too much information and I feel like theres an entire generation who's voluntarily just giving up their privacy.

> [P25] I just like say my piece, I go on my 5 minute rant about the privacy and they are like 'We don't care' and I'm like alright. Done.

The fundamentalist to some would represent the ideal user, one who cares deeply about PAS. Interestingly, there is some anecdotal evidence to suggest that fundamentalist eventually fatigue and could end up with traits similar to the technician cluster as they find their balance between cost and benefit.

5.3 Participant Cluster – The Struggling Amateur

Contributing Participants: P10 P14 P17 P18 P19 P21 P23

Dominate Traits:

Participants in this cluster like control over the information they put online including who they give information to and when.

> [P10] It would almost always be just my name and one of two emails I never put my phone number up there unless its someone that I'm communicating with for a job, it's strictly just email.

They are motivated in some aspects and can devote an unexpected amount of time to simple security methods. For example these participants could spend hours trying to set up a wireless network.

> [P14] *So do you change your email password every three months too? or just work?* Um I would say I change my email also, not as often, I would probably have to change it every six months, just in case,

Has some varied technical knowledge in order to understand basic technical terms like cookies but not processes such as SSL.

> [P10] I know that a cookie are part of a webpage that keeps minor information about what kind of information is on a website. I know that I have it set up to dump the cookies after I close my web browser.

Typically has one stronger password or mid level of layered password schemes, usually done by categories in which case passwords are personal or unique “to me”.

> [P19] Some of them I have the same password. So similar types of accounts would have, so if I have more that one email I would probably use like the same type of password or a combination or a little variation on it but it would kinda have the same underlying theme I guess. Then if it was a bank account it would have the same kind of theme but slight variations but some of them are exactly the same.

> [P10] *Are your passwords unique?* Usually no, they are unique to me but I think I have about a half dozen different passwords that I use that are all fairly similar but they are something if you don’t know me personally you probably aren’t going to know. But if anybody really wanted to sit down and crack my information the could probably easily do it.

This cluster of participants has odd cases of sharing passwords, including for work or with random people.

> [P14] I’m not sure if it’s for the telephone or the computer but work said let somebody know what your password is, incase something happens. So let’s, knock on wood, say I got hit by a car and killed, they would be able to access my files that way and they would be able to continue my work. It’s kinda morbid to hear it like that but I mean it makes sense, from a business perspective I mean.

Has a basic trust of all wireless network and does not maintain their usual wireless network. Typically, uses the school’s network or their landlord’s network.

> [P18] Uh I don’t know, either. It’s just all the things are set by my landlord I just clicked it and I use it and I just typed in the password.

> [P14] Because for example if I was buying concert tickets using wifi I would still have to put in all my passwords and everything I don’t think it’s any less secure... I don’t think so.

Believes they are secure because of diligent monitoring of their finances on a monthly basis.

> [P14] It’s important to monitor but I haven’t had a problem I mean I don’t think anyone has stolen my identity because I’ve checked my purchases and, yep, those are mine.

Make changes to their PAS protection methods based on weak or inaccurate advice from others.

> [P14] I remember my boyfriend telling me, ‘Oh, take your birthdate off of facebook and don’t put where you live’ and I said why? cause people like sending you birthday wishes right? ‘Oh cause i heard ... they can use your birthdate and where you live to get your credit card information. I don’t know how true that is but I was a little concerned and I thought why not be proactive.

Because of limited technical knowledge or motivation these participants are seemingly stuck in their set up. For example, encountering fallback authentication question has cause them not to log in.

> [P21] There is also a security control with my bank they will ask a security question if I’m using another computer instead of the one I registered so I am so lazy to remember all those questions so I try to avoid all those questions and to keep using my own laptop on my home computer.

Finally, this cluster maintains specific and sometimes inaccurate view that others are uneducated or insecure

> [P14] And my parents its so funny are like ‘You do online shopping? Your going to get caught with fraud, people are going to steal your identity’ and I’m like, I’ll worry about it when it happens, I haven’t had a problem.

The struggling amateur can be seen as part of the previous pragmatic majority. This group has some motivation, unlike the marginally aware, but typically is limited by knowledge.

5.4 Participant Cluster – The Technician

Contributing Participants: P2 P6 P20 P26 P27 P30 P32

Dominate Traits:

Chooses to preserve privacy over being social online.

> [P26] No not like personal information because right now there was some kind of auctions or some kind of privacy messages that don’t put some information because there are some people who are stealing the information and then they can use it for the credit cards or just stealing your information

Even when using social networks these participants are passive in their participation.

> [P6] I would say I don't really use them consistently. Like Facebook I see what people are saying and then I hop right off but yeah I don't use the whole functionality of any of them. I guess in some geek terms I would be considered like a lurker. I lurk, I stay back and I watch and I don't get into it too much.

This cluster has more concern about physical risks and threats.

> [P6] It breaks down into my whole thought on security, it's the people you talk to the people who have actual access your equipment, they are the ones who actually have the most chances of tearing you down.

These participants create passwords based off one personal seed.

> [P30] How it works, I have a word and then everything is based off that word and it has a certain number that I like at the end of it, if it needs a number. And then if it wants capitals I take the first letter of that word and I capitalize it.

Very specific and special cases for writing down passwords.

> [P26] Before I would write down just to remember but now there are really the ones I know and if there will be some new ones I would change.

Limited trust of privacy settings on website like Facebook, chooses to censor information regardless of the website settings.

> [P27] I think so because there is the option for you to pick the friends of friends who might be able to see it I think I set it to friends.

When determining levels of trust with websites or service the 'look and feel' of the site or the reputation or rating systems are most informative.

> [P32] Most of the time I'll either go on to the publishers site and order from there or I'll go to something like Amazon I try to stay away from sites that have been poorly designed web background cause that to me is kinda sketchy. *what is sketchy?* I dunno the ones that have kinda typos or grammatical errors or rocky fonts or black backgrounds with white fonts with pink and things that dont seem authentic.

> [P30] I try to use the rating system, I've heard that you can screw around and try to make it whatever but like, I sort of trust it despite, sometimes I look at the reviews depending on the item sometimes, like if they have some really rare item that it's like, why would they lie?

This cluster has a sense of knowledge when it comes to PAS claiming not only that they used to worry about those things, but now they know danger when they see it.

> [P2] I would like to think that most of the time I don't usually go to sites that would, um make it a problem, or download weird things. There was that one time I downloaded a screensaver and my computer immediately told me it was a virus and I felt stupid.

> [P6] Ever since I've had my Mac, which is almost 5 years old, I haven't worried about it. When I was on the PC I was way way more worried about that stuff.

Primary learning source includes news and blogs pertaining to PAS.

> [P20] Some of it is professionally, sort of word of mouth type thing. I'm sorta plugged in now but I'm not a pro... but im plugged in now so its easy now and I try to keep on top of things

Has made changes based on sound advice from friends or more knowledgeable people.

> [P27] Somebody told me it's dangerous. If you need to go for something like, go into your bank account so probably these details I wait until I go back home like unless it's really urgent but mostly I would just go web browsing and in my email. Because I don't have a lot of personal information in the email so I hope it would be okay.

The technician cluster represents another segment of the pragmatic majority, of higher motivation and mix of low to high knowledge a smaller middle ground that previously defined. In many cases they walk both sides of the line between secure and insecure.

5.5 Participant Cluster – The Lazy Expert

Contributing Participants: P3 P11 P13 P24 P28 P31

Dominate Traits:

Participants in this group typically choose convenience over security and being social over privacy.

> [P28] I just figure with email you are going to have to trust it to someone anyways someone is going to be storing it on their computer unless you set up your own domain name and server. Which I don't really have the inclination to do.

