8-2022

Evaluating Privacy Adaptation Presentation Methods to support Social Media Users in their Privacy-Related Decision-Making Process

Moses Namara
Clemson University, mosesn@clemson.edu

Recommended Citation

Namara, Moses, "Evaluating Privacy Adaptation Presentation Methods to support Social Media Users in their Privacy-Related Decision-Making Process" (2022). *All Dissertations*. 3087.
https://tigerprints.clemson.edu/all_dissertations/3087

Evaluating Privacy Adaptation Presentation Methods to support Social Media Users in their Privacy-Related Decision-Making Process

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Human Centered Computing

by
Moses Namara
August 2022

Accepted by:
Dr. Bart Knijnenburg, Committee Chair
Dr. Kelly Caine
Dr. Nathan McNeese
Dr. Brygg Ullmer
Dr. Pamela J. Wisniewski, University of Central Florida

Abstract

Several privacy scholars have advocated for user-tailored privacy (UTP). A privacy-enhancing adaptive privacy approach to help reconcile users' lack of awareness, privacy management skills and motivation to use available platform privacy features with their need for personalized privacy support in alignment with their privacy preferences. The idea behind UTP is to measure users' privacy characteristics and behaviors, use these measurements to create a personalized model of the user's privacy preferences, and then provide adaptive support to the user in navigating and engaging with the available privacy settings—or even implement certain settings automatically on the user's behalf. To this end, most existing work on UTP has focused on the "measurement" and algorithmic "modeling" aspect of UTP, however, with less emphasis on the "adaptation" aspect. More specifically, limited research efforts have been devoted to the exploration of the *presentation* of privacy adaptations that align with user privacy preferences. The concept of "presentation" goes beyond the visual characteristics of the adaptation: it can profoundly impact the required level of engagement with the system and the user's tendency to follow the suggested privacy adaptation.

This dissertation evaluates the potential of three adaptation presentation methods in supporting social media users to make "better" privacy protection decisions. These three adaptation presentation methods include 1) *automation* that involves the automatic application of the privacy settings by the system without user input to alleviate them from having to make frequent privacy decisions; 2) *highlights* that emphasize certain privacy features to guide users to apply the settings themselves in a subtle but useful manner; and 3) *suggestions* that can explicitly inform users about the availability of certain settings that can be applied directly by the user. The first study (Chapter 3) focuses on understanding user perspectives on the different configurations of autonomy and control of the examined three privacy adaptation presentation methods. A second follow-up study (Chapter 4) examines the effectiveness of these adaptation presentation methods in improving user awareness and engagement with available privacy features. Taking into account social media users' privacy decision making process (i.e.,they often make privacy-related decisions), the final study (Chapter 5) assesses the impact of privacy-related affect and message framing (i.e., tone style) on users' privacy decisions in adaptation-supported social media environments. We offer insights (Chapter 6) and provide practical considerations towards the selection and use of “optimal” privacy adaptation methods to provide user-tailored privacy decision support.

Dedication

To

Rehemah, my dear wife and constant uplifter.

My sons Aiden and Kairo, may you grow to achieve all your life's dreams.

My siblings, you all are the source of my inspiration.

My parents, in recognition of your unwavering support towards the pursuit of my life's dreams.

To the reader of this dissertation:

"Education is the great engine of personal development. It is through education that the daughter of a peasant can become a doctor, that the son of a mineworker can become the head of the mine, that a child of farmworkers can become the great president of a great nation. It is what we make out of what we have, not what we are given, that separates one person from another."

—*Nelson Mandela*

Long Walk To Freedom

Acknowledgments

The decision to do my Ph.D. at Clemson University will forever remain one of the best life decisions I could ever make. The journey has been challenging, inspiring, and extremely rewarding[1]. Indeed there is “something in these hills.” I am thankful to the many people who have supported me in this process.

First and foremost, I would like to thank my adviser, Dr. Bart Knijnenburg, for his guidance, mentorship, unwavering support, and a keen interest in my growth as a researcher and individual over the past five years (Jan 2017 - May 2022). He has seen me transform from a young man to husband, father, and now Dr.! I have hugely benefited from his knowledge and wide range of research skills in the process. His “hands-off approach with strategic interventions” advising approach has served me well. It has particularly allowed me to learn how to independently think, conduct research, and know when and how to seek help or, better yet, ask the right questions. Along the way, he has helped me learn how to communicate, write, present, mentor, and lead others. On the lighter side, I would also like to think we had a lot of fun, for I remain indebted (pun intended) for the countless times he helped me correct my placement of apostrophes (e.g., users vs. user’s vs. users’)—I hope he never comes to collect! I will forever be grateful that he took a chance on me.

I would also like to thank my committee members Dr. Kelly Caine, Dr. Nathan McNeese, Dr. Brygg Ullmer, and Dr. Pamela Wisniewski, whose valuable feedback was instrumental in shaping the work in this dissertation. I am sincerely grateful for the opportunities that I had to collaborate with each one of them. Dr. Caine taught me the importance of doing a good thorough literature review. I also greatly benefited from taking her class and reading her book on research methods. Dr. McNeese, gave me the opportunity to learn how to teach. In the process, we formed a special friend-

[1] During my time at Clemson University, I was recognized among 2021’s MIT Tech’s top 30 Under 30 innovators, 2022’s outstanding graduate research student in the College of Engineering, Computing, and Applied Sciences, and Clemson University’s outstanding graduate research student for the year 2022.

ship, collaborated together, and continue to have honest but valuable conversations. Dr.Ullmer, taught me the value of self reflection and how to strive to be a better beneficial member of society. Dr. Wisniewski, I will forever be grateful for the generous feedback on papers and research problems. I learnt alot about the craft of doing research and working through the down-times, especially in circumstances when papers get rejected.

During my time at Clemson, I also had the great fortune of working with many mentees. In particular, I am thankful for Henry Sloan, who worked with me on major parts of this work. He was a high school student when we started working together and is now graduating from college with an undergraduate degree in Computer Science.

I am thankful for all my colleagues within the Clemson Humans & Technology Lab (HAT-Lab): Daricia Wilkinson, Reza Ghaiumy Anaraky, Lijie Guo, Dr. Byron Lowens, Dr. Yang He, and Dr. Paritosh Bahirat. You all never hesitated to lend me a hand when I needed one.

I am also thankful to my parents, who emphasized the importance of education at a very young age. Thank you for supporting me in the attainment of a terminal academic degree. I am incredibly thankful to my father Dr. Namara Suleiman, who firmly believed that I had what it took within me to do a Ph.D. His advice, encouragement, and constant check-ins to ensure that I diligently worked and managed through the ups and downtimes were fundamental to completing this Ph.D. Most importantly, I am thankful for the everlasting gift of an education.

Last but not least, an enormous amount of thanks goes to my family for giving me purpose and strength to power through this Ph.D. journey. In particular, to my beautiful, strong, and supportive wife, Dr. Rehemah Nansamba Namara. You were the main reason I could finish writing this dissertation. You tirelessly pushed me to write, uplifted me when I was down, motivated me when I was out, took me out when I needed a break, fed me when I was hungry, calmed me when I was angry, and loved me when I was hard on myself. For that, I am thankful.

I am grateful for financial support from a Facebook Emerging Scholar Award and a Facebook Fellowship. These fellowships were invaluable in funding and allowed me to fully concentrate on my graduate research work, fund travel, attend and present various portions of this work at different conferences, network, make rent, and attend Clemson University. Any opinions, findings, conclusions, and/or recommendations made in this dissertations are by the author and do not necessarily reflect the views of the funders.

Table of Contents

List of Tables

List of Figures

Prologue

Readers of this dissertation will probably not know, but arriving at this dissertation's final focus/subject was a process. In particular, the focus of this dissertation was made more apparent after the proposal and invaluable feedback from the committee.

In early September 2021, I wrote and presented the dissertation proposal to my committee: Dr. Bart Kninjnenburg, Dr.Kelly Caine, Dr. Nathan McNeese, Dr. Brygg Ullmer, and Dr. Pamela Wisniewski. The dissertation proposal titled *"The Influence of privacy-related user emotion and trust on the adoption, use, and making of privacy decisions on Modern Online technologies*" was primarily focused on the influence of privacy-related affect and trust in the adoption, use, and privacy decision-making process on modern online technologies. In other words, the proposal suggested examining the role that privacy-related affect and trust played in the users' decisions to adopt and use various modern technologies. While the dissertation proposal draft helped fuse most of my research work on *technology adoption* [171, 168] and *privacy* [170, 169] decision-making processes together in one document, the core contribution of the dissertation remained unclear. This was primarily due to the disjoint and lack of an underlying common theme that would connect all the four included studies in the proposal. For example, each study focused on a different modern technology (e.g., Facebook [170, 169], Virtual Private Networks (VPNs) [171], and groupware applications [168]). Therefore, it was not quite clear how they could each be comprehensively studied under one dissertation.

In reviewing the dissertation proposal, the committee members realized this lack of cohesion and focus. For instance, Dr. McNeese correctly asserted that the studies included in the dissertation proposal seemed to be more broadly focused on adoption and use. Thus, I was trying to overlay an emotion and trust framework on top of them. In particular, it seemed like I was force-fitting the studies [171] & [168] into the dissertation yet they did not fit the overall thesis of the dissertation. In light of this, Dr.Wisniewski and the rest of the committee proposed that I consider taking studies [171] & [168] out of the dissertation and instead focus on privacy decision-making on social media (i.e., [170], [169]). This would strengthen the dissertation and in turn the proposed study (Chapter 5) into a more cohesive study .

I followed the recommended advice which helped turn this dissertation into a better product. During the dissertation defense in May 2022, the committee commended me for having the courage to pull out the two afro-mentioned studies, which made for a more focused and meaningful dissertation.

What follows is that dissertation.

Chapter 1

Introduction

As the world reels from a ravaging pandemic of COVID-19 that has affected over 500 million people, claiming over 6 million lives, online privacy remains on people's minds, given their increased reliance on technology to communicate and conduct business during this challenging period. This concern for online privacy is further exacerbated by the lack of universal laws regulating how online systems collect, store and manage user personal information [172]. As such, research examining how technology users can safeguard their online privacy remains critical to facilitate the beneficial use of technologies that they have come to rely on, such as social networking sites.

Over the past decade, online social network sites (SNS)[1] (e.g., Facebook, Instagram, LinkedIn, Twitter, etc.) have experienced tremendous user growth and popularity [44, 38]. These social networks have enabled users to maintain or create new connections, communicate, socialize with friends and family, network and search for career opportunities, read relevant news, express feelings, share thoughts, information, opinions, stories, pictures, and or videos [184, 81, 233]. By their social nature, social media sites emphasize the curation of a personal profile and disclosure of personal information for authentic self-representation, establishment of connections, and interactions with others [38]. Otherwise, the benefits presented by use and functionalities embedded within social media sites become infeasible to achieve [206, 207]. Inevitably, social media users have shared a tremendous amount of personal information ranging from personal contact information to open political or religious beliefs [7]. This proliferation of personal data has led to a broader array of privacy concerns (i.e., about the collection, use or potential misuse, unauthorized secondary use and improper access

[1] Throughout this dissertation I use the terms SNS/social media interchangeably.

to personal information) that have shaped how people adopt and use social media [79, 182, 241]. As a result, social media users have to make a myriad of privacy decisions concerning their information with regards to the management of their social media privacy [4, 176, 177]. In particular, users have to make privacy decisions with regard to *what*, *who*, *to whom* coupled with other system- or purpose-specific factors such as *how much* data collection is justifiable [189, 29, 122]. In making these decisions and engaging different privacy management techniques, social media users are able to relatively assert control over their information privacy and self-disclosure (i.e., determine what information to share/withhold, and control who accesses it), manage their online reputation (i.e., control the way they come across to other), and the access to oneself (i.e., control availability and accessibility others have to them) [180, 241].

However, research has shown that online privacy decisions–like most decisions–are inherently complex due to a number of factors [5, 4, 127]. First, people's mental ability to acquire and analyze all the relevant information is limited—making it "difficult for [social media] users to determine how much of their data may be collected and how it might be used" [5, p.44:2]. Second, people have difficulty picturing the consequences of privacy violations—making it "nearly impossible for [social media users] to fully assess what privacy vulnerabilities they might expose themselves to if they decide to interact with a given [social media]" [5, p.44:2] platform. Third, users' privacy decisions are not always rational (i.e., users do not always weigh the perceived costs against the perceived benefits of information disclosure [8]). Instead, users rely on heuristics such as following other people's privacy decisions [141], default [privacy] settings and framing of information requests [14, 150], their feelings/emotions, among many others [9, 8]—leading to "regrettable decisions that can range from over-sharing to increased exposure to" [5, p.44:2] privacy violations. Fourth, privacy is rarely an end-user's primary task, especially on social media where the goal is to foster relationships between users [7, 233]—leading to "regrettable actions" [226, p.1] from the underestimation of their online activities Given these complexities, social media users report feeling helpless and overwhelmed by the privacy decision-making process required to effectively manage their social media privacy [180, 105, 173].

Cognizant of these complexities involved in the privacy decision-making process and hurdles involved in the management of information disclosures across multiple social contexts [176], social media applications (e.g., Facebook) provide a plethora of privacy features within the platform to enable user achieve their desired privacy [240, 242, 207, 239, 85]. For example, Facebook provides granular audience selection privacy features that contain different audience categories from which users can select and determine who can access their posts/information [182], that it has endeavoured to make easy to access [196]. However, research finds that most users remain unaware of most of these privacy features [92], find them confusing [105], and encounter difficulties in discovering and engaging them in part due to the "inherent uncertainty, and sometimes ambiguity, associated with" [5, p.44:2] applying them. As a result, these privacy features remain underutilized despite efforts to improve users' awareness and subsequently engagement [74, 173].

In recent years, several researchers have investigated ways in which to improve users' awareness, engagement, and utilization of privacy features to reduce the burden associated with privacy decision-making [242, 239, 116, 121, 23, 148, 234]. One prominent approach advocated by privacy scholars to help reconcile users' lack of awareness, privacy management skills, and motivation to use the available privacy controls (i.e., features) is the **User-Tailored Privacy (UTP)** framework [129, 122]. UTP involves the measurement of users' privacy preferences and behaviors, use of those measurements to create a personalized model, and then provision of adaptive privacy decision support to privacy decision-making easier [122]. For example, to make privacy features easier to use, an adaptive privacy-setting interface would involve assessing and modeling user privacy preferences and then tailoring the system's user interface to provide privacy features that match those preferences [121, 239, 148]. Herein, rather than putting the full burden of using privacy features to achieve a desired level of privacy on the user, the idea behind UTP would be to instead measure the user's privacy preferences and behaviors, use these measurements to create a personalized model of the user's privacy preferences, and then provide adaptive support to the user in for instance navigating to the privacy settings—or even automatically implement certain settings automatically on their behalf [121]. This way, the wide variance in users' privacy preferences is taken into account, and they are supported or given the necessary tools they need to decide for themselves on how to meet their desired privacy goals [122].