> [P24] I put all my information online, because you know here is a tradeoff, if you want your friend to know you more, you should put more information online but I know I know this is a security and privacy problem here but I have no other choice. So I put everything online.

They treat the web like its public domain.

> [P11] Anything I wouldn't say in a crowd I wouldn't put on Facebook.

These participants rationalize a lower level of concern and thus their lower motivation.

> [P31] I have a Mac so no. I think about this sometimes because I heard someone recently talking about anti-virus for Mac and I was like, oh if that exists I probably should get it.

They have mostly unique passwords and share passwords, rarely with trusted people.

> [P28] What I do is have a master password and then Ill add a string to the end of it that is based on the site or the account

> [P13] I might have shared a password, the home computer one that but that s not important and only with my fiance.

Participants in this cluster are careful with passwords.

> [P28] I was actually just making a list of [accounts] cause I always forget my usernames, if I remember my user name I can usually remember my password.

They trust and maintain their own home network not worrying about rare threats.

> [P28] Yeah we have a secured wireless network, We use WPA because I'm aware of the security flaws in WEP

They believe they are not personally a target for computer crimes.

> [P3] I'm going to say most of them arent out to get your average Joe Smoe bank account they are looking for large corporations that have more important things to do and the government.

Often this cluster is responsible for helping others with PAS.

> [P13] my mom, I dunno she gets a virus every other week she'll call me up and be like it says I have a Trojan horse what's that? And I'll look at it and see if I can fix it.

Being a final segment of the pragmatic majority, the lazy expert is a unique and unexpected cluster. Having advanced technical knowledge similar to that of the fundamentalist, but unmotivated to act in a secure way, the cluster presents interesting design challenges.

5.6 Shared Traits

Not all traits were contained in one category of users. While the dominate traits are overwhelmingly in one P-cluster, shared traits could appear in two or three. This is to be expected based on the clustering of participants. For example, the fundamentalist and the lazy experts share higher PAS knowledge, and as such share some traits that are based on higher knowledge. In fact it was found that the shared traits divided in the two dimensions knowledge and motivation along the grades of low, medium or high.

Shared traits of higher knowledge were observed among the fundamentalist, the lazy experts and the technicians. Specifically, fundamentalists and lazy experts had advanced security knowledge through independent learning and had strong opinions against fallback authentication questions.

> [P4] I hate them! I detest them I feel like they don't provide a real sense of security in fact they weaken the security because its not hard for anyone to find out where I was born or what my favourite colour is are all things that would not be hard for a malicious person to figure out if they were willing to. I really don't like them because I think they weaken security.

All three, the fundamentalists, the lazy experts and the technicians, could identify technical terms like security certificates, SSL or Firewalls, had some knowledge of advanced software manipulation, and identified not only anti-virus software but other forms of software protection that they were using.

Demonstrating the technicians' mix of knowledge, they shared the generalization of concerns trait with the marginally aware and the struggling amateur clusters. Unable to point to any specific threat, citing instead "the hackers" or "criminals" or "identity thieves", these participants have a generalization of concerns that is graded as a lower knowledge trait.

> [P27] *Who are the bad guys?* Um hackers I guess. They try to log into your computers, they upload viruses to your computers, they steal your information from your computer. I guess?

Other lower knowledge traits were shared between the marginally aware and the struggling amateurs. These two clusters were more likely to write passwords down insecurely, less likely to be able to define more technical security terms like firewalls, and often admitted that to not understanding a concept but not bothering to look it up either.

> [P12] I've never changed my privacy settings [on Facebook] no. Um I know you can and I know I could probably figure it out if I looked at it but I've just never taken the time to look at it.

Shared traits divided along motivation lines as well. The technicians often choose security over convenience, a trait shared with the fundamentalists, demonstrating dedication to PAS and higher motivation. On the other hand, the technicians also believe "they can't find me" or in being protected simply by obscurity, a trait of medium motivation shared with the struggling amateurs.

Lower motivation traits were shared among the marginally aware, the lazy expert and occasionally the struggling amateur clusters. All three clusters share the trait of strongly trusting their banks to refund their money in the case of security breach, this trait causes them to be more relaxed during online shopping. The marginally aware and the lazy experts both expressed opinions that someone monitoring their computer behaviour would be okay. This is expected because both also believe that their information doesn't matter, and that the honest man has nothing to hide.

> [P31] I don't search anything like how to make a bomb or anything like that I search like Forever 21 and Farmville so I'm not concerned about that at all.

5.7 Discussion

A surprise in the data analysis was the clarity with which dominant traits emerged during the analysis, and how controlled shared traits were. Given the participant clustering technique, i.e. finding like concepts and linking participants along these like concepts, identifying a set of primary reasons for the links between participants occurred naturally. Where ambiguity existed, it was frequently the result of several participants sharing traits with two or more groups, and these relationships between P-clusters are captured in the *shared traits*. Moreover the shared traits themselves were easily divided along the axial coding scales, by returning to the knowledge and motivation dimensions.

An interesting aspect of the reorganization of the data from the first theory to the second, is the change that occurs in the different groups of participants. Using the initial information on participants' self-assessments, only two fundamentalists were identified: P14 and P31. On the reorganization, six participants were classified as fundamentalists, and P14 and P31 were classified into different groups. P14 shared the most commonalities with the group of participants that exhibited mixed knowledge levels and low to medium motivation. On the other hand P31 shares the most in common with the cluster with high knowledge and lower motivation.

P14 does exhibit higher motivation than many of the participants in the cluster he has medium motivation but he is definitely not an outlier in motivation rankings for this cluster. P31 was found to have lower motivation than those considered fundamentalists in the actions she took to proactively secure herself. Both participants differ from fundamentalists in the software they use, the advice they take, and their password schemes causing them to share more traits with other clusters. Reminiscent of Consolvo's fundamentalist who all claimed to heightened concerns but ended up giving away more private information [10], P14 and P31 have a disconnect between what they believe they did and what they were actually doing. By expanding the scope and looking at the participant more broadly, i.e. by reviewing software protection methods, password schemes, learning sources a more accurate picture of the participant is observed.

While the cluster of fundamentalists was disrupted by the modified clustering scheme, other categories were less disrupted. For example, the participants with low knowledge and low motivation were virtually unchanged. Only P28 was moved from the cluster of marginally concerned participants to a cluster of participants with high-knowledge and low motivation. This move is unsurprising, as participants' self-assessment (Chapter 4) focused largely on motivation. Adding a second dimension would be expected to fragment some pre-existing clusters of participants along divisions in knowledge level.

Another interesting attribute of this re-clustering is how evenly balanced participants are across categories. This was not by design, a much less balanced distribution was expected. However, having larger clusters of participants for each category provides confidence that the sample size is sufficient so to saturate the space of participants around the dimensions of knowledge and motivation.

The next two chapter utilizes personas as a tool for interpreting and working with these newly formed categories of PAS users.

Chapter 6

Design Implications - Personas

Given the analysis of the participants, one open question is how best to apply an enhanced understanding of users when designing PAS tools. A common problem with any qualitative result is transforming the data into actionable information that drives design. In this chapter, personas are introduced.These personas have been useful in discussing and further analyzing the categories of users. Rather than treating the categories and category traits as abstract entities, the personas have enabled focused conversations about user needs and aptitudes.

Mulder and Yaar discuss practical ways to create qualitative personas [35] specifically:

1. Conduct interviews with potential users.

2. Create the segmentations based on the interview data.

3. Make the personas real and believable.

The qualitative data and analysis represent the first two steps in constructing personas. By analyzing the relationships between the participant quotes and the participants themselves, five clusters of potential users were observed. I label these potential users categories as different types of PAS personas: Mark, the marginally aware, a persona of low knowledge and low motivation; Robert, the fundamentalist, a security persona with both high knowledge and high motivation; Allison, the struggling amateur, low-high knowledge and lower motivation; Patricia, the technician, low-high knowledge and higher motivation; Henry the lazy expert, with high knowledge but low motivation.

While the qualitative results of chapter 5 and table 5.1 provide dominant and shared traits for the five security personas, a set of information must be synthesized from interview data. Qualitative personas typically consist of a name and descriptive label, a quote, personal demographic information, domain specific information including, objectives and motivations, and a photo (obtained from online stock images). Each persona should also have a profile, which summarizes traits, while describing how the person interacts with the company/product [35]. Once the personas have been specified, they can be used in the creation of scenarios and to focus the discussion of design.

The personas described here constitute a basic persona, with the intention to be malleable into product specific personas depending on the context they used in. The information left to be completed includes the goals, intentions and domain specific information of each persona. In some cases not all personas may be of importance to the application. Even identification of the primary persona would depend on the context of the application. This information is left to be customized for individual applications that make use of the personas.

The next sections introduce each of the personas. As a reference table 5.1 provides an overview of the attributes that went into the profile descriptions here.

6.1 Mark — The Marginally Aware

"I changed my password because when I was signing up it wouldn't let me use my normal one."

Knowledge Low, Motivation Low

Mark is the first to admit he's not very good with computers. "If things go wrong I don't really know how to maneuver around and kinda figure things out. If things go wrong I kinda get someone else to help me or I just avoid it." He picked up a lot of his computer terminology from watching the news or CSI, but that's never stopped him from participating online. He rarely worries about security. "I know there are precautions to take, and to a certain extent I take them but at no point in my day do I worry about them." Like his friend Henry, he believes that he's really not important enough or scandalous enough to have to worry.

To protect himself, Mark does run an anti-virus scanner, and has recently upped his passwords from 123456 to his birthdate. When it comes to passwords, Mark has maybe three, but one is heavily favoured and used for most of his web accounts. Unfortunately, the password requirements keep making him modify his password. Recently he has been thankful for fallback authentication, because he knows he can get back into his account that way.

He likes to shop online particularly for books and DVDs. He places his trust in what the site says, and how popular the site is. "When you first go into the banking site, they show all the awards and everything they have won for security and they could be totally fictional but it still makes me feel better to know they have won awards for that." Overall, Mark has no problem shopping anywhere anytime, because if anything happens to his credit card he's convinced that his bank will refund him. The same as they did last time.

6.2 Robert — The Fundamentalist

"I try not to use any websites that might store my personal information in plaintext."

Knowledge High, Motivation High

Robert has been interested in computers for a very long time, and reads about security and privacy issues in his spare time. His concerns extend beyond his personal security; he worries about the tracking of information and the censoring of information on a global scale. He doesn't participate in social networking because he believes the trade-off isn't worth it. He would feel compelled to monitor his information on these sites, and he doesn't feel the effort is worth it.

To protect himself on his home computer, he maintains his network as best he can, "I believe even WPA is cracked now. I, you know, do my best." Certain files or portions of his hard drive are encrypted, and he runs an anti-virus scan on his Mac once a month. While not every password is completely unique, he maintains a multi-layer scheme, and all his important accounts have their own password. When looking to trust a website with his information, he specifically looks for their security claims. But if a website still uses fallback authentication questions, he may not use it at all.

Robert is reluctant to help other people; he feels like he has been nagging people and no one seems to get it. This could stem from his view that most people are acting insecurely. "... Because some people just don't know what a computer is, let alone security."

6.3 Allison — The Struggling Amateur

"I Google my name a lot and I don't really see anything come up. So unless people are using a different name?"

Knowledge Low-High, Motivation Low-Medium

For the most part, Allison wouldn't consider herself knowledgeable about computers. However she believes that her computer is set up safely with a firewall. She is aware she shouldn't do her banking on public computers, and she doesn't check her bank on public computers, regardless, because she cant remember the answers to the authenticating questions. The most common network she uses is the University's network, "So I believe it's quite safe because usually universities are quite strict about that." Although Allison may be aware of what software she's using to protect her computer, she's doesn't always know what the software is for.

Allison will make changes to her security or privacy habits based on advice from others. The advice isn't always completely accurate, but she doesn't try to discriminate between the quality of the different advice she receives. "I remember my boyfriend telling me, 'Oh take your birthdate off of Facebook and don't put where you live' and I said why? cause people like sending you birthday wishes right? 'Because I heard that they can use your birthdate and where you live to get your credit card information.' I don't know how true that is but I was a little concerned and I thought why not be proactive."

To protect her privacy, Allison relies on online anonymity, "I Google my name a lot and I don't really see anything come up. So unless people are using a different name I can't really think of anything that I would be worried about." When it comes to passwords, Allison may only identify with one password, but she actually has a small number of password layers based on the type of website she's using. She has occasionally written down her password and "Once our neighbour in the apartment asked if she could use our Internet, so we gave her the password. It was no big deal."

6.4 Patricia — The Technician

Patricia has been refining her security habits based on reasonable advice from friends and what she has learned from news and blogs. She is fairly confident and comfortable with the

"The Nigerian princes and stuff. That stuff, I know it when I see it now."

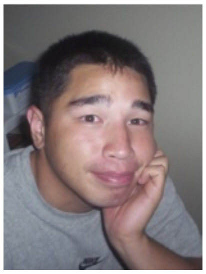
Knowledge Low-High, Motivation High-Medium

decisions she's made, like saving her passwords on the computer. Anything that she's found questionable she asks her friend Henry or does a search on it. "I would like to think that most of the time I don't usually go to sites that would cause a problem, or download weird things. There was that one time I downloaded a screensaver and my computer immediately told me it was a virus and I felt stupid." On top of her anti-virus and firewall, she's also downloaded Ad Block and NoScripts for Firefox.

Patricia actually only has a couple of passwords, a low security and a high security one. Both are based on the same personal thing in her life that she believes no one would be able to guess. Because she saves her passwords on the computer, she is hesitant to let other people use it. She has given a password to her boyfriend before, but only on rare occasions.

When it comes to online shopping, Patricia puts a lot of trust in the rating systems on Amazon and eBay. If she is not using those sites, she trust her intuition about the site: Provided it's not sketchy looking she uses it. Patricia's Facebook page isn't empty, but she is reluctant to participate in online social networking, preferring her privacy. She has spent time finessing the privacy settings on Facebook.

6.5 Henry — The Lazy Expert



"I'm sure hackers are still a concern, just more for major corporations and not the average person"

Knowledge High, Motivation Low

Henry loves using his computer and is an expert user. He has learned about encryption in school, and while he thinks that it's a cool idea, he doesn't believe it's worth the effort. He treats the information on his online social networks as if its public, because it's on a server controlled by someone else. By the same token, he doesn't really mind being monitored at work because, in the end, his information isn't that interesting and he's not likely to be the target of an attack. Being aware of the problems with WEP was not reason enough for Henry to change his network setting. "If some unix script kiddie wants to crack my WEP, then he can get in because I've honestly had trouble with WPA just being a pain." The most recent change in Henry's habits was to stop torrenting because he downloaded a virus which was hard to clean from his computer.

When it comes to passwords, Henry has developed a system that allows nearly all of his passwords to be unique. This comes with the price of occasionally writing down his password or username, but that information is always kept securely in a secret location in his room. He gave a password to his mom, once, because he needed directions while driving.