Most existing work on UTP has covered the "measurement" and "modeling" aspects of the framework [150, 148, 121, 234, 23, 190, 84]. This prior work has been instrumental in the development of personalized algorithmic models to uncover user privacy preferences and recommend appropriate privacy actions [122]. However, while this prior work has identified methods or created personalized models that can be used in the "adaptation" aspect of UTP that involves " tailoring the privacy [features] of the system" [122, p.381] to the user, limited research efforts have been devoted to the ways in which such adaptations can be *presented* to aptly support users in making privacy decisions that align with their preferences [148, 50, 236, 225]. The concept of "presentation" goes beyond the visual characteristics of the adaptation and can have a profound impact on the required level of engagement with the system and the user's tendency to follow the suggested adaptation. For example, while some prior works propose adaptation methods that include fully automating the privacy decision-making process (e.g. [198]), others have implemented adaptive suggestions (e.g. [148], or suggested the use of personalized nudges (e.g. [227]) or interface adaptations (e.g. [236]). Thus, a comprehensive understanding of the presentation of the ''optimal" adaptation method is essential if systems are to help users meaningfully engage with the available privacy features without overwhelming or misleading them. Furthermore, since these adaptation methods would vary in the autonomy and control they afford users in the privacy decision-making process, they could lead to different user reactions, user engagement patterns with the available privacy features, privacy protection outcomes and trust levels in the social media platform. To this end, my dissertation focuses on understanding user perspectives on the different configurations of autonomy and control of the examined varying privacy adaptation presentation methodologies. More specifically, I examine the effectiveness of these adaptation presentation methods in improving user awareness and engagement with available privacy features, overall privacy protection outcomes, and level of trust in social network sites. Furthermore, I assess the potential impact of privacy-related affect and tone in providing adaptation-supported privacy decision help. Fundamentally, I seek to answer the following research question:

> **What is the *"optimal"* adaptation presentation method that can be used to support users and alleviate their privacy decision-making burden on social media?**

Researchers assert that to alleviate the user burden inherent with privacy decision-making, it is essential to proactively strike a personalized balance between users' desire for privacy and their need for online interaction without overwhelming them with privacy features [239]. Nonetheless, social media sites like Facebook have a plethora of privacy features, making it impossible for users to engage and apply all the features in line with their privacy preferences [239, 92, 150]. Depending on the privacy feature, the possible privacy adaptation presentation method implemented is likely to vary in the level of autonomy and control it affords to users in the privacy decision-making process [203]. However, this reduction in user burden can also reduce users' perception of control, increasing their anxiety and ultimately decreasing the acceptance of the privacy adaptations [216, 30]. Therefore, a necessary first step in assessing for an "optimal" privacy adaptation presentation method involves understanding *which* privacy features can be tailored/adapted to the users' privacy preferences and *how* to implement such adaptations.

I start my inquiry in **Chapter 3**, by examining user preferences for privacy adaptation presentation methods used to adapt 19 Facebook privacy features. More specifically, I examined user preference to the three increasingly autonomous privacy adaptation methods: 1) *suggestions* explicitly inform users about the availability of certain settings that can then be applied directly by the user; 2) *highlights* emphasize certain privacy features to guide users to apply the settings themselves; and 3) *automation* involves the automatic application of the privacy settings by the system without user input [170]. By doing so, we were able to understand how users would respond to the different possible privacy adaptation implementations of the stated privacy features, and under what terms each of the method was preferred or undesired. We found that the "optimal" (i.e., preferred) adaptation method depended on the users' familiarity with the privacy feature and how they used it, and their judgment of the awkwardness and irreversibility of the implemented privacy functionality. Participants generally disliked the full *automation* method, except for privacy features they used frequently and perceived as inconsequential, where it could alleviate some of the behavioral onus and effort of managing one's privacy. The *highlight* method was appreciated for its ability to unobtrusively raise users' awareness about a privacy feature and was thus most suitable for features users only used occasionally. Finally, the *suggestion* method was preferred as a means to teach users privacy features they were unfamiliar with, unless this resulted in awkward suggestions of behaviors with negative social connotations. In summary, we found that different Facebook users were (un)familiar with different features, and thus each preferred adaptation method for each feature differed per user. Based on these findings, we recommended that the adaptation method itself be tailored to the user as well.

Based on the findings and recommendation mentioned above, in **Chapter 4**, I conducted a follow-up experimental study (N = 406) to examine the effectiveness of the different adaptation methods—*automation*, *highlights*, and *suggestions*—in improving users' engagement with the available privacy features and overall privacy protection outcomes. In particular, we tested three proposed *"adaptation methods"* (automation, suggestions, highlights) in an online between-subjects user experiment in which 406 participants used a carefully controlled SNS prototype. We systematically evaluated the effect of these adaptation methods on participants' engagement with the privacy features, their tendency to set stricter settings (protection), and their subjective evaluation of the assigned adaptation method. We found that the *automation* of privacy features afforded users the most privacy protection, while giving privacy *suggestions* caused the highest level of engagement with the features and the highest subjective ratings (as long as awkward suggestions are avoided).

The works in **Chapters 3 & 4**, revealed privacy suggestions as the preferred adaptation presentation method that could be used to enhance and support users' in their privacy decision-making process. However, to improve the acceptance and effectiveness of such privacy suggestions in enhancing users' engagement with privacy features, prior work asserts that they be designed to align with how users make privacy-related decisions [73, 4]. In making privacy-related decisions, research shows that users rely more on heuristic (rather than analytic, systematic) processing of the conveyed information [73, 4, 6]. These decision heuristics are susceptible to factors such as user affect (i.e., how users feel) [142, 43]. Furthermore, in presenting recommended privacy actions, privacy scholars also call for the careful consideration of the framing and structuring of privacy information (i.e., the way information is presented to the user) [50].

Taking into account the recommendations from this prior work, in **Chapter 5**, I conducted an experimental study (N = 750) to systematically understand the unique impact of privacy-related affect and what framing (i.e., tone style) privacy suggestions should embody if they are to more effectively encourage users "better" manage their social media privacy. The primary goal was to examine the privacy suggestion tone style that would "better" encourage users to engage with privacy features to achieve their desired privacy, considerate of users' feelings about social media privacy (i.e., privacy-related affect). Furthermore, I wanted to assess the impact on users' experience with the platform (i.e., perceived decision helpfulness and trust in the platform) based on the engagement with the provided privacy suggestions. I manipulated the privacy suggestion tones and used priming to put participants into various (privacy-)affective states to evaluate the most appropriate framing (i.e., tone) for each and ultimate privacy decision outcomes. I found that the examined three varying privacy suggestion tone styles (i.e.,*neutral*, *passive*, *assertive*) indeed influence users' privacy decision outcomes. However, the nature of the effect significantly differs based on users' pre-existing privacy-related induced affective states (or lack thereof), i.e., the mood a user is in before the actual privacy protection decisive situation occurs. For instance, when users are in a positive privacy-related affective state, the neutral tone tends to work better at encouraging them to make “better” privacy decisions. In contrast, the assertive tone tended to work best when users were in a negative privacy-related affective state. Furthermore, we observe that the tone that privacy suggestions embody not only influences users’ behavior regarding these suggestions and/or the privacy actions they recommend; they also impact users’ other privacy actions (i.e., actions that are not subject to suggestions by the platform), indicating that tone has a robust, system-wide effect. These findings suggest that considering users’ privacy-related affect (i.e., how users feel about their social media privacy) is crucial in determining the tone style that system designers can use to craft and present personalized privacy suggestions.

In Chapter 6, we conclude with a discussion of the contributions of this dissertation and possibilities for future work.

Chapter 2

Related Work and Motivation

As more users have become accustomed to the use of social media sites (SNS) such as Facebook, the frequency with which they have had to make privacy decisions has also risen. Social media privacy is rooted in the work of Westin [232], who defined *privacy* as individual control over the disclosure and subsequent uses of personal information. As such, privacy scholars conceptualize social media privacy as a boundary regulation mechanism "where [social media] users seek to strike a balance between being too open or disclosing too much versus being too inaccessible or disclosing too little" [180, p.124]. I adopt this conceptualization of privacy for the work in this dissertation, and interchangeably use the terms "social media privacy" or "privacy".

In this chapter, I review related work corresponding to the origins, use and proliferation of social media (see **Section 2.1**), highlight the proliferation and rise in use of social media (see **Section 2.2**), the privacy challenges related to social media use (see **Section 2.3**), privacy decision-making process(es) on social media—where users have to consider making a trade-off between the possible privacy threat of information disclosure/data collection and the benefits that might accrue (see **Section 2.4**), prevalent social media privacy protection behaviors and decision-making strategies (see **Section 2.5**), the origins and application of user-tailored privacy—as a concept to alleviate the user burden inherent with privacy decision-making, and the presentation methods (i.e., "adaptation methods") of personalized privacy adaptations (see **Sections 2.6 & 2.7** respectively).

2.1 A Brief History of Online Social Network Sites

In his 1929 short story called “Chain-Links” [106], Frigyes Karinthy pondered about a well-connected mind-game that he constantly found himself playing not only with human beings, but with objects as well. He wrote that *”the strange mind-game that clatters in me all the time goes like this: how can I link,with three, four, or at most five links of the chain, trivial, everyday things of life. How can I link one phenomenon to another? How can I join the relative and the ephemeral with steady, permanent things - how can I tie up the part with the whole?”* [106, p.3]. The game was premised on the fact that planet earth was small, and thus *”anyone on Earth, at my or anyone’s will, can now learn in just a few minutes what I think or do, and what I want or what I would like to do.”* [106, p.1] Unbeknownst to Karinthy, it would not be until 6 decades later, that the human separation concept (i.e., six degrees of separation) which asserts that humans are connected to each other through a series of chains of acquaintances less than six connections away from the other, would first be implemented online.

The six degree separation theorem, popularized by a 1990’s eponymous play *Six Degrees Separation* by John Guare [80], was the first to imagine a way in which people from different spheres of the globe could be interconnnected. In the winter of 1997, Andrew Weinreich, touted the idea of online social networks which could connect the world within a single network [38, 138]. Inspired by the six degrees theorem, Weinreich believed that with a free, web-based networking service, people could volunteer information about their interests, jobs, and connections, which would make it easy to index their relationships in a single place [138]. Weinreich subsequently launched SixDegrees.com—less than 10 years after the invention of the internet [138]. SixDegrees allowed users to create profiles about themselves, list their friends, surf their friends list, connect and send messages [38]. However, SixDegrees was short-lived—Weinreich believes that it was simply ahead of its time, given the state of technology at the time, that could not nurture the kind of connectivity needed for a social network to thrive [138]. Nevertheless, SixDegrees proved the concept of social networking and pioneered several aspects that would come to be part of virtually all social networking sites (SNS) today [38]. The web-based social networking model it conceptualized was adopted and modified by other influential SNS, most re-markedly Facebook[1] [201, 138, 38].

[1] Refer to [38] for a brief synopsis of other popular SNS in the early 2000’s

Founded by Mark Elliot Zuckerberg in February 2004, Facebook launched as a social networking site meant to foster online connections, specifically among students (e.g., Havard-only SNS) [138]. By 2006, Facebook had gradually expanded to foster connections among all internet users [38]. Facebook allowed users to curate profiles, share photos and videos, and make connections with other people online. Additionally, it allowed other outside third-party developers to build applications that would run on top of it. These features, coupled with how it managed user privacy, helped differentiate it from other social networking sites at the time [38, 37]. Over the years, Facebook became very popular and grew drastically as a social networking site where new relationships and connections could be forged and old one's easily maintained. In Zuckerberg's own words, Facebook grew to become a "powerful new tool [that people use] to stay connected to the people they love, make their voices heard, and build communities and businesses" [246, p.8]. In doing so, Facebook might have helped realize Karinthy's "chain-link" vision and surpass SixDegrees' much earlier attempt at creating a successful social networking service. Today, 72% of the U.S public reports using some type of social networking site (SNS) [45], with about 69% of the U.S adults actively using Facebook [44]. Facebook remains the most used online social network worldwide with roughly 1.91 billion daily active users[2] [164].

Based on Facebook's popularity and tremendous effect on online social networking, interactions and communication, for all the studies in this dissertation (i.e., Chapters 3, 4 & 5), I leverage it as a case example of a social media platform to empirically examine the feasibility of *"adaptation methods"*—ways in which predicted privacy behaviors (i.e., privacy adaptations) can be presented to the user. Henceforth, I interchangeably use the term "social media" or "social networking site" with specific reference to Facebook.

2.2 The Use and Proliferation of Online Social Network Sites

The rise of social networking sites (SNS) fundamentally shifted how people organize, interact and socialize online [207]. People could easily form new online connections that otherwise would not have been possible without the emergence of SNS. According to Boyd and Ellison, social network sites (SNS) were different from other forms of computer-mediated communications based on three main characteristics that they afforded users. The ability to: "(1) construct a public or semi-

[2] As of June 2021

public profile within a bounded system; (2) articulate a list of other users with whom they share a connection; and (3) view and traverse their list of connections and those made by others within the system." [38, p.211]. These characteristics allowed users to articulate, control and make visible their social networks. For example, Facebook afforded users the ability to curate a profile, control and customize the access to how and who they allow to access their personal information [53]. Herein, each user profile differed based on who was in the users' network, demographics, interests, and the information shared. Most importantly, users had the express ability to control how their personal information was shared or accessed [150].

These characteristics allowed online social networking, especially Facebook, to evolve from a niche phenomenon to mass adoption. Earlier studies on the use and proliferation of Facebook highlighted the reasons behind the high rates of acceptance among specific demographics of the populace [83, 78, 133]. More specifically, in a 2005 survey of students at Carnegie Mellon University, Gross and Acquisti found that 74.6% of the undergraduate students had a Facebook profile [78]. In a 2007 survey of first-year college students, Hargittai et al. [83] found that 80% of the students reported using Facebook, specifically to enunciate their offline relationships, build new relationships and connect with other people in their existing network. Similarly, Lampe et al. [133] in their study on changes in use and perception of Facebook among undergraduate college students from 2006-2008, found that reasons for using the site remained relatively constant over time: most students used the site to primarily connect. Connecting using Facebook involved using the site to search for other people to date (i.e., "social browsing"), check out people met offline/socially (i.e., "social searching"), learn more about other people in the same class or living near (i.e., "social searching"), and keep in touch with old friends (i.e., "keep in touch") [133]. Ultimately, Facebook drastically grew from about 100 million users in August 2008 to over two billion users worldwide today [3] [164, 44].

However, in their work, Lampe et al. [133] also found that while users were positive about the site in its early years (2006-2007), the way they perceived the audience for their user profiles and general attitudes about the site differed over time (2006-2008) [133]. What could possibly explain this possible change in user perception and attitude overtime? While Lampe et al. [133] initially postulated that this difference could be because users had maxed out the utility for being present on the site, several researchers [7, 78, 88] alluded to the prevalent over sharing pattern and privacy implications that early Facebook users seemed "oblivious, unconcerned, or just pragmatic

[3] As of May 2022

about" [78, p.78].