Most people know that Henry is good with a computer, so he has often been called on for IT support. "My mom, I dunno she gets a virus every other week shell call me up and be like 'it says I have a Trojan horse what's that?' And Ill have to look at it and see if I can fix it."

6.6 Summary

This chapter presented a set of five personas representative of the five categories of users discovered as a result of the study. The personas are: Mark - the marginally aware, Robert - the fundamentalist, Allison - the struggling amateur, Patricia - the technician, Henry - the lazy expert. These personas are meant to feel like real users but are actually fictional characters developed from a set of traits. The traits were originally discovered as a result of the affinity diagram and the participant clustering process. Each persona's description converts the basic set of traits into a coherent story about the character to enable the designer, engineer or researcher to better visualize the potential users.

One challenge that remains is to evaluate the participant clusters and the personas. In the next chapter begins to address this challenge through thought experiments and by analyzing a separate set of independently interviewed users in light of these user categories.

Chapter 7

Applications for Requirements and Design

Researchers have frequently evaluated the use of personas in requirements and design, seeking to determine whether personas aid in the process [8, 9, 31, 33]. However, the goal is not to evaluate whether these personas are valuable a central premise of this work is that there does exist a value to personas. Instead, the goal is to determine whether the categorizations and personas developed here are a useful representation of the design space in PAS. This chapter explores the potential applications of the PAS personas. First an analysis of a separate set of participant, demonstrating the application of the new user categorizations. Second the personas are used to evaluate existing designs and then used to explore related the aspects of the personas within the study that could encourage new design or requirement research avenues.

7.1 Evaluate User Categories

One way to evaluate and use the personas is to look at a separate set of participants. To perform this analysis, my colleagues made available a set of participants collected from a study of WiFi security practices [44]. Table 7.1 describes these participants.

7.1.1 Study Method

The WiFi study was a two-part study. Participants, recruited at random from area cafes with open wireless access points, were questioned about security practices, and given a demonstration of a packet sniffing attack in the first phase of the study. 3 – 4 weeks later,

Table 7.1: Participants in the WiFi study.

ID	Occupation	Age
W1	Mathematics Ph.D. student	29
W2	English student/retail employee	22
W3	Retired sales manager	67
W4	Government employee	24
W5	MBA student	26
W6	MBA student	29
W7	Chemical Engineering/MA student	23
W8	Investment analyst	23
W9	Physiotherapy/Recreation student	24
W10	Sociology MA student	26
W11	Behaviour therapist	30
W12	Security expert	35+

participants were re-interviewed to determine whether WiFi behaviour had changed, and what factors influenced whether participants did or did not change behaviour.

7.1.2 Results

The participants were coded by independent researchers and in this section some of the observations that determined the participants' placement within a grading of knowledge and motivation are explained. As one component of the study, my colleagues questioned participants' on their knowledge of security technologies. One participant, W12, had extensive knowledge of security and privacy. Four other participants, W1, W2, W4, and W8 had mixed knowledge of security technologies. Other participants, W3, W5, W6, W7, W9, W10, and W11 had limited to no knowledge of technologies such as Secure Socket Layer (SSL) protocols and Virtual Private Network (VPN) Connections. Table 7.2, shows the placement of this set of WiFi participants on the knowledge to motivation scale.

In terms of motivation for security, participant W12 was both highly knowledgeable and highly motivated, similar to Robert (the fundamentalist) of the persona set. In transcripts, it was observed that this participant describes how he avoids using webmail in a browser because of "man-in-the-middle" attacks. He also avoids doing online banking on public WiFi, noting that he "knows enough of security to know it's reasonable," but had some concerns about highly sophisticated potential attacks on SSL connections.

Table 7.2: WiFi participants grouped according to motivation and knowledge

High Knowledge			W12
Mid Knowledge	W1, W8	W2, W4	
Low Knowledge	W5, W9	W3, W6, W7, W10, W11	
	Low Motivation	Mid Motivation	High Motivation

W1 and W8 both had mixed levels of security knowledge, but very little motivation. During the follow up interview, W8 noted that:

> [W8]Well, the way I see it is, if somebody is out there logging what websites I visit and sells it, that's fine.

W1 initially noted that he was a "careless WiFi user" and saw little reason to change. After all, anyone who actually tried to eavesdrop on someone would have to have "psychological problems" according to W1, and so it just wouldn't happen. While these participants don't directly align with Henry, our lazy expert (they have slightly less knowledge) they share one important characteristic. Both Henry and these new participants believe that, regardless of the fact that they may be exposing themselves to risks and an awareness of those risks, the chances of something happening to them are low, and it's not worth the bother to change any behaviours. They are not a target, and so do not need to change.

W5 and W9 know very little about security, but do not see any reason to change, even in light of a packet sniffing demonstration. W9 notes:

> [W9]I'm very flexible, so if people want to know where I'm going, OK. I don't care.

These two participants echo actions and traits of Mark (the marginally aware).

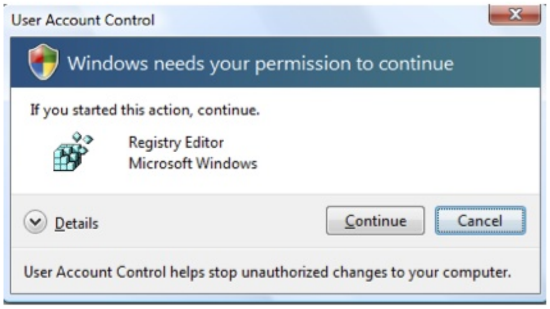
W2 and W4, both with moderately higher knowledge of security, were classified as having medium motivation because of actions they took proactively to protect themselves. Both noted expired security certificates, and only went to websites they trusted with these expired certificates, a behaviour echoed by W12. As well, both were aware of SSL connections and paid attention to 'https' and padlock icons to indicate secure web pages. While their area of interests are slightly different from Patricia (the technician), they share Patricia's focus on a single area of concern.

Initially, positioning W4 was somewhat challenging; identifying this participants motivation resulted in some back and forth between researchers. This participant was very careful about login prompts checking for encrypted connections and was reluctant to engage in online banking or other potentially risky behaviours on public WiFi. However, W4 lacked a sophistication regarding his security habit and did not proactively educate himself about security, and he was very focused on login security but did not consider other aspects of PAS. W4 has higher motivation but shares more characteristics with Patricia than with Robert.

Finally, W3, W10, and W11 all reported changes in behaviour, particularly an increased caution using open WiFi as a result of the packet sniffing demo. Demonstrating moderate motivation and limited knowledge much like the struggling amateur Allison. W11 taught her coworkers about SSL connections and worked with her peers to ensure everyone was being more secure. W3 was reassured about his online banking, and paid careful attention to connections after the demo. Finally, both W6 and W7, despite a lack of knowledge, proactively took steps to protect themselves, justifying their placement in medium motivation. These participants knew little but all wanted to be secure. They typically welcomed advice from someone like Henry about how best to protect themselves, and generally tried to follow Henry's advice.

7.2 Evaluate Design

An aspect of personas is their utility in discussions of PAS tool design or analysis. This section is a thought experiment examining two common pieces of security software bundled with modern operating systems. The first is User Account Control (UAC) in Windows Vista, which notifies the user when any administrator-level task is initiated through a dialog, for example figure 7.1. The second is Windows Vista Firewall, which has default settings and a set of customization screens.



User Account Control

Windows needs your permission to continue

If you started this action, continue.

Registry Editor
Microsoft Windows

User Account Control helps stop unauthorized changes to your computer.