More specifically, Hew [88] found that students on Facebook tended to disclose more personal information about themselves, potentially attracting unknown privacy risks about themselves. Gross and Acquisti [78] indicated that based on the richness and amount of the personal information users disclosed on their Facebook profiles, coupled with the visibility, public linkage to the members' real identities, and scope of the network, users were putting themselves at risk for a a variety of attacks on both their physical and online persona. For example, the authors pointed out that depending on the accuracy of the publicly disclosed information (e.g., hometown, current residence, phone number), it would not be too difficult to reconstruct one's social security number or steal their identity [78]. Additionally, the blur of public versus private sharing contexts (i.e.,"context collapse") inherent within SNS made it more complex to manage personal information disclosure across multiple contexts (e.g.,"once a "friend" has been added to one's network, maintaining appropriate levels of social interactions in light of one's relationship context with this individual (and the many others within one's network) became even more problematic" [180, p.119]) [207]. Gradually, as Facebook grew, there was an increase in privacy concerns, change in perception, attitudes and use patterns [7, 53, 207].

2.3 Privacy Challenges on Online Social Network Sites

The inherent nature of social media involves the disclosure of personal information. This personal information is a basic requirement to help link and identify the social profile and shared information to a physical person in the real world [13, 78]. However, the disclosure of personal information presents privacy challenges and raises privacy concerns related to its collection, storage and access [61, 40]. While the collection and use of such information helps in the provision of numerous benefits for users (e.g., formation of new/maintaining old connections, new product/service recommendations, social support, influencing others, reputation, enjoyment, etc.), they worry that their personal information can easily be misused or transferred to other third-party entities without their express permission or knowledge [179, 7, 18, 20]. As such, social media users express high levels of privacy concerns about the collection and use of their information [44, 20, 179]. Of most concern is the fact that all the personal actions users perform via their social media profiles can easily be linked back to them (i.e., based on their real-world identities) [13, 78]. Thus, privacy violations that could occur on social media, are most likely to result into regret and real world consequences for users [226, 73, 205]. As a case example, to highlight the prevalent privacy challenges associated with SNSs, I provide a brief genesis of some of the privacy challenges peculiar to Facebook, one of the most prominent social networking sites today [44].

Launched in 2004, Facebook began as a network-centric SNS where users' profiles and content were only visible to members within the same network [37]. Users joined using a valid email addresses associated with their respective institutions (e.g., a university, high school, workplaces, etc) [38]. This requirement helped keep the site relatively closed and contributed to "users' perception of the site as an intimate, private community" [38, p.218], where only peers and close online connections were perceived to be the "audience" rather than strangers or casual acquaintances on the site [133]. Hence, in creating and curating their personal Facebook profiles, users were encouraged to provide their personal details such as address, telephone number, photograph, interests and other details to foster these connections. More specifically, users were encouraged to label the people they already knew on the site as "Friends" or send out email invites to others who were not on the site [38]. As a result, users mostly used the site to find other people to date, meet new people, and learn about people living near them [133].

In 2006, as Facebook grew and expanded to everyone on the internet, the network-centric approach could not scale well, and thus the requirement for users to join close-knit networks was rather de-emphasized [37]. Instead, a "Newsfeed" that aggregated and provided updates about new profile changes of friends was launched [133, 88]. New privacy features were also unveiled to enable users have greater control over their audience and access to their personal information on the site. In particular, users were provided with features to determine what could be shared with whom, using audience control categories such as ("No one", "Friends", "Friends-of-Friends", or a particular "Network"—which later evolved to be more granular and became "Everyone","All Users", or "Public") [37, 180]. Additionally, users were provided with the options to have other people become their "friends" or "followers" [180]. However, the default settings of these controls were always set to enable sharing broadly [37]. These new constant platform changes and practices presented new information privacy challenges that increasingly contributed to users change in attitude about the platform [133]. Hence, while Facebook users seemed oblivious or unconcerned about these privacy issues at the onset of using Facebook [78], overtime, they exhibited heightened privacy concerns, especially related to the third party access to their information [179]. These privacy concerns were made more apparent by the sustained popular coverage of SNS with regards to user privacy [38, 86].

The rapid development of Facebook also evolved users' privacy norms [207]. For example, in 2010, Zuckerberg believed that the new privacy norm had evolved to the point where "people [had] gotten comfortable not only sharing more information and different kinds but more openly and with more people." [115]. However, contrary to that belief, Facebook users at the time were found to instead "exhibit increasingly privacy-seeking behavior and progressively decreasing the amount of personal data they shared publicly." [207, p.1]. In their study on privacy trends among Facebook users between 2010-2011, Dey et al. [53] found that users in their study had dramatically become more private (i.e., 17.2% and 52.6% users had reported hiding their friend lists between March 2010 and June 2011 respectively). Similarly, in their study on Facebook users' perceptions on privacy, O'Brien, and Torres [179], found that 75% of the users at the time reported changing their privacy settings towards tighter controls, largely prompted by privacy concerns related to third-party access to their personal information. In another study on Facebook users' perceptions on privacy across five years (i.e., 2010-2015), Tsay-Vogel et al. [214] found that after 2012, users harbored higher levels of privacy concerns that would heighten if users were exposed to incidents of privacy violation.

In 2018, an extreme user privacy violation on Facebook did occur. Personally identifiable information of more than 87 million users was illegally accessed by an external data analytics firm Cambridge Analytica [98, 204]. Following the Cambridge Analytica privacy violation incident, reports found that users exhibited a more negative perception of Facebook and greater privacy concerns [89, 87]. However, few users went ahead to delete their Facebook accounts or even update their privacy settings [89]. A pew research survey of U.S Facebook users found that only 54% of them had at least adjusted their privacy settings, 42% had taken a break from the site, and only 25% had deleted the application from their phone [186]. Otherwise, users seemed reluctant to even change their privacy settings purportedly due to the endless data breaches and updates on Facebook [89]. Nevertheless, due to the gravity of the Cambridge Analytica privacy breach, Zuckerburg was compelled to testify in front of a joint Congress hearing in April 2018 [246]. Here, senators quizzed him about the seemingly negligent privacy practices of Facebook. More specifically, Hon. John Thun, senator from South Dakota, tasked him with the responsibility to unveil "without delay about what Facebook and other companies plan to do to take greater responsibility for what happens on their platforms. How [he would] protect users' data? How [he would] inform users about the changes that he was making? And how [he] intend[ed] to proactively stop harmful conduct [from happening] instead of being forced to respond to it months or years later?" [246, p.3]. Additionally, Hon. Chuck Grassley, senator from Iowa, expressed deep concern that "consumers may not fully understand or appreciate the extent to which their data is collected, protected, transferred, used and misused." [246, p.6]. These sharp remarks clearly highlighted the persistent privacy challenges all-pervasive on SNS, especially Facebook.

Several researchers have tried to examine the various privacy attitudes, perceptions and behaviors of SNS users to uncover ways in which some of these privacy challenges can be alleviated [180, 7, 181, 78, 95]. On Facebook, we learn that the common privacy challenges involve: (1) insufficient control of information/self disclosure—choosing what kind of information to share with other people on the site); (2) context collapse and problems related to imagined audiences—the blur between public versus private sharing contexts make it difficult to manage online relationships given the existence of multiple and different connections on the site ; (3) appropriate reputation management—controlling how to present oneself to different groups of people on the site to avoid regret, (4) access to onself—controlling access that others have to a user on the site; and (5) privacy loss due to third party access to information [180, 44].

2.4 Privacy decision-making on Online Social Network Sites

On SNS, privacy decision-making is a burdensome and complex task that users have to perform to garner the social benefits related with the use of social media [13, 7]. These privacy decisions are typically related—but not limited to—boundary management where users seek to control the visibility/access of their data or information [180]. More specifically, Alemany et al. [13] assert that the privacy decision-making process on social media "is [typically] composed of the impulse to share something, the choice of channel [(i.e., what social network to use as well as communication within the network)], the composition of the message, the choice of receivers, and the feedback assessment." Along this process, there are varying potential privacy risks (e.g., unauthorized access, relationship breaks, context collapse, stalking and identify theft, misuse of personal information, etc.) that users must consider before performing an action [13]. Thus, privacy decisions are performed to limit the potential costs/risks related with performing certain actions on the social media platform [240, 78]. Majority of these privacy decisions are made using a set of predefined interface privacy controls/features to take action (i.e., whether to accept/reject/ choose from a category of privacy options to indicate a preferred privacy level) [13, 240, 242]

However, privacy decisions–like most decisions–are inherently complex because they have "delayed and uncertain repercussions that are difficult to trade-off with the possible immediate gratification of information disclosure" [127, p.2]. Studies that have tried to examine, explain, and predict how individuals make privacy decisions assert that these decisions are carefully considered by way of conscious-analytic and profit-loss calculations [28, 178]. In other words, privacy decision-making is a rational process where decisions are carefully considered based on the risk and benefits related with information disclosure [8, 2, 9, 28]. Within the privacy field, this principle is encoded as the "privacy calculus" [135]. The theory postulates that in making decisions about online disclosure of personal information, people go through a calculation process in which expected loss of privacy is carefully weighed against potential gains from that behavior [11]. However, Acquisti et al. [4] argue that contrary to this traditional theory, the authors suggest that privacy decision-making is influenced by several factors that affect how people make privacy-sensitive decisions [5, 4]. Some of these factors include incomplete information to make appropriate privacy decisions, the inherent purpose of social media use, and the bounded ability to make judgments about uncertain events (e.g., privacy violations) [4]. Social media users' have incomplete information about how their disclosed information or data could be used (e.g. if it will be shared with a third-party) or how disclosed information or collected data could be used. Even when such information is provided (e.g., within privacy policies), research shows that users seldom read it [161]. Nonetheless, even if users wanted to read the provided privacy policy information, it would take them an estimated 54 billion hours/year to read, and in turn, cost the American economy over three quarters of a trillion dollars [161]. As a result, users blindly accept or ignore such information, making it difficult for them to make informed privacy decisions related to information disclosure [10]. Furthermore, the primary end-user task on social media is not to "manage privacy" but instead to connect with others [7, 233]. As such, when the privacy implications of the online activities undertaken are underestimated, "regrettable actions" can occur [226]. Users are also usually uncertain about the privacy risks associated with disclosure or misuse of their personal information [4]. Instead "users perceive privacy risks as an abstract problem that is psychologically distant and more related to the distant future" [13, p.23]. Such judgements require considerable cognitive effort and information [4]. Thus, rather make rational calculated privacy decisions based on the costs versus benefits of disclosure, some users lean on heuristics or take shortcuts [4, 5]. For example, in their work examining why people make posting decisions, Ferwerda et al [65] found that when social media users' are uncertain about some decisions (e.g., whether to share a post), they resort to heuristic such as deciding to self-censor (i.e., not share at all). Anaraky et al [73] found such behavior can be observed more among younger adults (rather than older adults) who tend to make heuristic decisions rather rational calculated decisions. Some privacy decisions such as–managing access to self (e.g., tagging or friend requests)—are also more liable to heuristics rather than rational decision making based on the default and framing of the privacy choice [14].

Therefore, in the process of engaging with the available privacy features to make privacy-related decisions, users are faced with a huge burden with regards how to apply them to make decisions that align with their privacy preferences [30, 239]. Consequently, users avoid the hassle of exploiting and using the available privacy features [92, 173]. Only users with a sufficiently strong motivation in pursuance of their privacy protection eventually make changes to their privacy settings different from the set defaults [128]. Furthermore, these few motivated users are often met with an overload of privacy features, and with privacy instructions that are hard to read and comprehend [128]. Thus, despite the noble efforts towards alleviating privacy concerns through the provision of privacy features/controls, users remain unaware about these privacy features, find them confusing, and encounter difficulties in discovering and engaging them [92, 7, 105, 173]. As a result, these privacy features remain underutilized despite efforts to improve users' awareness and subsequently engagement [74, 173].

2.5 Overview of Facebook Users' Privacy Feature Engagement & Use Patterns

Facebook users use the platform to communicate and socialize with friends and family, network and search for career opportunities, share thoughts, relevant news, feelings, emotions, news, stories, pictures and videos of various life events [184, 81]. To successfully support all these use cases, Facebook offers users a number of privacy features to control how they interact and share information with each other [181]. These privacy features are supposed to help users set their desired level of privacy in sufficient detail. However, user awareness of these privacy features remains low

[92, 242], and most users end up not using the available privacy features [7, 240, 242, 92].

For instance, in their study on user awareness of News Feed controls on Facebook, Hsu et al.[92] found that 49% of Facebook users were not aware of the existence of many of the features that they could use to personalize their feed. Even when users had the desire to use the existing controls, they struggled to find them. The authors also found that there was a misalignment between users' expectations and the actual functionality provided by the features [92]. Liu et al. [150] also found that available privacy features matched users' expectations only 37% of the time, and that incorrect expectations almost always meant that users underestimated the extent to which their information was exposed to others. As a result, they estimated that about 36% of all content on Facebook is posted with a privacy setting that shares it to more people than expected. This lack of awareness and misalignment of privacy features has important ramifications for both new and experienced Facebook users [226].

2.6 A Self-Adaptive Privacy Approach: User Tailored Privacy (UTP)

Several privacy scholars have investigated ways in which to alleviate the burden associated with privacy decision-making [122, 13]. In particular, researchers have sought to examine how a system's privacy setting can be tailored to a level of privacy that each individual user is most comfortable with, so that rather than putting the full burden of managing these settings on the user, the privacy decision making process is more personalized based on the users' privacy preferences and behaviors [122, 121, 119, 227]. As such, one key concept that has been suggested to achieve such a user-centric solution is **User-Tailored Privacy** (UTP) [122, 121, 119]. The idea behind UTP (see Figure 2.1) is to measure users' privacy preferences and behaviors, use these measurements to create a personalized model of the user's privacy preferences, and then provide adaptive support to the user in navigating the available privacy settings—or even automatically implement certain settings automatically on the user's behalf [121]. Researchers assert that, by proactively striking a personalized balance between users' desire for privacy and their need for online interaction, such *adaptive approaches* could alleviate the privacy decision-making burden and help users achieve the privacy they want without overwhelming them with privacy features [239].

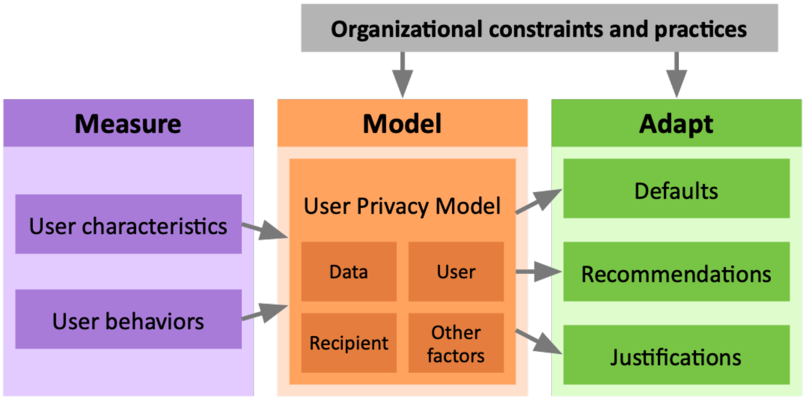


Figure 2.1: A schematic overview of User-Tailored Privacy (Adapted from [119])

Consequently, a growing body of research has focused on UTP's application. However, majority of this reseach work has focused on the "measurement" and development of personalized "models" that align with users' privacy preferences [4] [158, 218, 150, 213]. For instance, on the measurement aspect of UTP, prior work has focused on detection and categorization of personal information shared on social media and its privacy sensitivity. For example, Hirschprung et al. [90] suggested a method for estimating people's privacy preferences albeit in financial terms. Mao et al. [158] analyzed tweets to build a classifier that could determine the privacy sensitivity of the information that users' disclosed, the privacy threats posed by the revealed information, what kind of users leaked information and how they leaked it. They found that based on users' tweets, private information (e.g., *who*, *where*, and *what* time a person would go for vacation) and topics (e.g. sexuality, expressed emotions, confessions, bodily harm, illegal activities, and disrespectful behaviors) could easily be detected and categorized. The authors built a machine learning (ML) based classifier with an accuracy of 78% to automatically detect such private information and categorize topics. Vanetti et al. [218] also built a rule-based classifier that analyzed shared textual information on users' timelines to distinguish between personal and non-personal information. The resultant classifier was 80% accurate and could help predict violent, vulgar, offensive, hateful and sexual textual information. Similary, Wang et al. [224] also built a content-based classifier to help classify sensitive tweets into 13 pre-defined topic categories, so as to help users develop privacy protection mechanisms that align with their privacy preferences.

Prior work has also focused on personalized model to help users manage their privacy deci-

[4] Refer to [13] [199] & [122] for an extended overview on these aspects

sions. For example, Liu et al.[150], analyzed privacy preferences and permission-granting behaviors of 4.8 million Android users. They found that although people's mobile app privacy preferences are diverse, a small number of profiles could actually be identified to simplify their privacy-decision making processes. Similarly, Wijesekera et al. [234] built a classifier that could make mobile app permission decisions on the user's behalf by detecting a change in their context, and when necessary, inferring user privacy preferences based on their past decisions and behaviors. The resultant classifier accurately predicted users' privacy decisions 96.8% of the time.

Studies in the context of location-sharing applications have also developed personalized models that align privacy settings with users' privacy preferences. For instance, Toch et al [213] found that people who tend to visit a wider variety of places tended to be subjected to a greater number of requests for their locations. However, users were only comfortable granting permission if the location was typically visited by a large and diverse set of people. Benisch et al [30] found that privacy-setting schemes were more accurate at capturing users' location sharing preferences if they were dependent on both time and location. Ravichandran et al [190] found that decision-tree and clustering algorithms could be used to provide users with a small number of basic default policies to choose from to alleviate the burden involved in sharing locations with location-based apps and abstract away user-specific elements (e.g., a user's default schedule or canonical places such as "work" and "home").

A series of studies in the broader context of the Internet of Things built similar user models clustering users' privacy decisions into a number of privacy profiles [23, 84, 195]. For instance, Bahirat et al. [23], developed a set of three "smart" default profiles that captured users' preferences towards sharing data with public IoT systems. He et al. [84] used a similar approach to predict users' smarthome privacy preferences with five profiles, and Sanchez et al. developed a four-tier profile-based system to predict users' privacy preferences in the context of wearable fitness trackers. In each case, the profile-based solution was able to capture users' preferences with an accuracy of around 82-85%.

In the context of social networks, Fang and LeFevre [64] used a similar profile-based approach in the development of a privacy wizard that automatically assigns privileges to a user's Facebook friends. The evaluation of the wizard with privacy preference data collected from 45 real Facebook users revealed that the it could generate highly accurate settings to automatically assign to a user's friends with minimal user input. Yang et al. [244] proposed a utility-based trade-off framework that modelled and quantified users' adaptive sharing requirements as utility of potential privacy risks and social benefit, to automatically recommend users a subset of a select sharing circle each time they would initiate an information-sharinng action. Similarly, Vidyalakshmi [220] built a model for calculating a privacy score metric based on users' personal attitudes toward privacy and communication information.

While this prior work has identified methods to detect or classify privacy sensitive information, and create personalized models that can be used to adapt a system's privacy settings to the user's preferences, limited research has focused on the design and presentation of these adaptations [148, 50, 236, 225]. This work is important, since the optimal "adaptation method" can help users to meaningfully engage with the available privacy features without overwhelming or misleading them. My dissertation work seeks to address this gap by examining user preferences, effectiveness, and framing of various adaptation methods that can be used to present privacy adaptations to users.

2.7 Presentation of Privacy Adaptations

Limited research effort has been devoted to the exploration of the *presentation* of privacy adaptations that align with user privacy preferences. The concept of "presentation" goes beyond the visual characteristics of the adaptation and can have a profound impact on the required level of engagement with the system and the user's tendency to follow the suggested adaptation. For example, while some propose to fully automate the privacy decision-making process (e.g. [198, 218]), others have implemented adaptive suggestions (e.g. [148], or suggested the use of personalized nudges (e.g. [227]) or interface adaptations (e.g. [236]).

Vanetti et al. [218] stressed the importance of examining the usability of an interface tool that would automatically recommend privacy trust values for contacts users did not know personally (based on those users' actions, behaviors, and reputation in the SNS).

Liu et al. [148] found that mobile app permission setting suggestions based on user privacy preferences were perceived to be helpful and largely adopted by users. Most importantly, the suggestions increased user engagement with the privacy settings.

Warberg et al. [227] reaffirmed the importance of examining the possibilities of tailoring privacy nudges to align with individual differences in decision making and personality, especially among large organizations such as SNS that typically have a large number of users.

Wilkinson et al. [236] recognized that the privacy features on social networks are often more than one click away, and explored the idea of adapting the social network User Interface (UI) in such a way that it increases the salience of those features that fit the user's personal privacy management strategy (cf. [242]).

While this existing work has explored different methods of adaptively assisting users with their privacy management practices, few researchers have examined user preferences for the methods as well as compared the various adaptation method in terms of their effectiveness at enhancing user engagement and overall privacy protection [50]. My dissertation work seeks to address this gap.

Chapter 3

Exploring Adaptation Methods To Better Support Privacy Decision-Making

An SNS like Facebook provides its users with a plethora of privacy features to enable them control and manage their privacy [240, 242]. Several researchers find that despite the availability of these features, users remain unaware about these privacy features, find them confusing, and encounter difficulties in discovering and engaging them [92, 7, 105, 173]. As a result, these privacy features remain underutilized inspite of the noble efforts towards improving user awareness and subsequent engagement with the features [74, 173]. Hence, in this chapter [1], I seek to answer what can be done—if anything– to improve user engagement with privacy features in a way that aligns with users' privacy preferences.

More specifically, I present a study on the potential presentation methods (i.e., *"adaptation methods"*) that can be used to improve privacy protection and management on modern online technologies such as social networks like Facebook. The study was motivated by the feasibility of successfully implementing User-Tailored Privacy (UTP) features [119]. UTP is a privacy-enhancing adaptive approach used to measure users' privacy preferences and behaviors, use these measurements to create a personalized model of the user's privacy preferences, and then provide adaptive support

[1] Published as [170]

to the user in navigating the available privacy settings—or even automatically implement certain settings automatically on the user's behalf [121]. Several privacy scholars have advocated for the use of UTP as an approach to improve users' awareness of, and engagement with, privacy features [121, 119, 227]. As a result, a growing body of research has focused on the algorithmic development and implementation of personalized models that can be used to adapt a system's privacy settings to the users' privacy preferences [150, 234, 190, 213, 30]. However, limited research has focused on the design and presentation of these adaptations [148, 50, 236, 225]. This study sought to address this gap by exploring the user reactions, perceptions and feasibility of three adaptation methods (*Automation*, *Highlight*, and *Suggestion*) that vary in the level of autonomy and control they afford to users in the privacy decision-making process.

For this study, we developed adaptive versions of 19 Facebook privacy features, and for each user-tailored feature, we tested users' reaction to the features adapted using the three adaptation methods that can be used to implement the suggested behavior (i.e., privacy adaptation). The three adaptation methods were: 1) *Automation* —involves the automatic application of the privacy settings by the system without user input, 2) *Highlights*—emphasize certain privacy features to guide users to apply the settings themselves; and 3) *Suggestion*—explicitly informs users about the availability of certain settings that can then be applied directly by user. We found that for users, amongst these three adaptation methods, the optimal adaptation method depended on the their familiarity with the privacy feature, how they use them, and their judgment of the awkwardness and irreversibility of the implemented privacy functionality. Based on our findings, we provide design recommendations for the implementation of user-tailored privacy on modern online technologies such as social network sites like Facebook.

3.1 Background

User-Tailored-Privacy: User Tailored Privacy (UTP), proposed by Knijnenburg et al. [119], is one of the recommended privacy adaptive approaches that several privacy scholars have advocated as a means to improve users' awareness of, and engagement with, privacy features [121, 119, 227]. The approach involves modelling user privacy preferences and automatically tailoring privacy settings to match these preferences [121, 239, 148]. More specifically, UTP is composed of three main parts; the *measurement* of users' privacy preferences and behaviors, then using the measurements to create a personalized *model* and finally *adapting* the user interface to the predicted privacy preferences by changing the default privacy settings, giving an explicit recommendation, and/or providing a context-based justification for the predicted behavior (see Figure 2.1). The goal of UTP is to support and/or complement users' privacy management strategies beyond simple *"settings"* towards a utilization of distinct, coherent subsets of numerous privacy features in a way that alleviates the cognitive burden related with the engagement of these privacy features. UTP arguably provides users with just the right amount of control and useful privacy related information so as not to be overwhelming or misleading.

In this study, we tested the *adapt* part of UTP [119]. We used UTP to assess the optimal adaptation method. Furthermore, since modern online technologies like Facebook have a plethora of privacy features that can all be potentially be adapted to the user's preferences, we also investigated the feasibility of tailoring Facebook's privacy features. More specifically, we answered the following research questions:

RQ1: Which features should be tailored to the user's preferences?

RQ2: How should such adaptations be effected?

3.2 Methods

To answer our research questions: *which* features should be tailored to the user's preferences, and *how* should such adaptations be implemented? We created 19 mockups of "user-adaptive" versions of Facebook privacy features. Implementing each adaptive feature with three different adaptation methods (Automation, Highlight and Suggestion) at varying levels of automation. We carried out a series of semi-structured user interviews with 18 participants, showing them paper prototypes of our adaptive privacy features, and asking them to judge the presented adaptive capabilities and the three adaptation methods. In this section, we describe our participant recruitment and interview procedures.

3.3 Recruitment and Participants

Between October and December of 2017, we recruited adult self-reported Facebook users with the purpose of collecting their feedback on our adaptive privacy features and adaptation methods. They were recruited through flyers around a university campus and the surrounding area, and via email using university student email listservs. 18 participants each completed the 45-minute interview session; their demographics are shown in Table 3.1.

ID	Gender	Age group	Features Shown
A	F	18-21	1,4,9,15,16,17
B	M	21-25	4,7,9,10,13,17
B	M	21-25	4,7,9,10,13,17
C	M	21-25	4,7,8,10,11,13
D	M	18-21	4,6,8,11,14,19
E	F	18-21	5,6,12,14,18,19
F	M	18-21	4,5,8,11,12,18
G	F	35-40	4,8,11,16,18,19
H	M	18-21	2,5,7,12,16,18
I	M	25-30	2,3,5,12,14,18
J	M	25-30	1,6,9,13,14,17
K	M	25-30	1,5,6,10,14,16
L	M	25-30	5,9,10,13,16,17
M	F	25-30	2,3,7,9,12,13,17
N	F	20-25	2,4,5,7,8,11
O	M	20-25	2,5,7,12,15,16
P	M	20-25	1,2,4,15,16,17
Q	M	25-30	1,4,6,14,17,18
R	M	25-30	5,6,12,14,18,19

Table 3.1: Participant demographics (gender, age group, and experimental treatment (features shown)).

3.3.1 Interface Mockups

The instrument for our study was a set of paper-based mockups of the 19 Facebook privacy features listed in Table 3.2. The choice of these features is inspired by Wisniewski et al. who map out an exhaustive set of boundary regulation mechanisms on various social network sites in [240], and identify Facebook features implementing these boundary regulation mechanisms in [242]. We called our system “Fakebook” and used cartoon-style renderings to have the participants focus on the presented mechanism rather than the specific graphical implementation and the feasibility of the adaptive technology. For each feature, we created a mockup of the default non-adaptive version currently available on Facebook, plus three adaptive version, with each version implementing a different adaptation method: automation, highlight or suggestion. These three adaptation methods implement varying degrees of automation [149], and are further discussed below.

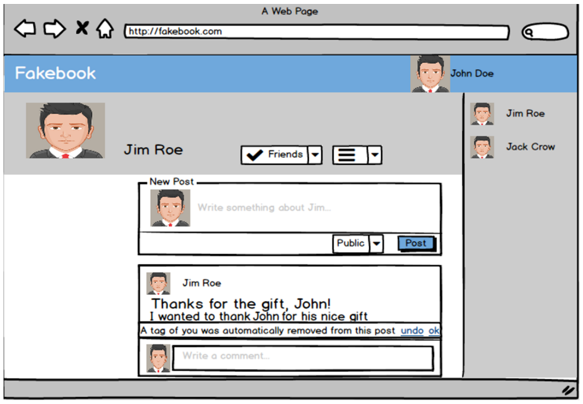
#	**Description**
1	Restrict the audience that can view your photo albums
2	Block or unblock an app or game
3	Ignore future event requests from a friend
4	Block or unblock people from seeing your timeline posts
5	Place friends into custom lists
6	Turn the chat on/off
7	Add/remove your contact information
8	Restrict the audience of a post to friends on a custom list
9	Delete a post
10	Hide a post
11	Turn on/off game and app notifications and invites
12	Restrict who can look you up using your email address or phone number
13	Untag yourself from posts
14	Place friends on the “restricted” list
15	Give feedback and/or report a post
16	Limit the default audience that can view your posts
17	Restrict who can posts on your timeline, and who can see what others post on your timeline
18	Follow or unfollow a friend
19	Add/remove your personal information e.g. date of birth, languages, political views

Table 3.2: Participant demographics (gender, age group, and experimental treatment (features shown)).