Figure 7.1: Microsoft's User Account Control

Consider, first Windows Vista UAC. It is difficult to articulate how this control appeals to any security persona developed here. Some (Robert, the fundamentalist, Patricia, the technician) dislike the UAC because it doesn't provide them with enough information. Others (Mark, the marginally aware, Allison, the struggling amateur) find it irritating because it asks questions that they cannot really answer. Finally Henry, the lazy expert, finds it annoying because it prompts him when he would rather ignore PAS as much as possible.

Windows Vista Firewall, in contrast, allows lower motivation personas (Mark, the marginally aware and Henry, the lazy expert) to ignore it. Mark may not even know it exists. Allison, the struggling amateur, is reassured by its presence, and only attends to it when she receives a security warning. Finally, Patricia, the technician, finds the simple GUI for customizations useful. Only Robert, the fundamentalist, might find the application of limited usefulness for him based on the lack fine grained customization.

In summary, Windows user account control is poorly designed. Even its target persona, arguably Patricia, receives insufficient information to make informed decisions. In contrast, Windows Vista Firewall seems functional for most users. Even people like Robert, recognize the fact that, for other people who know less than him, it's probably a good idea to have a basic firewall with simple settings.

7.3 Focus of Concerns

Other interesting dimensions in the data seem to cut across participant category boundaries. Many participants focus on *either* privacy *or* security as their primary concerns. For example, a participant who uses MacOSX justifies his safety as a result of the fact that Apple computers are rarely targeted by viruses seem to place primary importance on security, not privacy. On the other hand, a participant who dislikes monitoring are frequently more concerned about privacy. Another cross-cutting aspect was global versus local concerns: Are people more concerned about preventing hackers from attacking their system (global threats) or physical access to their machines, i.e. securing it from their co-workers in an office environment (local threats). Taken together these aspects represent four potential areas of focus for users. The four areas of focus being local security, global security, local privacy and global privacy.

Local Security: These people tended to worry mostly about the physical security of their computer system. They worry about having their computer or cell phone stolen and they generally will not let untrusted people near their system.

> [P6] It breaks down into my whole thought on security, it's the people you talk to the people who have actual access your equipment, they are the ones who actually have the most chances of tearing you down the rest of the Nigerian

> princes and stuff. That stuff I know it when I see it now. And that's stuff you can protect against without too much work.

Global Security: People who are worried about security on a global level are more concerned about threats they can't see. They will have stronger passwords but are more likely to write them down.

> [P3] I'm a firm believer in writing down passwords because who you need to protect yourself from is not people in your office, because they can get into anyones computer anyways. Especially the system administrator. So what you are protecting yourself from is people outside your physical building and that what the password is for.

Local Privacy: For most participants if they were worried about privacy it seemed to be on a localized level. They were worried about very specific threats like their boss or their parents reading their information. Usually these people craved more control over their information but had no real concerns about putting the information out there.

> [P2] Well I was actually just reading an article today from the Ontario college teachers magazine saying that don't put anything on facebook that you wouldn't want your principle to see. So I no longer have pictures of me drinking or otherwise being silly, you know just in case a student or principal or whatever was to see it. I'm more careful with that now, but I do have my privacy settings on Facebook fairly secure.

Global Privacy: Those participants that worried about privacy on a more general level were concerned about things like information retrieval at the data mining level. They were concerned about what companies such as Facebook or Google were learning about them. Most participant who were worried about this were less active online and tended to censor their information more.

> [Moderator] Who are the bad guys? [P10]Anybody who has ads because they are always trying to get as much information to see who you are and where your going, that it's unnecessary... I guess I can't really consider them bad guys just everybody who uses the internet to try and get information through non-traditional ways

Very few participants would focus there attention in all 4 areas, in fact most participants only had 1 or 2 areas that they were particularly verbal about. User's with Robert's traits are likely to focus their attention towards multiple, if not all, areas. On the other hand, Mark, and other marginally aware users may not be concerned about any. Personas of the middle ground, Henry, Patricia and Allison, are likely to pick one or maybe two areas of concern. Which ever areas the participants seemed to focus on their PAS mechanisms followed. Future work could include further exploring the idea and implications of local and global areas of interest. Ideally, researchers could investigate way to increase the user's focus in all four areas.

7.4 Explain, Understand and Predict Users

A common selling point of personas is their potential to explain or even predict users actions when interacting with a new product or service. In Pruitt and Grudin work with personas they observed this basic human ability to understand people in this way [38]. They wrote:

> "If team members are told, 'Market research shows that 20% of our target users have bought cell phones', it may not help them much. If told them 'Alan has bought a cell phone' and Alan is a familiar Persona, they can immediately begin extrapolating how this could affect behaviour. They can create scenarios. We do this kind of extrapolation all the time, we are skilled at it, not perfect, but very skilled."

Throughout this research project the personas have been helpful to describe extended phenomena that was observed in the data. For example, there were unique relationships that existed between the personas. Henry, the lazy expert and users like him often took on the role of helpers and educators of others like Allison, the struggling amateur or Mark the marginally aware. In some cases Henry gives advice that he does not follow himself. Similarly, Robert, the fundamentalist, if he takes the time to educate anyone it would be someone like Patricia, the technician, or possibly Henry, the lazy expert. This ensures for Robert that whoever he is educating will understand what he saying and why it is important. Based on the observed behaviours, Henry, the lazy expert, becomes central and important to the dissemination of knowledge. Just like targeting the early adopters of technology can be vital to achieving critical mass on a new service, Henry is a vital persona to spread information. Depending on the context of application the relationships between users could be leveraged to improved usability.

Another area where the personas proved useful was in examining various cost benefit ratios. Most significantly was the balance between security and convenience or being private and being social. While it is easy to see that users with higher motivation such as Robert value security and users with lower motivation like Henry and Mark prefer convenience. There remains a spectrum between these two extremes for Allison and Patricia who are trying diligently to find the balance. For these two personas each PAS decision could be influenced one way or the other and is particularly dependant on the user's focus of concern as previously described.

When it comes to online participation, including social networks or online shopping, there exists a spectrum of how much information the personas are willing to let go. On one end Henry and Mark are paying very little attention to what information they are putting out. They are not likely to even remember how many online accounts they maintain, only that the number is big. On the other extreme, Robert is comfortable completely opting out of a service if he feels uncomfortable with the site or the site's practices. Again, in the middle, there is a tension between censoring information and controlling information. Patricia prefers to censor her information and put less information online, not surprising since she prefers privacy over being social in most cases. Allison, the struggling amateur, wants to maintain control over her information. So when using social networks such as Facebook, Allison is more trusting of the website and the settings that she has in place to keep her information from people she doesn't know. However by having lower motivation, Allison's settings are rarely updated and marginally understood, and indicate potential PAS areas of improvement. The tension between censorship and control is of interest to not only those fighting for privacy but the companies asking for participation.

Finally, as others have noted, context often has an effect on participants behaviours [10]. Some participants are more security conscious at work because of company policies or greater personal liability, and some participants are more PAS conscious when acting in the role of a helper to other users. Context concerns were overwhelming observed in participants of the lazy expert cluster resembling Henry because Henry's low motivation often switched based on his context. All of these additional data points provide an opportunity for enriching the categorization of participants.

All of the above extended phenomena were aided by the imagery of the personas. It is my belief that other context and specific application can be improved as well either using these personas or by building them up further as described in the next chapter.

Chapter 8

Conclusion

This thesis has presented findings from a qualitative interview study of users' opinions, practices and justifications for behaviour relating to privacy and security. The observations made during the interviews allowed for the creation of five user classifications. This final chapter reviews and highlights the study's goals and results. This chapter also highlights some limitations of this work and makes suggestions for future work to extend and validated the research put forth here.

8.1 Goals

At the outset of this research it seemed apparent that the users of PAS tools have been described in many fragments but rarely as a heterogenous community. Knowing that a better understanding of the user community often leads to quality informed design, the goal of this research was to look beyond specific problem areas and and understand the breadth of the community. The research goals outlined the desire to understand:

- How are people currently interacting and dealing with computer PAS?
- If the current classifications of users provide an accurate view of the user community?
- How can differences between people allow for a better understanding of the community?
- Which attributes allow for clustering of individuals on both a usable and relevant level?
- Which design implications could be explored by better understanding and embracing the differences between people?

The qualitative interview study of users' PAS opinions, practices and justifications was designed to counterbalance two properties of previous work. First, in classifying users many researchers conduct a survey and take quantitative approach. This method can allow for details of of the user to be lost in the process including specific motivations, justifications or questions outside of the scope of the survey that may influence the user's decisions. Secondly, previous work often tries to isolate privacy from security or vice versa. Privacy and security as concepts are so closely related and dependant on each other that it is very difficult to consider one without the other. For these reasons the study was designed as a set of qualitative in-depth interviews covering a variety of topics of privacy and security.

8.2 Results

The results summarize 32 semi-structured interviews that were conducted, the participants answers were reviewed and analyzed for themes within participants and similarities and differences across the group of participants. Open coding allowed for the formation of the affinity diagram consisting of 24 concepts and 85 subconcepts. The concepts were Q-clustered and graded around dimensions knowledge and motivation. At the same time the participants were P-clustered by similarities and along the dimensions of motivation and knowledge. This co-axial analysis of the participants and data became crucial to understanding the breadth of the community. Either dimension taken alone would only capture a fragment of the community.

As such, the data revealed five participant clusters, the marginally aware, the fundamentalists, the struggling amateur, the technician, and the lazy expert. Aside from the generalized motivation and knowledge assessments, clear traits were observed for these P-clusters. On average ten dominant traits were found for each P-cluster. Furthermore some traits were shared equally among two or three P-clusters. Shared traits were heavily influenced by either knowledge or motivation and easily divided among the P-clusters by these influences.

As a visualization and for better usage of the clusters and cluster traits, personas were created. Five personas named Mark, Robert, Allison, Patricia and Henry were described in basic form. Each of the personas represent a composite of the cluster but could still be further developed depending on the context of the application.

8.3 Limitations and Future Work

In any qualitative study, the goal is to build themes grounded in the data that have high likelihood of generalizing to a larger population. In this study, there exists two potential risks to the impartiality of the data. The first involves personal biases as a security researcher, and the second involves the demographics of the study participants.

In any data coding exercise, researchers use their experience with data to assign interpretations. In the analysis of data, I make several judgments on the data, including what factors represent high versus low knowledge, how reasonable advice that participants have acted on is, what types of password schemes represent different motivation levels, and many others. Furthermore, in guiding the selection of primary categories for axial coding, I may have been influenced significantly by past research in this area (e.g. [17]). I cannot claim to be unbiased in my interpretations of the data. However, also note that, in clustering participants, the bias is partially control for by using concrete participant comments to define the concepts, rather than subjective judgments of knowledge or motivation. As stated in Chapter 3, if participant quotations were similar in concept or subconcept on the affinity diagram, the participants were linked. Then the strength of these connections were used to define P-clusters. Any ambiguities in similarity between various participant quotations and actions were resolved collaboratively by researchers within the same lab and with access to the raw data. Subjective judgments of knowledge and motivation were dimensions used to combine concepts on the affinity after P-clustering, not to provide the primary clustering of participants. One litmus test of the quality of the subjective judgments is whether the attributes listed on Table 5.1 align well with the knowledge and motivation levels attributed to each of the personas.

The second risk is a demographic risk. Researchers are split on the importance of demographics in an analysis of user characteristics [22, 41]. Unfortunately, the participants were typically younger (average age 26.3, SD 5.9 years). As such, the participants were, without exception, digital natives, users who grew up with the internet in the online world. Sheehan might argue that these participants are more likely to be pragmatists, and that a group of more mature digital immigrants might increase the size of fundamentalists (Robert) and marginally awares (Mark). In future work, I plan to conduct additional interviews with more mature participants to partially address the shortcomings of this work.

Beyond further developing the personas, future work includes using the personas to drive experimental design. For designers of PAS tools interests lie in what tools would help Mark, the marginally aware or Allison, the struggling amateur make significant changes to their security habits. While password requirements frustrate the these categories, they do slowly improve their passwords. New password technique like graphical passwords may significantly enhance the security of these users, while having less effect on users like Robert and Henry, the high-knowledge categories.

8.4 Contributions and Summary

Through qualitative interviews an interesting interaction between the motivation and knowledge of our participants was observed. By clustering sets of participants around levels of motivation, levels of knowledge and like-minded ideas, five categories of PAS users are established. Using traits of these users drawn from affinity diagrams, five unique personas were created. There are five major contributions of this thesis:

1. A study with the intent of examining the differences between users, without pre-specified concepts for classifications.

2. A critical examination of previous classifications attempts and potential shortcomings of their methodology.

3. A unique categorization of the PAS community based on two dimensions, knowledge and motivation.

4. Five user personas, developed based on the categorization.

5. Two thought experiments demonstrating the use and power of the PAS personas.

As a result of the research approach, this thesis represents a significant advance in the state-of-the-art. The groups of users are new and provide a more nuanced picture of types of users than has existed in past research. Personas, while popular in design, have not been used in the security and privacy domain. The goal is to examine the differences between users, to broaden the discussion about users, to inspire informed design for usable privacy and security tools in the future.

APPENDICES

Appendix A

Terminology

Affinity Diagram A technique adopted in grounded theory used to organize ideas and data. The defining property of affinity diagramming is working from the bottom up that is, from quotes to concepts and eventually to theories.

Cookie A cookie is a relatively harmless file stored by a web browser on the user's personal computer to allow a website to restore a personalized state of a website.

Firewall A firewall is a software protection method that exerts control over network communication of a personal computer.

Open Coding The initial step in grounded theory, conceptualizing the data from its raw form into usable levels of abstraction.

PAS Privacy and Security

Persona Personas are a HCI tool adapted from Marketing in which fictional characters are created to represent the different user types of a demographic or user group.

Security Certificate Information issued by a certificate authority and used to establish a secure connection through the SSL or TLS protocol.

SSL or TLS Secure socket layer (SSL) and Transport layer security are protocol to establish encrypted communications over a network.

Appendix B

Participant Profiles

Participant #	Year of Birth	Location	Occupation
P1	1982	Toronto, ON	Programmer
P2	1984	Sarnia, ON	High School Teacher
P3	1985	Sarnia, ON	Programmer
P4	1986	Cupertino, CA	Programmer
P5	1983	St. Paul, MN	Retail Associate
P6	1982	St. Paul, MN	Freelance Web Designer
P7	1984	Waterloo, ON	Morgage Specialist
P8	1986	Waterloo, ON	Programmer
P9	1986	Kitchener, ON	Social Worker
P10	1983	St. Paul, MN	Retail Manager
P11	1984	Windsor, ON	IT Support
P12	1982	Kitchener, ON	Elementary School Teacher
P13	1982	Waterloo, ON	Web Developer
P14	1986	Hamilton, ON	Planning Technician and Part Time Student
P15	1986	Waterloo, ON	Student
P16	1985	Waterloo, ON	Student
P17	1960	Kitchener, ON	Admin Assistant and Part Time Student
P18	1986	Waterloo, ON	Student
P19	1984	Waterloo, ON	Student
P20	1965	Waterloo, ON	Librarian and Part Time Student
P21	Withheld	Waterloo, ON	Student
P22	1987	Waterloo, ON	Student
P23	1977	Waterloo, ON	Student
P24	1982	Waterloo, ON	Student
P25	Withheld	Toronto, ON	Programmer and Part Time Student
P26	1985	Waterloo, ON	Student
P27	1984	Waterloo, ON	Student
P28	1985	Waterloo, ON	Student
P29	1986	Waterloo, ON	Student
P30	1985	Waterloo, ON	Student
P31	1983	Waterloo, ON	Student
P32	1986	Waterloo, ON	Student

Appendix C

Study Materials

C.1 Recruitment Script

Hi,

As part of my MMath Thesis, Im doing research on opinions and attitudes towards computer security, under the supervision of Professor Dan Berry. The findings from the project will allow us to better understand how the general public views security measures on their computer.