3.3.1.1 Automation

The Automation adaptation method implements adaptations without first requesting permission from the user. This adaptation method has the highest degree of automation, as it can operate completely outside of the user’s awareness. In our implementation, the user is not explicitly notified of the automatic adaptation, but they are able to see that automated action has occurred when they arrive at the location where they would have done the action themselves. For example, when a user is untagged from a post, a participant shown the Automation method would see the tag removed and replaced with a message informing them that they were automatically untagged (see Figure. 3.1).

The Automation method substantially reduces the onus of privacy decision-making but can feel like a significant loss of control [126, 203, 183]. Indeed, Vihavainen et al.[221] studied the implications of full automation on social interaction on social network sites (SNS) and found that the loss of granular control leaves users feeling powerless to adjust the specifics of what is being



Thanks for the gift, John!
I wanted to thank John for his nice gift

A tag of you was automatically removed from this post undo ok

Figure 3.1: Mockup of the Automation version of feature 13: *“untag yourself from posts”*.

disclosed. Optimization of such details of disclosure is a task that users still feel cannot readily and correctly be transferred from them to a system. Hence, in our designs, the message indicating the automated action has an undo button, allowing the user to reverse the action. The undo button makes the Automation method more similar to the other adaptation methods (which require user intervention before the adaptation is enacted) and is predicted to increase the perceived control of this adaptation method, without harming its inherently unobtrusive nature.

3.3.1.2 Highlight

The Highlight adaptation method increases the visual prominence of the action that the adaptive procedure predicts the user would want to take. This can be done either through a color change, or by giving the recommended action a more prominent location on the screen. In our implementation, we give the recommended action a yellow background color, and change its ordering in the list of options, if appropriate. The Highlight method implements a moderate degree of automation: it gives users a clear indication as to what action they should consider—reducing their cognitive load without reducing their control. As some privacy features in the Facebook interface are hidden behind a button or a menu, our Highlight implementation can also highlight the element that gives access to the adapted feature. For example, when a user is missing important basic information such as political views (See Figure. 3.2), a highlight on this missing information and of the feature that enables users to edit this basic information could be necessary. The highlight provides guidance to users in cases where the adapted feature is not prominent.

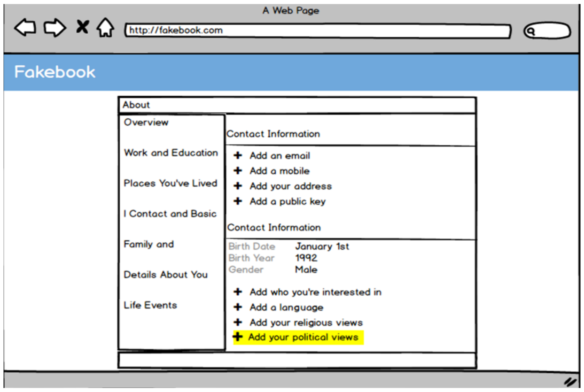
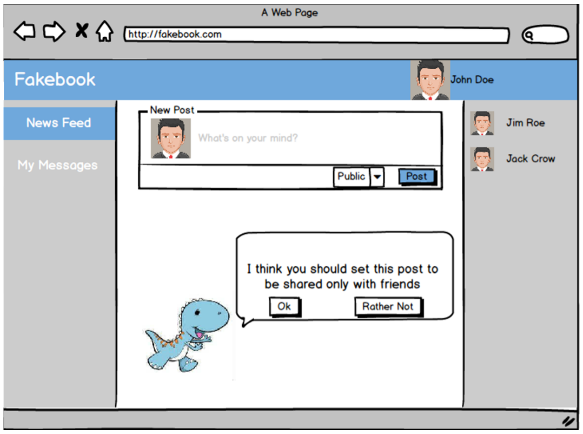


Figure 3.2: Mockup of the Highlight version of feature 19: *"Add/remove personal information e.g., date of birth, language, political views"*.

3.3.1.3 Suggestion

The Suggestion adaptation method displays an "agent" (virtual character) that verbally suggests a recommended action to the user. Our implementation is based on Facebook's "Privacy Dinosaur", which the Facebook platform currently uses to display "Privacy Check-up" notifications to the user. The Dinosaur provides suggestions in a general form of, *"I think you should..."*, increasing the personal nature of the interaction (see Figure. 3.3). The provided options are **"Ok"** and **"Rather Not"**, allowing the user to either accept or reject the recommended action. Users were told that if they selected "Ok", the setting would automatically be changed however they would still be taken to the appropriate setting as well. By asking for an explicit decision, this adaptation method implements our lowest degree of automation.



I think you should set this post to be shared only with friends

Figure 3.3: Mockup of the Suggestion version of feature 8: *"restrict the audience of a post to friends on a custom list"*.

Personalized and anthropomorphic agents have been shown to have beneficial effects on the acceptability of recommendations [127]. That said, the suggestions will likely be perceived as relatively intrusive: they take up space and time, potentially creating an undue onus. On the other hand, the explicit suggestions provide a safer alternative to the other methods, as they give the user explicit control over the adaptation.

3.3.2 Interview Procedure

Each interview session lasted about 45 minutes, and participants were compensated with a $5 Starbucks gift card for their time. An IRB-approved interview protocol was adopted to ensure consistency across all sessions. After obtaining informed consent from participants, the sessions were audio recorded and later transcribed. The interview with participant O was conducted remotely using video conferencing and screen-sharing, while the remaining interviews were conducted face-to-face. After building rapport with participants and introducing them to the study, they answered two questions for each of the privacy features in Table 3.2. The first question asked how familiar participants were with each feature (using a 5-point scale: not at all familiar–extremely familiar) and the second question asked how frequently they used each feature (5-point scale: always–never).

Next, participants were presented with a paper-based user interface mockup of a randomly selected privacy feature. They were given a scenario to fully understand the use of the feature. The scenario was:

> *"You are John Doe from Fresno, California. You are 22 years old, and regularly use Facebook for business and leisure. You are currently looking for a job and are trying to keep a clean Facebook account. You would like to < use privacy feature > to achieve < some goal >".*

Participants were then first shown the default non-adaptive version of the feature, and asked if they were aware of the feature, and how often they used it (on Facebook). If they had used the feature before, they would be asked for what purpose they used the feature. If they were completely unfamiliar with the feature the scenario would again be invoked to help them better understand the use of the feature. Next, participants were shown a randomly selected adaptive version of the same feature, and asked for their opinion on the presentation, functionality, pros, cons and comfort with the adaptive feature, and the method with which it was implemented. This procedure was repeated for a total of six times per participant. The subset of features shown to each participant is listed in the last column of Table 3.1; we ensured that all participants encountered each adaptation method at least twice (semi-random) with a different privacy feature, and endeavored to cover all the privacy features equally among all the participants.

After completing six features, participants were given an exit survey, asking them to select their preferred adaptation method (Automation, Highlight, Suggest, or As is) for each of the features. This helped us gain a broader overview on whether participants would want to use any of these adaptive features beyond the in-depth interview, and if so, which adaptation method they would prefer. The findings for each feature are presented in Table 3.3. Note that the exit survey was only completed by 10 of our 18 participants.

Feature #	Automatic	Highlight	Suggestion	As is
1	1	0	6	3
2	2	5	3	0
3	3	2	4	1
4	0	0	6	4
5	1	1	6	2
6	0	3	3	4
7	0	3	3	4
8	0	2	5	3
9	0	0	2	8
10	2	3	3	2
11	1	3	3	3
12	2	2	3	3
13	1	3	3	3
14	0	3	5	2
15	2	5	1	2
16	1	1	4	4
17	2	1	2	5
18	1	2	5	2
19	2	1	3	4

Table 3.3: The overall distribution of preference for each adaptation method per privacy feature.

3.4 Findings

Based on our analysis of the interview data, we find that Suggestion was the most preferred adaptation method, followed closely by Highlight, with Automation being the least preferred. However, we find that the preferred adaptation method for each specific feature largely depends on the user's awareness and usage of the feature, and in some cases on whether the feature results in awkward or irreversible privacy behaviors. We discuss these findings in detail below.

3.4.1 Automation

3.4.1.1 Automation and Frequency of Use

We find that participants generally dislike the Automation method, especially for features they never use or are unaware of. As participant M stated when shown the Automation version of the privacy feature that enables one to block app invites (feature 11 in Table 3.2):

> *"I was not aware you can block game app invites because I have not explored Facebook properly. Maybe if I knew this particular feature existed, I would prefer doing it manually than automatic because you never know who is getting automatically blocked."* (Participant M)

On the other hand, participants are more accepting of the Automation method for privacy features they use frequently, just as Participant C stated about the automatic removal of a tag (feature 13)

> *"It saves me a lot of time and [. . .] effort because I do not have to look through 100 posts that all my friends have tagged me in [. . .] In terms of situations where I am applying for a job or like applying for school or something maybe taking those precautionary measures has a certain cognitive load on me, so it kind of takes that off [. . .]. It follows along the line of 'prevention is better than cure' [. . .] So it kind of prevents a wrong, rather than have a wrong thing out there and then cure it. [. . .] Better safe than sorry !"* (Participant C)

Nevertheless, participants stressed the need for additional control over the automated feature, e.g., they would want to be able to turn it on or off. When shown the Automation version of the audience selector tool used to control who can see a photo album (feature 1) participant A expressed:

> *"I feel like it should be a choice for people to have stuff like this automated for you. I personally would not care for it because I feel it does not save you that much time and I can set my intended audience in a few seconds."* (Participant A)

This indicates that the ease of use is an important reason to like the Automation method, and that the absence of cognitive load reduces the need for fully automated adaptations. Furthermore, participants are worried about the accuracy of the adaptation for features they use only occasionally. For example, participant B, who only occasionally uses the "block people" feature (feature 4), argued:

> *"It means I am relying on the system to detect someone that I know needs blocking. So essentially, I am believing the system understands me perfectly. Maybe to some degree the system can learn what kind of people I block [but] I am not so sure that it's just learnable like that."* (Participant B)

3.4.1.2 The Presumed Irreversibility of Automation

With the Automation method, participants are wary that the system will reduce their ability to make their own privacy decisions. Combined with the fear that the system might get their privacy preferences wrong, they worry that the Automation method will implement privacy behaviors that are irreversible, leading to persistent negative consequences. As participant A put it when shown the Automation version of the blocking app invites feature (feature 8):

> *"I am kind skeptical of the automatic option that it automatically picks who's going to see this, because again, it could pick the wrong person and I do not notice, and then you are not able to know who sees the picture type of thing."* (Participant A)

Similarly, participant M stated about automatically blocking app invites (feature 11):

> *"Say you have a close friend who is into this game stuff, then automatically blocking him would not be nice, you would lose a friendship there"* (Participant M)

Finally, participant B made a comment about automatically blocking people (feature 4) that really shows how this fear is related to a potential loss of control:

> *"Let's say for example, I block two people that posted something about politics, so if the system understands that ok he does not like things regarding politics. I think that's kind of assumed just because two things were related to politics. I really want to know what algorithm it uses to understand my character in terms of what kind of people I block."* (Participant B)

While our implementation of the Automation method gives users the possibility to undo the automated action, this did not alleviate participants' concerns. Many stated that it might already be too late to undo the automated action by the time they take note of it. For example, participant I on the possibility to undo the Automation of the friends list management feature (feature 5) stated that:

> *"I should not have to undo. It should not do unless I tell it to. Some things cannot be undone. [...] What if it assumes that this person is my friend, yet he is my boss and I happen to share an inappropriate post with that person? Now am fired! Sure, you can undo the setting, but you cannot undo the damage."* (Participant I)

Others mentioned that having to always check to make sure the system got their preference right would only increase the cognitive load. For example, on the Automatic version of the friend list management feature (feature 5), participant I stated:

> *"Sure, you can undo the setting but [...] doesn not that even cause more or the same amount of work? I thought the point of this was to make it easier, but this makes it harder. Now I have to go through and check to make sure all is good."* (Participant I)

This responsibility could even spill over into their other social network activities. As Participant A stated about the Automation of the audience selector (feature 8):

> *"Personally, if it says 'do you want to share with friends' I would not undo [it], because most of the time I share with friends anyway. [But] if it got it wrong, it would make me be conscious about the text I post, making me read it over and over again."* (Participant A)

3.4.1.3 Automation for Actions with No Consequences

Our findings suggest that Automation is only appropriate when the automated action has no big consequences for the use. As participant L stated about the Automation of hiding a post (feature 10)

> *"If you are already going so far as to like make decisions about automatically hiding posts or what not which Facebook already does in the backend obviously why are u telling the user about that in the first place"* (Participant L)

Participants expected that the adaptation would have to be very accurate for there to be no negative consequences. For example, participant A stated that she would be comfortable with the Automation of audience selection (feature 8) if it was very advanced:

> *"I guess if it could [. . .] guess who is in the picture and what the picture is about, then you can set it to an audience. But I do not think the technology is probably there yet. It's like if a picture has 5+ people, it would probably analyze 'Oh you are at a party.' then you probably should [share it with] your friends [only].' If this was automatic I think I would rather have the automatic [version]"*(Participant A)

Similarly, participant B would only be comfortable with the system automatically blocking people (feature 4) if it were very accurate:

> *"Once I believe that the system is [. . .] very good at understanding the kind of people I block, then I would be comfortable. But if you are asking me to use it right now, am not so sure the system knows my character very well. I would use this fully if I have proof the system is good at its job of understanding what kind of people I block. Having automatic blocking kind of gives me the assurance that I am going to look clean in the eyes of the people."* (Participant B)

3.4.2 Highlight

3.4.2.1 Highlight for Unobtrusive Awareness

Participants appreciated the Highlight method for its ability to unobtrusively raise users' awareness about a privacy feature. When shown the Highlight version of the friend list management feature (feature 5) participant L stated:

> *"I think it's not obstructive to seeing the rest of the screen but it also gives a visual cue to say like we recommend this choice or information"* (Participant L)

Similarly, when shown the Highlight version for adding/removing contact information (feature 7), participant C stated:

> *"I will have to agree 100%, that it definitely makes me more aware, because otherwise I am just seeing plain text, and I do not really care about what information I am putting*

> *out there. This kind of makes me [...] more aware of what I am putting out there and what it's asking me for [...] It helps me be aware or control my privacy to a degree better."* (Participant C)

The same participant also remarked that the Highlight method is less cluttered and less taxing than the Suggestion method because it allows him to "selectively choose to ignore the highlight feature." Also comparing Highlight to Suggestion, participant K stated regarding the "hide a post" feature (feature 10):

> *"if I scroll through a bunch of posts to hide, suggestions would be annoying but if I was scrolling through and it had a highlight then that would be ok. If it was highlighted yellow or something then that would draw my attention."* (Participant K)

3.4.2.2 A U-shaped Relation with Familiarity

We find that participants' preference for Highlight depends on their familiarity with the privacy feature. On the one hand, expert users of a certain privacy feature may find Highlight a redundant adaptation method, and prefer full Automation instead. For example, participant A regarding the reporting of a post as spam (feature 15) stated that:

> *"It's a redundant adaptation to have. I understand that your trying to raise awareness that 'oh this is the spam button,' but [...] if you wanted to report it in the first place then you would report it as spam, but if you did not want to mark it as spam regardless then you would not."* (Participant A)

On the other hand, participants could easily get confused with the Highlight method if they are unfamiliar with a privacy feature, resulting in a perceived loss of control. They are instead more likely to prefer a Suggestion that provides some more information. When shown the Highlight version of the friend list management feature (feature 5) participant L stated this downside:

> *"It cannot really show a justification for why it is being highlighted over something else [...] I mean that's less information being given to the user."* (Participant L)

In sum, our findings suggest that there is a nuanced U-shaped relationship between the participant's familiarity with a privacy feature and their preference for the Highlight method: while it may be redundant for expert users of the feature, and confusing for novice users, it unobtrusively provides an optimal level of awareness to those who occasionally use the feature.