Participation will take approximately 60 minutes of your time and will be scheduled for a time that is mutually convenient. The interview can be completed in person, by phone or over Skype. Following the interview you will be asked to complete a short questionnaire regarding your computer usage.

If you agree to participate, you will be asked a series of questions regarding your opinions and behaviours regarding using your computer. Additionally, participants will receive a $5 gift certificate for Tim Horton's or Dunkin' Donuts as a token of appreciation. If you're interested in participating, please email me back so we can set up a mutually convenient time.

Thanks for your help, Janna-Lynn Weber

This project was reviewed by, and received ethics clearance through, the Office of Research Ethics at the University of Waterloo. Should you have any comments or concerns resulting from your participation in this study, please contact me or my advisor, Dr. Dan Berry at dberry@cs.uwaterloo.ca or Dr. Susan Sykes in the Office of Research Ethics at 519-888-4567, Ext., 36005. Or ssykes@uwaterloo.ca

C.2 Interview Script

C.2.1 Goal

To gather a better understanding of the different ways people think and act in regards to security and privacy. In order to further develop and validate a set characteristic themes or personas that will aid in the design of secure software.

C.2.2 Introduction Script

Thank you for helping me out. Im interested in learning about the ways people think about security and privacy. With you permission Id like to record the session so that I can review it later. Just for my use only.

Do you have any questions regarding the information letter or the consent form? Do you agree to volunteer?

So with your permission, Im going to begin recording now. Remember that this is to learn about you, there are no wrong answers and every bit of information you can give me is helpful.

C.2.3 Potential Question List

Tell me a little about you? What do you do? How old are you? What do you remember most about your first computer? When did you get email? Do you still use that account?

General Computer Familiarity and Internet Usage

- Do you own a laptop or a desktop?
- What operating system do you use?
- What web browser do you normally use?
- Have you tried other web browsers?
- How many hours were you online yesterday? Is that an average day?
- How many hours were you online last week? Is that an average week?
- What do you like to do online during the day?
- What is your current job?

- What do you typically use a computer for at work?
- How often do you typically use the computer outside of work?
- Have you ever purchased something online?
- Think about the last thing you purchased
- Where was it from? Have you bought from them before?
- Are you a member of any online social networks?
- Do you have a blog?
- Have you ever been concerned about who might be reading your blog or social network profile?
- Do you know if there are pictures of you or that you have taken online?
- Do you use your cell phone to access the internet?
- How often would you use your cell to access the internet? What are you typically doing with it?
- How often would you say you use public WiFi?
- Have you ever use a neighbors unsecured WiFi?
- In your opinion, whats so great about computers and the internet?
- What are some examples of things of yours that you wouldnt want to be online?
- Do you have any concerns with the current state of computers or the internet?

Online Security

- How many email addresses do you currently check?
- How many online accounts do you think you have?
- How do you feel about online banking?
- Tell me about the last new account you signed up for
  - What was it? How did you find it? Why did you sign up?
  - How long to did it take you to sign up?

  - When did you last log into it?

- If you had to guess, how many passwords and userIDs would you say you have?
- How many unique passwords do you have?
- Do you have a method for remembering them or keeping track of them?
- Have you ever written a password down?
- Have you ever shared a password?
- Walk me through what you would do if you were asked to create a new password
- For an email account
- For a bank account
- For a online newspaper
- Have you ever changed your browser settings? When? Why?
- Can you tell me what a cookie is?
- Can you tell me what a certificate is?
- Can you tell me what a Firewall is? Do you use one?

Physical Security

- How would you describe the atmosphere at your workplace with regards to security?
- When you are at your computer at work do you have to login to access it?
- Do you turn your computer off at the end of the day?
- When you get up for a bathroom break do you lock your computer?
- When you go for lunch would you lock your computer?
- Are there any websites or applications that are blocked at your work?
- Do you or anyone you know have a way to access those sites anyways?

Home Practices

- How have you protected your computer at home?
- Do you use any software to protect your computer?
- Do you use a wireless network?
- Have you secured your wireless network?
- Can you tell if your neighbours have secured their networks?

Concerns & Practices

- When we think about computers and the internet, who are the bad guys? How do they profit?
- What is a bigger concern for you: unknowingly spreading your own information or having a stranger spread it for you?
- If you knew there was a 50% chance, that at any given time, someone was watching everything you did on your computer, would you change your behaviours? How so?
- Which is the bigger concern for you?
  - Hackers trying to get into your account or your ex-lover trying to get into your account
  - Co-Worker looking at your computer or Spies looking at your computer
  - Your boss looking at your Facebook or the government looking at your Facebook

Experiences

- Tell me about the last time you forgot your password or userID
- Tell me about a time when you had one of your accounts compromised
- Do you believe youve ever had a close call with someone trying to scam you?

Wrap up

- Where did you learned about all the security measures you take?
- Who would you say is less secure than you? Why?
- Who would you say is more secure than you? Why?
- Any other comments?
- If I had any follow up questions would you mind if I emailed you?

C.3 Thank You Email

Dear Participant,

We would like to thank you for your participation in our study. As a reminder, the purpose of our project was to conduct a qualitative study on the security and privacy habits of real people. If you are interested in learning about the results of this study, email me and I will keep you up to date on the projects progress.

Sincerely,
Janna-Lynn Weber
J6Weber@cs.uwaterloo.ca

As a reminder, any data pertaining to you as an individual participant will be kept confidential. This project was reviewed by, and received ethics clearance through, the Office of Research Ethics at the University of Waterloo. Should you have any comments or concerns resulting from your participation in this study, please contact Dr. Susan Sykes in the Office of Research Ethics at 519-888-4567, Ext., 36005 or ssykes@uwaterloo.ca.

Appendix D

The Affinity Diagram

The affinity diagram created from over 500 participants quotes can be downloaded as a PDF from: `http://se.uwaterloo.ca/~dberry/FTP_SITE/students.theses/janna.weber/AffinityDiagramPhotos.pdf`

Bibliography

[1] M.S. Ackerman, L.F. Cranor, and J. Reagle. Privacy in e-commerce: examining user scenarios and privacy preferences. In *Proceedings of the 1st ACM conference on Electronic commerce*, page 8. ACM, 1999. 4, 6, 15, 20

[2] A. Acquisti and R. Gross. Imagined communities: Awareness, information sharing, and privacy on the Facebook. *Lecture notes in computer science*, 4258:36–58, 2006. 1, 14

[3] A. Adams and M. A. Sasse. Users are not the enemy. *Commun. ACM*, 42(12):40–46, 1999. 1, 2, 11, 14

[4] A. Adams, M.A. Sasse, and P. Lunt. Making passwords secure and usable. *People and Computers*, pages 1–20, 1997.