3.4.3 Suggestion

3.4.3.1 Convenience or Nuisance

All but one of the participants prefer the Suggestion method for at least one of the presented privacy features. Like Highlight, Suggestion raises users' awareness about the privacy feature. For example, participant E stated the following about the Suggestion version of the adaptive "restricted" list feature (feature 14):

> *"yes, restricted lists are not what I always think of. I think of blocking more than restricted list and thus having the suggestion pop-up brings it more to mind."* (Participant E)

Suggestions are convenient, because they provide a shortcut to the functionality. Participant J was shown the Suggestion version of the adaptive untag feature (feature 13) and he stated:

> *"I do not have to go through the settings[..]I do not have to click the drop down and find anything in the settings menu. I am given a clear choice about the tag to either keep it or remove [..] It really focuses me in on the thing that might be important."* (Participant J)

Similarly, when shown the Suggestion version of the hide post feature (feature 10), participant C brought up both increased awareness and convenience benefits:

> *"I would definitely save a lot of time because this would pop up and I would click "ok" for the posts that I do not care about [...] and it will take care of all similar posts. It's going to catch my attention more than a hide icon."* (Participant C)

Several participants appreciated the idea of getting privacy advice from a virtual character. For example, when participant A was shown the Suggestion version of limiting the default audience that can view one's posts (feature 16), she expressed:

> *"I think it's a cute dinosaur for starters. It really does grab your attention because if I am about to post and something like this pops up, I am definitely going to look at it [...] so it reminds you before you post."* (Participant A)

Similarly, when shown the Suggestion version of the feature that turns game and app notifications and invites on/off (feature 11), participant C stated:

> *"I prefer the suggestion, because comics and pictorial representations are more than just text [. . .] comically depicted speech bubbles kind of engage my mind and bring my immediate focus and attention into this [. . .] it's drawing me towards fixing the need of the hour."* (Participant C)

On the other hand, some participants did not like the virtual character, suggesting it was somewhat childish, and not serious enough for the topic of privacy. For example, participant F commented on the Suggestion version of the follow/unfollow a friend feature (feature 18):

> *"It looks like a blue bunny almost [. . .] It's a little childish I guess [...] I think a little more professional presentation would be in order."* (Participant F)

Some participants also suggested ways in which the virtual character could be improved or made better e.g. participant A, feature 16:

> *"I guess it would be better if you can have an option to change or customize it to maybe something like a privacy dog or a self-resembling avatar to [make it] seem like I am reminding myself."* (Participant A)

Furthermore, we find that participants tended to dislike the Suggestion method for features they use frequently. This is because too many suggestions require considerable attention from the user to successfully be dealt with. As participant C continues to explain about the hide post feature (feature 10):

> *"I do not want to see more than two of these at a particular instance for two consecutive posts [. . .] It gets repetitive [. . .] I do not want to see a suggestion saying the exact same thing on three consecutive posts, even if those posts are things that I do not care about."* (Participant C)

Similarly, participant F commented about the follow/unfollow feature (feature 18):

> *"It would be okay if it was every once in a while. I really do not want it to be like 'oh you should do this this this and this!' [...] But I think [I would like it] if every once in*

> *a while, it was like 'you have not spoken to this person in three years maybe u should unfollow'."* (Participant F)

3.4.3.2 An Opportunity for Explanation

For features that participants are unfamiliar with, Suggestion has an added advantage: the opportunity to explain the privacy feature and the adaptation to the user. These explanations give users a reason for the Suggestion, thereby actively helping them learn something about Facebook privacy. Combined with the ability to either follow or ignore the suggestion, such explanations may help users feel more in control of their privacy. For example, when participant H was shown the "follow or unfollow a friend" privacy feature (feature 18), he stated:

> *"I think it would be helpful if it gave a reason, the tough part is counting on the person to follow through. But I think people value their privacy and I think it would be successful because there [are] lots of fake accounts."* (Participant H)

Similarly, when participant I was shown the privacy feature that restricts who can look him up using his email address or phone number (feature 17), he stated:

> *"I feel like if you are going to suggest something to me, you should give me a reason."* (Participant I)

3.4.3.3 Awkward Suggestions and Social Norms

Our findings show that suggestions can break certain social norms, especially when applied to private behaviors that carry a negative social perception, such as deleting posts and unfollowing users. For example, when participant I was shown the "follow or unfollow a friend" privacy feature (feature 18), he stated:

> *"I might like someone's posts a lot but not follow them. Thus the system can suggest that I follow them based on those likes. However it should not make a suggestion that I unfollow anyone, because common sense dictates [that you should not suggest to me to unfollow people]."* (Participant I)

Similarly, participant L stated about the deletion of a post (feature 9):

> *"I do not want Facebook to suggest what I should delete because that would be a weird decision to make for me."* (Participant L)

Indeed, some participants mentioned that the Privacy Dinosaur adds to the awkwardness of Suggestions that carry a negative social perception. For example, when participant I was shown a privacy dinosaur that came with the suggestion to restrict who can look him up using his email address or phone number (feature 17), he stated:

> *"Why is the dinosaur giving me a suggestion and not just insight? [. . .] I do not think the dinosaur should suggest. I think the dinosaur should just give me options, or [tell me] what different options do or something [. . .] But giving me a suggestion without actually giving me reason why it's suggesting would probably be a reason I would be uncomfortable [with it], because I would feel like this dinosaur knows more than it's giving me information about."* (Participant I)

This comment also suggests that explanations can potentially reduce the awkwardness of Suggestions by carefully explaining the reasoning behind them. Without explanations, though, certain suggestions felt unsolicited or even rude. As participant H expressed about the Suggestion to turn on/off game and app notifications and invites (feature 11):

> *"The system could notice how much I have been clicking 'NO'. It would then be helpful to have a suggestion that says, 'we noticed you say NO a lot, do you want to block the app invite?' It's kind of a call to action I guess."* (Participant H)

3.4.4 No Adaptation

3.4.4.1 No Adaptation Rather than a Different Method

We have already discussed several situations where participants preferred the traditional untailored privacy features to our user-tailored alternatives. This preference was most pronounced for seemingly irreversible actions (especially when participants saw such features paired with the Automation method, e.g. participant I, feature 5: *"Sure, you can undo the setting, but you cannot undo the damage"*) and for actions with a negative social perception (especially when participants saw such features paired with the Suggestion method, e.g. participant L, feature 9: *"I do not want Facebook to suggest what I should delete because that would be a weird decision to make for me."*)

In both cases, participants did not prefer a different adaptation method, but rather opted for no adaptation at all.

3.4.4.2 The User Is the Best Adaptation Algorithm

Beyond this, the preference for 'no adaptation' also seemed to correlate with participants' trust in the system's ability to learn the user's preference (again, this was most pronounced when participants saw the Automation method, e.g. participant B, feature 4: *"It means I am relying on the system to detect someone that I know needs blocking"*), and finally, this preference seemed to correlate with participants' familiarity with the privacy feature. For example, when shown the Suggestion to have the chat feature turned off (feature 6), Participant J stated:

> *"I feel like if I turn off the chat it's because I want to be temporarily without notifications and I will come back and turn it back on later. But I think more likely I will just put my phone on silent [...] I want chat all the time—like, that's my main use of Facebook. I would not want some automatic process to turn it off. And if it suggested I turn it off, I would not listen."* (Participant J)

In sum, when participants distrusted the algorithm behind a certain adaptive privacy feature, or when they were already intimately familiar with the privacy feature, they essentially considered themselves to be a better adaptation algorithm than the system. Hence, in these cases they preferred the traditional untailored version of the privacy feature.

3.5 Discussion

Our findings summarized in Table 3.4 answer and shed an interesting light on our research questions. We find that the preferred adaptation method for the different privacy features depends on users' awareness and usage of those features (RQ2). Since different Facebook users are (un)familiar with different features, this means that the preferred adaptation method for each feature differs per user. The adaptation method itself should thus be tailored to the user as well.

Moreover, we find that the preferred adaptation method may sometimes not be suitable, in which case users end up preferring the untailored version (RQ1). This limits the extent to which user-tailored privacy can be implemented on Facebook.

Automatic/Irreversible?	**Awareness/Usage**		
	Unfamiliar/Do not use	***Occasional Use***	***Frequent Use***
Yes	As is	Highlight	As is
No	Suggestion	Highlight	Automation

Table 3.4: Preferred adaptation methods given adaptation effects and user privacy feature awareness or usage

3.5.1 Unfamiliar/Infrequently-Used Features

When Facebook users are unfamiliar with a privacy feature, they prefer the Suggestion method, mainly because our implementation of Suggestion (the "Privacy Dinosaur") allows for the adaptive behavior to be explained. The infrequent use and unfamiliarity makes the *load* of using them more cognitive rather than physical. With proper explanation, the Suggestion method actually reduces this load.

Moreover, its superior level of *control* turns the Suggestion method into a "privacy education" tool that introduces users to a privacy feature they were previously unaware off. Normally, introducing users to a new privacy feature can be daunting or confusing: because the user is unfamiliar with the feature, they may not know how to interact with it (for example: if the user has never "blocked" another user, they may not know when it would be appropriate to do so). The adaptive behavior solves this problem, though, by not only introducing the feature to the user, but also suggesting to the user how to interact with it, thereby reducing the cognitive load. In effect, the adaptive nature of the Suggestion makes it a very accessible tool for education.

However, users do *not* prefer the Suggestion method when it gives the wrong suggestions that they are likely to find awkward such as blocking a friend. Such a suggestion is considered to be against the norm of social interaction. Therefore, rather than opting for one of the other adaptation methods (which lack the desired explanation of the adaptive behavior), users prefer that the privacy feature remains untailored.

3.5.2 Occasionally-Used Features

When Facebook users use a feature occasionally, they may prefer the Highlight method. This preference is mainly a compromise: Suggestion would significantly be a destruction for features that are used with some regularity (the privacy dinosaur would show up too frequently), while Automation would significantly reduce *control* (users are not familiar enough with these features to comfortably allow the system to take over altogether).

3.5.3 Frequently-Used Features

When users use a feature frequently, users prefer Automation, suggesting that they are willing to give up some *control* in return for a reduction in the *effort required for proper* privacy management. Frequent users already know what to do with a feature, so their main effortful load is rather physical than cognitive. In effect, neither Highlight nor Suggestion would sufficiently reduce this load. Moreover, users seem to have an intuitive understanding that their frequent use of a feature likely improves the quality of the adaptive behavior. This gives them a certain amount of "indirect" control over the Automation method.

However, users do *not* prefer the Automation method when the resulting automated privacy decision feels irreversible. For example, Facebook users would not appreciate the system automatically unfriending or blocking their friends, deleting their posts and setting their post audiences. Even though our implementation of Automation provides a clear mechanism to "undo" the decision, making every decision technically reversible, users are uncomfortable when a system automatically implements a decision that "feels" irreversible without asking the user.

3.6 Design Implications

We offer the following insights for social network designers interested in implementing user-tailored versions of privacy features. While our study focuses on the Facebook platform, we argue that our insights are sufficiently generic to also apply to other social networks (or indeed, other information systems in general).

3.6.1 Selectively Automate Privacy Features

Our findings suggest that designers can automate privacy features to relieve some of the user responsibility in privacy decision-making. However, they are advised to only do so for features that users use frequently, and to avoid automating any privacy behaviors that are perceived as having irreversible consequences. Given the large variation in privacy feature usage among Facebook users [242], this means that the selective Automation of privacy feature should itself be tailored: the system should find out which features each user frequently uses, and only automate those features.

Accuracy is of utmost importance when fully automating privacy features: Many participants in our study portrayed a lack of trust in the system's ability to accurately tailor its privacy settings to their preferences. Unless the underlying algorithm is extremely accurate, users will likely believe that they themselves are much better at managing their own privacy (even though research shows this often not to be the case! [156, 150, 206]).

3.6.2 Selectively Apply Highlights

Designers can use subtle highlights recommending certain privacy behaviors as a means to assist users in making better decisions, but also to help raise their awareness of certain privacy features that they may have forgotten about. Designers can capitalize on the subtle awareness-raising capabilities of this adaptation method by using it primarily for privacy features that users only use occasionally. Again, this means that the application of the Highlight method should itself be tailored to the user.

3.6.3 Selectively Make Suggestions

Facebook already has a Privacy Dinosaur that makes privacy-related suggestions, so designers have the opportunity to leverage this functionality to make adaptive privacy suggestions or design a similar virtual character for other social networks/information systems.

The virtual character should be designed not only to make privacy recommendations, but also to explain those recommendations: several participants in our study suggested—unprompted—to include explanations of the adaptive behavior in the dinosaur's suggestion. Designers should avoid the potential awkwardness of suggesting privacy behaviors with negative social connotations (e.g. blocking or unfollowing people), though. That said, a good explanation can alleviate some of these concerns.

The opportunity for explanations also makes the Suggestion method particularly useful for introducing the user to privacy features they are unfamiliar with. Again, this means that the application of the Suggestion method should be tailored to the user's awareness of the various privacy features.

3.7 Limitations and Future Work

An obvious limitation of our study is that our adaptive privacy features were mere paper mockups, using cartoon-style renderings with less visually distracting features as compared to the actual Facebook. This might have given them a less realistic appearance, but also made it easier for the participants to concentrate on the presented adaptation mechanism and envision the use of the adaptive privacy features without getting hung up on design details. Moreover, whereas in real life such adaptive features would likely make the occasional mistake, our presented scenarios assumed that the adaptation methods presented to participants worked with 100% accuracy. That said, participants questioned the idea that the adaptive system would always get their privacy preferences right, and frequently brought this up as a potential reason to prefer the traditional untailored privacy features. As such, the potentially reduced accuracy of the presented adaptations in real-world systems is likely to significantly impact users' perceptions and may result in a reduced preference for adaptive privacy functionality.