[5] M. Aoyama. Persona-and-Scenario Based Requirements Engineering for Software Embedded in Digital Consumer Products. In *Proceedings of the 13th IEEE International Conference on Requirements Engineering*, page 94. IEEE Computer Society, 2005. 19, 20

[6] B. Berendt, O. Günther, and S. Spiekermann. Privacy in e-commerce: Stated preferences vs. actual behavior. *Communications of the ACM*, 48(4):106, 2005. 1, 4, 6, 14, 16

[7] H. Beyer and K. Holtzblatt. *Contextual design: defining customer-centered systems.* Morgan Kaufmann Pub, 1998. 2

[8] Y. Chang, Y. Lim, and E. Stolterman. Personas: from theory to practices. In *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*, pages 439–442. ACM, 2008. 19, 64

[9] C.N. Chapman, E. Love, R.P. Milham, P. ElRif, and J.L. Alford. Quantitative evaluation of personas as information. In *Human Factors and Ergonomics Society Annual*

Meeting Proceedings, volume 52, pages 1107–1111. Human Factors and Ergonomics Society, 2008. 19, 64

[10] S. Consolvo, I. E. Smith, T. Matthews, A. LaMarca, J. Tabert, and P. Powledge. Location disclosure to social relations: why, when, & what people want to share. In *CHI '05: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 81–90, New York, NY, USA, 2005. ACM. 6, 16, 17, 57, 71

[11] G. Conti and E. Sobiesk. An honest man has nothing to fear: user perceptions on web-based information disclosure. In *Proceedings of the 3rd symposium on Usable privacy and security*, page 121. ACM, 2007. 14

[12] A. Cooper. *The inmates are running the asylum: Why high tech products drive us crazy and how to restore the sanity.* Pearson Higher Education, 2004. 18, 19

[13] C. Courage and K. Baxter. *Understanding your users: a practical guide to user requirements: methods, tools, and techniques.* Morgan Kaufmann Pub, 2004. 6

[14] Lorrie Cranor and Simson Garfinkel. *Security and Usability.* O'Reilly Media, Inc., 2005. 2, 4, 12

[15] J.W. Creswell. *Qualitative inquiry & research design: Choosing among five approaches.* Sage Publications, Inc, 2007. 23

[16] N. De Voil. Personas Considered Harmful. 2010. 19

[17] P. Dourish and K. Anderson. Collective information practice: emploring privacy and security as social and cultural phenomena. *Human-computer interaction*, 21(3):319–342, 2006. 2, 13, 74

[18] P. Dourish, R.E. Grinter, J. Delgado de la Flor, and M. Joseph. Security in the wild: user strategies for managing security as an everyday, practical problem. *Personal and Ubiquitous Computing*, 8(6):391–401, 2004. 2, 13

[19] P. Dunphy, J. Nicholson, and P. Olivier. Securing passfaces for description. In *Proceedings of the 4th symposium on Usable privacy and security*, pages 24–35. ACM, 2008. 1, 14

[20] Katherine M. Everitt, Tanya Bragin, James Fogarty, and Tadayoshi Kohno. A comprehensive study of frequency, interference, and training of multiple graphical passwords. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, pages 889–898, New York, NY, USA, 2009. ACM. 12

[21] D. Florencio and C. Herley. A large-scale study of web password habits. In *Proceedings of the 16th international conference on World Wide Web*, page 666. ACM, 2007. 1, 14

[22] B. Friedman, D. Hurley, D.C. Howe, E. Felten, and H. Nissenbaum. Users' conceptions of web security: a comparative study. In *Conference on Human Factors in Computing Systems*, pages 746–747. ACM New York, NY, USA, 2002. 2, 13, 74

[23] S. Gaw and E.W. Felten. Password management strategies for online accounts. In *Proceedings of the second symposium on Usable privacy and security*, page 55. ACM, 2006.

[24] S. Greenberg and B. Buxton. Usability evaluation considered harmful (some of the time). In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 111–120, New York, NY, USA, 2008. ACM. 18

[25] J. Grimmelmann. Facebook and the social dynamics of privacy. *Iowa Law Review*, 95(4), 2009.

[26] J. Grudin and J. Pruitt. Personas, participatory design and product development: An infrastructure for engagement. In *Proc. PDC*, pages 144–161. Citeseer, 2002. 6, 19, 20

[27] C. Herley. So long, and no thanks for the externalities: The rational rejection of security advice by users. 2009. 2, 14

[28] J. Hitchings. Deficiencies of the traditional approach to information security and the requirements for a new methodology* 1. *Computers & Security*, 14(5):377–383, 1995.

[29] P. Hoonakker, N. Bornoe, and P. Carayon. Password authentication from a human factors perspective: Results of a survey among end-users. 2009.

[30] Imperva. Consumer password worst practices'. `https://www.imperva.com/lg/lgw.asp?pid=379`, 2010. 15

[31] P.T.A. Junior and L.V.L. Filgueiras. User modeling with personas. In *Proceedings of the 2005 Latin American conference on Human-computer interaction*, page 282. ACM, 2005. 19, 20, 64

[32] P. Kumaraguru and L.F. Cranor. Privacy indexes: A survey of Westin's studies. *Institute for Software Research International*, 2005. 15

[33] F. Long. Real or imaginary: The effectiveness of using personas in product design. In *Proceedings of the Irish Ergonomics Society Annual Conference*, pages 1–10, 2009. 6, 19, 64

[34] T. Miaskiewicz, T. Sumner, and K. A. Kozar. A latent semantic analysis methodology for the identification and creation of personas. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 1501–1510, New York, NY, USA, 2008. ACM. 24, 28

[35] S. Mulder and Z. Yaar. *The User is always right: A Practical guide to creating and using personas for the Web*. New Riders Publishing Thousand Oaks, CA, USA, 2006. 19, 58, 59

[36] David S. Platt. *Why Software Sucks ... And What You Can Do About It*. Addison Wesley, 2006. 12

[37] J. Pruitt and T. Adlin. *The persona lifecycle: keeping people in mind throughout product design*. Morgan Kaufmann, 2006. 6, 19, 20

[38] J. Pruitt and J. Grudin. Personas: practice and theory. In *Proceedings of the 2003 conference on Designing for user experiences*, pages 1–15. ACM New York, NY, USA, 2003. 19, 20, 70

[39] Robin Richards. Code breaking. 2, 3

[40] Bruce Schneier. The process of security. Information Security Magazine, April 2000. 7

[41] K. B. Sheehan. Toward a typology of Internet users and online privacy concerns. *The Information Society*, 18(1):21–32, 2002. 4, 6, 13, 15, 74

[42] D.J. Solove. I've got nothing to hide and other misunderstandings of privacy. *San Diego L. Rev.*, 44:745, 2007. 7

[43] A.L. Strauss and J. Corbin. *Basics of qualitative research: Grounded theory procedures and techniques*. Sage Newbury Park, CA, 1990. 20

[44] Colleen Swanson, Ruth Urner, and Edward Lank. Naive security in a wifi world. In *IFIPTM 2010: Fourth International Conference on Trust Management*, 2010. 64

[45] Tara Whalen and Kori M. Inkpen. Gathering evidence: use of visual security cues in web browsers. In *GI '05: Proceedings of Graphics Interface 2005*, pages 137–144. Canadian Human-Computer Communications Society, 2005. 12

[46] A. Whitten and J.D. Tygar. Why Johnny can't encrypt: A usability evaluation of PGP 5.0. In *Proceedings of the 8th USENIX Security Symposium*, pages 169–184. Citeseer, 1999. 10, 11

[47] K.P. Yee. Aligning security and usability. *IEEE Security & Privacy*, pages 48–55, 2004. 1, 10

[48] M.E. Zurko. User-centered security: Stepping up to the grand challenge. In *Proceedings of the 21st Annual Computer Security Applications Conference, Washington*, pages 187–202, 2005. 1, 10