On the other hand, we note that most existing work on adaptive privacy features evaluates their accuracy only, without testing the user experience of the resulting system or the usability of the mechanism by which the privacy recommendations are presented to the user (Liu et al. [149] and Knijnenburg and Jin [118] are notable exceptions). Our paper demonstrates that the method by which the recommendations are presented has a strong influence on the user experience. Hence, we encourage researchers and developers of adaptive privacy features to conduct usability and user experience tests.

Our study design relied on users' self-reported evaluations of the paper-based mockup designs we showed them. While this allowed users to critically reflect upon the consequences of the user-tailored functionality and the three adaptation methods, users did not have the opportunity to interact with the privacy features in a social network interface. This precludes us from making strong claims about the usability of the adaptation methods, and it may even mean that users' preferences for these methods change once they have the opportunity to interact with them. Thus, future research should explore the usability of different adaption mechanisms in an interactive test environment.

We also limited ourselves to a subset of prominent Facebook privacy features as previously identified by Wisniewski et al. [242]. They cover only a limited subset of the available privacy features and are restricted to the features on the Facebook platform. That said, we made sure that the selected features span the various "boundary protection mechanisms" covered in [240]—a work that also demonstrates that these mechanisms exist in various forms across a variety of social network sites.

Despite these limitations, the answers to our research questions constitute a clear pattern of user preferences, with Table 3.4 mapping out which situations call for adaptive privacy features, and which adaptation method would likely be preferred. We argue that these insights are sufficiently generic to apply to any social network site, or indeed any information system that may benefit from adaptive privacy features. In future work, researchers, developers and designers can leverage these insights for the development adaptive privacy features in research prototypes or real-world social networking sites.

3.8 Conclusion

This chapter demonstrated the viability of UTP and highlights users' initial perceptions towards the three proposed adaptation methods (*Automation*, *Highlight, and Suggestion*) in increasing user awareness, engagement and use of the available privacy features within some modern online technologies such as SNSs like Facebook. Our findings reveal that participants generally dislike the full Automation method, except for privacy features they use frequently and perceive as inconsequential, where it can alleviate some of the behavioral onus and effort of managing one's privacy. The Highlight method is appreciated for its ability to unobtrusively raise users' awareness about a privacy feature and is thus most suitable for features users only use occasionally. Finally, the Suggestion method is preferred as a means to teach users privacy features they are unfamiliar with, unless this results in awkward suggestions of behaviors with negative social connotations. As the familiarity with and usage of the various privacy features differs extensively per user, we argue that the choice of adaptation method itself needs to be tailored to the user as well.

An intuitive next step is to quantitatively understand the effectiveness of the three adaptation methods in improving users' engagement with the available privacy features and their overall levels of privacy. In Chapter 4, I describe an experimental study (N = 406) in which we developed and leveraged a functional but carefully controlled SNS UI prototype to test users' privacy management behaviors. The controlled but semi-realistic SNS environment implementing one of the adaptation methods, allowed us to gain an empirical understanding of the effectiveness of the adaptation method in improving user engagement and their overall level of privacy protection.

Chapter 4

The Effectiveness of Adaptation Methods on Privacy Decision-Making

In Chapter 3, I shed light on users' initial perceptions towards the application of three adaptation methods (*Automation*, *Highlights*, and *Suggestions*) in presenting suggested privacy behaviors. In particular, the *Automation* adaptation method was generally disliked, except for privacy features that were used frequently and perceived as inconsequential. In this case, automation could help alleviate the cognitive burden involved in privacy management. The *Highlight* adaptation method was appreciated by users for its ability to unobtrusively raise users' awareness about a privacy feature. The *Suggestion* adaptation method was preferred as a means of teaching users privacy features that they were unfamiliar with, unless the the feature was awkward to suggest or had a negative social connotation. However, it remains unclear if these adaptation methods would actually be effective at improving users' engagement with the available privacy features (to encourage active ownership over one's privacy) and their overall levels of privacy protection. As such, this begs the question: *which* adaptation method(s) are effective at improving user engagement with the privacy features and offer better overall privacy protection outcomes?

In this chapter [1], I present an experimental study aimed at examining the effectiveness of the

[1] Published as [169]

adaptation methods in improving user engagement and overall privacy protection on modern online technology such as an SNS platform. We systematically evaluated the effect of these adaptation methods on participants' engagement with the privacy features, their tendency to set stricter settings (protection), and their subjective evaluation of the assigned adaptation method. We found that the *automation* of privacy features afforded users the most privacy protection, while giving privacy *suggestions* caused the highest level of engagement with the features and the highest subjective ratings (as long as awkward suggestions are avoided). Based on our findings, we provide practical recommendations to improve user awareness of, and engagement with, privacy features on modern online technologies like social media platforms.

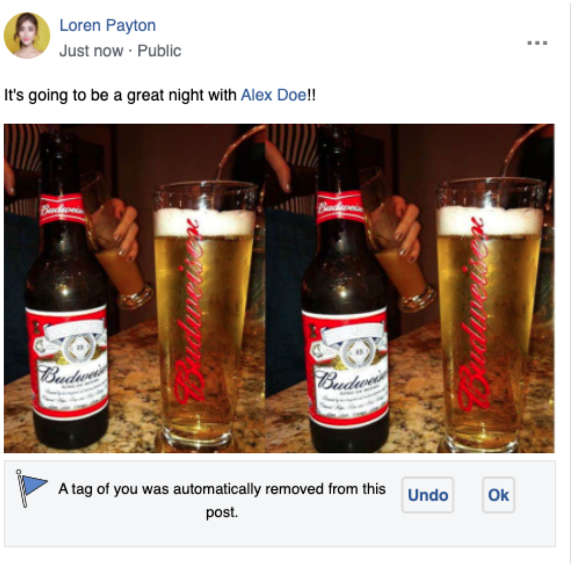
4.1 Background

Presenting Privacy Adaptations: While prior work has identified methods to create personalized models that can be used to adapt a system's privacy settings to the user's preferences, limited research has focused on the design and presentation of these adaptations [148, 50, 236, 225]. The concept of "presentation" goes beyond the visual characteristics of the adaptation and can have a profound impact on the required level of engagement with the system and the user's tendency to follow the suggested adaptation. For example, while some propose to fully automate the privacy decision-making process (e.g. [198]), others have implemented adaptive suggestions (e.g. [148], or suggested the use of personalized nudges (e.g. [227]) or interface adaptations (e.g. [236]). More specifically, Liu et al. [148] found that mobile app permission setting suggestions based on user privacy preferences were perceived to be helpful and largely adopted by users. Most importantly, the suggestions increased user engagement with the privacy settings. Warberg et al. [227] reaffirmed the importance of examining the possibilities of tailoring privacy nudges to align with individual differences in decision making and personality, especially among large organizations such as SNS that typically have a large number of users. Wilkinson et al. [236] recognized that the privacy features on social networks are often more than one click away, and explored the idea of adapting the social network User Interface (UI) in such a way that it increases the salience of those features that fit the user's personal privacy management strategy (cf. [242]).

While this existing work has explored different methods of adaptively assisting users with their privacy management practices, all but two papers (i.e. [170, 50]) have compared various adaptation method in terms of their effectiveness at enhancing user engagement and overall privacy protection. The first exception is the work highlighted in Chapter 3 where we identified three adaptation methods—Automation, Highlight, and Suggestions—that varied in the level of autonomy and control afforded to users (ranging from full control to no control [202, 203]) in managing their privacy. The second exception, Colnago et al. [50], adopted the adaptation methods used by Namara et al. [170] in the design of different automation levels for a personalized Internet of Things (IOT) privacy assistant (PPA). They found that in choosing an appropriate adaptation method, users weigh their desire for control against their fear of cognitive overload in making privacy decisions.

Thus, building on the work in Chapter 3 and Colnago et al.'s [50] , the work in this study implemented the same three adaptation methods but within a functional and carefully controlled SNS UI prototype to allow us to gain an empirical understanding of the effectiveness of the adaptation method in improving user engagement and their overall level of privacy protection. we iterate on the definitions of the three adaptation methods described in Chapter 3 (Section 3.3.1) to align them to the context of this study:

> **Automation:** The "Automation" adaptation method involves the automatic manipulation of a privacy feature without first requesting user permission. While this adaptation method can operate completely outside of the user's awareness, our implementation does leave a message on the privacy feature informing the user of the automated action taken by the system on their behalf. For example, when a user is automatically untagged from a post, the tag would be removed and replaced with a message informing them that they were automatically untagged (see Figure. 4.1). Coupled with the message is a small "*Undo*" button that allows the user to reverse the automated action if they are uncomfortable with the automated setting.



Loren Payton

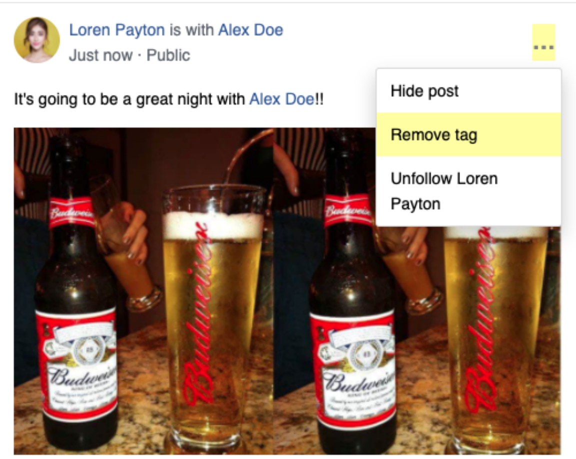
Just now · Public

It's going to be a great night with Alex Doe!!

A tag of you was automatically removed from this post.

Figure 4.1: The automation adaptation of the privacy feature for *"untagging oneself from a post"* in our SNS UI mockup

Highlight: The "Highlight" adaptation method involves increasing the visual prominence of a privacy feature—a subtle "nudge" that is meant to encourage the user to undertake a certain privacy action. This is achieved by highlighting the background of the privacy feature using a highly contrasting color (in our study: a yellow background color). Note that our SNS UI mockup is based on the Facebook UI, in which many privacy features are hidden behind menu options or have multiple navigation pathways. The highlight implementation therefore illuminates not only the privacy feature itself, but also the path towards it. For example, when a user is tagged in a post, the Highlight adaptation to untag the user emphasizes both the context menu button that contains the "Remove tag" feature as well as the feature itself (see Figure. 4.2).



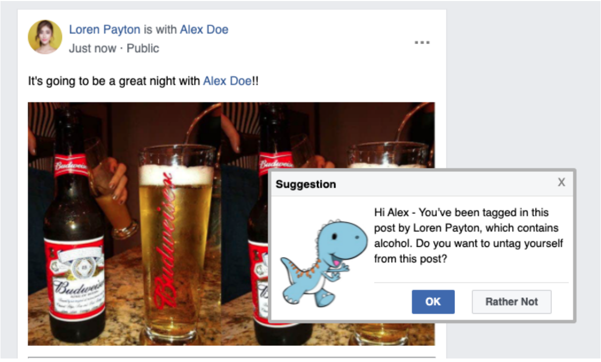
Loren Payton is with Alex Doe

Just now · Public

It's going to be a great night with Alex Doe!!

Figure 4.2: The highlight adaptation of the privacy feature for *"untagging oneself from a post"* in our SNS UI mockup

Suggestions: The "Suggestion" adaptation method involves proactively recommending the privacy action to the user. Namara et al. [170] display the recommendation message using a virtual character ("agent") to increase its prominence and to be more endearing [4]. Moreover, although recommendation messages can vary in tone and framing, Namara et al. use a positive framing (i.e. nudge the user towards taking the suggested action), giving the user the option to accept ("*Ok*") or reject ("*Rather Not*") the recommended action. We use the same implementation as Namara et al. [170] (see Figure. 4.3) because their particular implementation was well-received in their interview study. We leave the investigation of alternative versions of this adaptation method for future work. If users click "*Ok*" the suggestion is implemented directly. If the suggestion appears when the privacy feature is not on the user's screen, users are transferred to the page or point where the feature appears, so that they can verify the adaptation and adjust the setting if needed.



Loren Payton is with Alex Doe
Just now · Public

It's going to be a great night with Alex Doe!!

Suggestion

Hi Alex - You've been tagged in this post by Loren Payton, which contains alcohol. Do you want to untag yourself from this post?

Figure 4.3: The suggestion adaptation of the privacy feature for *"untagging oneself from a post"* in our SNS UI mockup

Facebook Users' Privacy Behaviors and Features: As mentioned above, the SNS UI mockup used in this study was based on the Facebook UI, which has many privacy features. Wisniewski et al. [240] identified and categorized an exhaustive set of prevalent boundary regulation mechanisms supported on social media platforms. They found that Facebook supported its users' privacy preferences through features that facilitated management of access to oneself (e.g., blocking other users, or hiding one's online status to avoid unwanted chats), management of personal information (e.g., withholding contact or basic info), management of interpersonal interactions (e.g., friending and unfriending), management of virtual spaces (untagging posts or photos or deleting unwanted content posted by others), and management of interactions between networks (e.g., hiding one's friend list from others). In a follow up study, the authors identified 36 privacy features users often used to perform these privacy behaviors [242]. They analyzed the behavioral patterns in a collected dataset and found that the users' engagement with the identified features loaded onto 11 distinct latent factors. Moreover, they were able to identify 6 groups of participants who employed distinctly different privacy management strategies to achieve their desired level of privacy. Namara et al. adopted 19 of the privacy features identified by Wisniewski et al. [242], making sure to include features from all 11 identified factors. To make our study more manageable, we further reduced the number of privacy features to 13, still keeping at least one from each of the 11 identified factors.

4.2 Methods

Our user experiment aimed to examine the effectiveness of adaptation methods—automation, highlights, and suggestions—in improving user engagement and overall privacy protection. Specifically, we sought to answer the following research questions:

RQ1: Which adaptation method(s) are effective at improving user engagement with the privacy features?

RQ2: Based on their default application and user engagement patterns, which adaptation method(s) offer better overall privacy protection outcomes?

RQ3: Which adaptation method(s) do users find most helpful?

Going beyond previous work [170, 50], we specifically examined the actions users took when privacy features were adapted and presented using these adaptation methods. The Clemson University Institutional Review Board (IRB) approved our study.

4.2.1 The SNS User Interface Mockup

Participants interacted with a carefully controlled working prototype of an SNS platform ("*FriendBook*", see Appendix A.1, Figure 1). To increase the realism and ecological validity of the experiment, the FriendBook UI was based on the UI of the Facebook web application[2] and populated with posts using the Tweet corpus collected by Cachola et al. [42]. Each user saw the exact same posts, friends, etc., thereby guaranteeing that all users had the same opportunities to engage with the various privacy features. Using FriendBook allowed us to manipulate how we applied the adaptation methods to the adapted privacy features; in some conditions we applied the same adaptation method to all features, while in other conditions we avoided adapting certain features and/or or tailored the adaptation method to the user's awareness and past usage of each privacy feature (see Table 4.1, and Section 4.2.4 for a description of the experimental conditions).

[2] FriendBook was developed before a new Facebook UI design was deployed in September 2020.

Conditions	Description	N
None (C1)	No adaptation is applied to any of the features.	54
$_{all}$Automation (C2)	All 13 privacy features are presented as having been automatically executed by the system.	49
$_{all}$Highlight (C3)	All 13 privacy features are highlighted using a yellow color.	45
$_{all}$ Suggestions (C4)	Suggestions are provided for all 13 privacy features.	47
$_{all}$Tailor (C5)	The adaptation method applied to each privacy feature depends on users' familiarity with and prior usage of the feature (on Facebook), as explained in Table 4.3.	61
$_{some}$Automation (C6)	The privacy features are presented as having been automatically executed by the system, except for the features deemed "irreversible" in Namara et al. [170] (i.e. the three Block features).	46
$_{some}$Suggestions (C7)	Suggestions are provided for the privacy features, except for the features deemed "awkward" in Namara et al. [170] (i.e., the three Block features, Delete post, and Unsubscribe from a friend).	40
$_{some}$Tailor (C8)	Like Condition C5, but automation is avoided for "irreversible" features and suggestions are avoided for "awkward" features (no adaptation is applied instead).	64

Table 4.1: Overview of the strategies used to adapt the 13 privacy features in each of the eight experimental conditions. Included are the number of participants (N) recruited in each condition. Note: There is no "some" variant of the Highlight condition, since Namara et al. [170] did not find any features for which its application was deemed problematic.

As outlined in Section 4.1, we implemented adaptive versions of 13 privacy features (see Table 4.2 for descriptions) inspired by Wisniewski et al.'s [242] inventory of Facebook's privacy features. The selected 13 privacy features cover privacy behaviors commonly performed on Facebook such as altering the News Feed, managing profile information, friend management, limiting access control, blocking people/apps/events, restricting chat, and friend management [242]. The privacy features were similar in design and functionality to those found on Facebook.

All user interactions with the privacy features (adapted or not) were recorded and used to assess overall engagement patterns and privacy protection outcomes (see Section 4.2.5).

Privacy Behavior	**Feature Name**	**Description**
Altering News Feed	Hide post	Hide a post from the timeline or newsfeed.
	Unsubscribe from a friend†	Stop seeing a person's posts in the newsfeed but remain friends with them.
Selective Sharing	Audience selection	Restrict the audience that can view posts.
Timeline/Wall Moderation	Delete Post†	Delete a post.
Reputation Management	Remove Tag	Untag oneself from a post.
Restricting Chat	Changing chat availability	Turn the online chat indicator (i.e., active status) on/off.
Managing Contact Info	Contact Info	Remove contact info (e.g email, phone number, home address).
Managing Basic Info	Basic Info	Remove basic info (e.g date of birth, gender, religious/political views).
Friend Management	Organize friends	Place a friend into a custom list.
Limiting Access Control	Control who can post on timeline	Restrict the audience that can post to one's timeline.
Blocking People	Block a person*†	Stop a person from seeing one's timeline.
Blocking Apps/events	Block app invites*†	Used to block future application requests from particular friends.
	Block event invites*†	Block future event invitations requests from particular friends.

Table 4.2: The 13 Privacy Features adapted using the 3 adaptation methods. $*$: deemed "irreversible"; †: deemed "awkward".

4.2.2 Study Setup

Participants were recruited between December 2019 and January 2020 via Amazon Mechanical Turk,a participant recruitment platform where people complete short tasks and receive automatic payments [193]. A total of 575 adult participants were recruited.

We restricted participation to people within the United States with a high "worker reputation" (i.e., those with a HIT approval rate greater than 95% with at least 50 approved past HITs) to ensure satisfactory response quality. We also included several attention check questions and quality checks to remove participants who spent little time (less than 1 minute) within the study environment or who did not carefully read/respond to the pre- and post-survey questions [132]. After discarding 169 participants who did not meet our participation requirements and data quality checks, the valid data used in the analysis was from 406 participants: (215 Men, 189 Women), aged between 18 and 60 (median category: 25-30).

4.2.3 Study Procedure

After reading the consent form and agreeing to partake in the study, participants completed a pre-survey. This pre-survey asked participants to indicate their awareness and past usage of each of the 13 privacy features (on Facebook). This was done by showing the participant an image of the privacy feature under examination and asking them 1) "Are you familiar with this Facebook feature: [Name of Feature]?" (response options: Yes, No) and 2) "How often do you use this feature?" (response options: Never Used, Used Once, Occasionally Use, Frequently Use). The responses to these questions enabled us to appropriately tailor the adaptation methods of each privacy feature based on their awareness and past use of the feature for the participants in experimental conditions C5 and C8 (see Section 4.2.4 for details on how the adaptation method was tailored in these conditions). Note that while this tailoring procedure was only implemented in conditions C5 and C8, all participants filled out the pre-survey to prevent this procedure from becoming a confounding variable.

A job search scenario was used as a motivating context in which participants could explore and manipulate the FriendBook profile used in the study. Specifically, participants were invited to imagine that:

> "*You are Alex Doe from Fresno, California and regularly use FriendBook (a social media site) for professional and leisure activities. You are currently looking for a job and have been advised by your mentor that employers monitor and scrutinize applicants' FriendBook profile before making decisions on whether to hire them or not. They have provided you with the following smart practices to consider about your profile as you go through the application process.*"

A list of smart practices (See example in Appendix A.2, Figure 2) was shown to participants following the scenario to ensure that they were cognizant of the types of tasks they could perform while on FriendBook. They were quizzed on this list to make sure that they paid attention to it. Together, the scenario and the list of smart practices helped participants navigate, engage, explore and review "their" profile on FriendBook. For easy recollection of the use context, the list of smart practices was also presented as a persistent sidebar throughout the user interaction process with

FriendBook. (see Figure 1). This list was carefully pilot-tested (N = 25) to make sure that participants were properly motivated to manage their profile without explicitly demanding that they would engage in specific privacy management practices. Responses in our pilot-test debriefing interviews convinced us that participants would interact with the privacy features that *they themselves* thought to be the most appropriate ones to engage with.

Participants were subsequently asked to explore and interact with their profile on FriendBook, with the goal of ensuring that they were okay with what is on it, given the imagined upcoming job interview. In this phase, participants explored the various posts/friend, profiles/settings and—where appropriate—made changes using the available privacy features[3]. Depending on the experimental condition, (a subset of) the 13 privacy features would be adapted to the user using the designated adaptation method(s).

Upon completing the FriendBook task, participants were asked to evaluate the overall usefulness of the FriendBook platform (based on a scale adopted from [52]) and the perceived level of decision help they believed the platform provided (based on a scale adopted from [124]). Each participant was compensated with $3 for participating in the study.

4.2.4 Experimental conditions

We developed a total of eight experimental conditions, with each condition applying the adaptation methods to the privacy features in a unique manner (see Table 4.1 for an overview). Condition C1 served as a baseline where no adaptations were applied at all. In conditions C2-C4, all 13 privacy features were adapted to the user, using one of the three adaptation methods (Automation, Highlight, Suggestions, resp.).

Condition C5 was motivated by the results of Namara et al. [170], who concluded that it likely would be expedient to tailor the adaptation method itself to the user's prior knowledge and usage of the feature. Hence, in this condition the application of one of the adaptation methods was conditional upon participants' answers in the pre-survey regarding their familiarity with and usage of the privacy features (on Facebook): the Automation adaptation was applied to any privacy features the participant used frequently on Facebook; the Highlight adaptation was applied to any privacy features the participant used only occasionally; no adaptation was applied if the user had

[3] Participants who spent too little time (<1 minute) interacting with FriendBook were removed from the analysis. The remaining participants spent an average of 5 minutes on FriendBook.

consciously decided not to use the privacy feature (i.e., they were aware of the privacy feature, but never used it or used it only once and then abandoned it); and the Suggestion adaptation was applied if the user was not aware of the adaptation (see Table 4.3 for an overview).

Aware of privacy feature?	**Usage of privacy feature**	**Adaptation Method**
No	N/A	Suggestion
Yes	Never Used/Used Once	Default
	Occasionally Use	Highlight
	Frequently Use	Automation

Table 4.3: Adaptation method selection rules for the Tailor conditions (C5 & C8) as suggested by Namara et al. [170]

Condition C6 constituted a variant of condition C2, where the Automation adaptation method was applied to all privacy features *except* those features whose effect participants in Namara et al.'s study had deemed "irreversible", i.e., the three Block features (see Table 4.2). Similarly, Condition C7 constituted a variant of C4, where the Suggestion adaptation method was applied to all privacy features *except* those features for which participants in Namara et al.'s study had indicated that a suggestion would be "awkward", i.e., the three Block features, Unsubscribe from a friend, and Delete Post. Finally, condition C8 constituted a variant of condition C5, where the adaptation method of the privacy feature was tailored to the user, but where the Automation adaptation was avoided for "irreversible" features and the suggestion adaptation was avoided for "awkward" features (in those cases, no adaptation was applied).

4.2.5 Measurement

We recorded all user interactions with the privacy features to measure their engagement:

Manual accept: The participant "manually" interacted with a privacy feature that was not adapted, or they rejected the adaptation initially but then manually restricted their privacy after all.

Explicit accept: The participant explicitly accepted the adaptation, either by approving the suggestion (by clicking "Ok"), engaging with the highlighted feature, or verifying the automated adaptation (by clicking "Ok").

Implicit accept: The participant ignored an automated adaptation, thereby implicitly accepting it.

Implicit reject: The ignored highlighted feature or the suggested adaptation, or simply did not interact with the privacy feature at all.

Explicit reject: The participant explicitly rejected the suggestion (by clicking "Rather Not") or the automated adaptation (by clicking "Undo").

Based on these user actions, we assessed the overall engagement patterns (Section 4.3.1) and subsequently the privacy protection outcomes (Section 4.3.2) across all the eight experimental conditions. We define **positive engagement** as the sum of participants' manual engagement with the privacy features and their explicit acceptance of adaptations, and **negative engagement** as the explicit rejection of adaptations. We define **privacy protection** as the sum of participants' manual engagement with the privacy features, their explicit acceptance of adaptations, and their implicit acceptance of adaptations. For these three metrics, we used multilevel logistic regressions with a random intercept for participant to compare the odds of engagement / protection between the eight experimental conditions.

The post-study questionnaires assessing **perceived decision help** and the **perceived usefulness** of the platform were submitted to a confirmatory factor analysis. Both factors had a good reliability and convergent and discriminant validity [4]. Table 1, Appendix A.3 shows the factor loadings, as well as Cronbach's alpha and average variance extracted (AVE) for each factor.

We compared each adaptation condition (C2-C8) against the none condition (C1), compared between the adaptation conditions (C2-C5 for "all" and C6-C8 for "some"), and compared between the indiscriminate ("all") and selective ("some") variants of Automation (C2 vs. C6), Suggestions (C4 vs. C7) and Tailor (C5 vs. C8). Since we made a total of 19 comparisons per outcome variable, we corrected for familywise error using the Benjamini-Hochberg method[5] [211].

[4] Cronbach's alphas > 0.8 indicate good reliability. AVEs > 0.5 indicate convergent validity, and $\sqrt{AVEs}$ higher than the inter-factor correlation indicate discriminant validity.

[5] A post-hoc method that reduces α to account for family-wise error "by sequentially comparing the observed p-value for each of a family of multiple test statistics, in order from largest to smallest, to a list of computed B-H critical values" [211, p.78].

4.3 Results

Figure 4.4 shows the distribution of user actions in the eight experimental condition (C1-C8). Below, we first analyze the significant differences in **user engagement** between conditions, followed by the differences in **privacy protection**. We end this section with an analysis of users subjective evaluations (**perceived decision help** and **perceived usefulness**) between conditions.

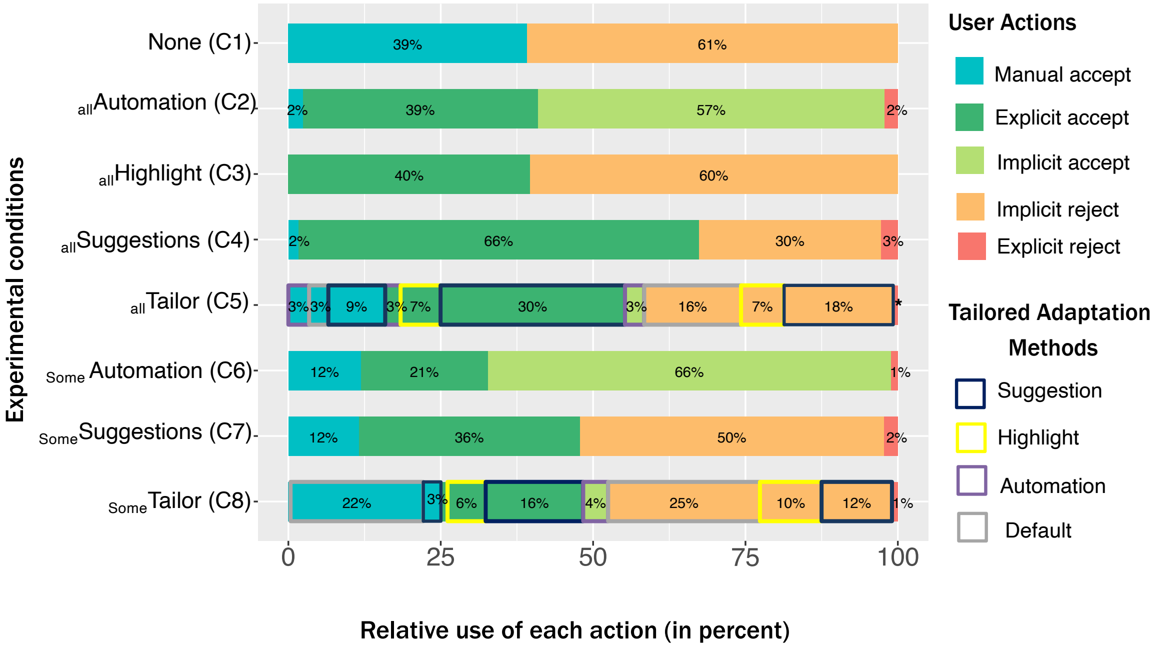


Figure 4.4: Actions taken by participants across the eight experimental conditions. The level of positive user engagement is assessed by proportion of actions that are either manual or explicit accept, while the level of privacy protection is assessed by the proportion of actions that are either manual accept, explicit accept, or implicit accept ($*$ represents action counts $< 1\%$).

4.3.1 Engagement Patterns

Figure 4.4 shows that participants rarely explicitly rejected an adaptation (by clicking "Rather Not" in a suggestion or by undoing an automated adaptation)—the prevalence of such **negative user engagement** in the conditions where it applied was only around 2%, and there were no statistical differences in negative engagement between these conditions ($\chi^2(5) = 10.756$, $p = .0564$). In the remainder of this subsection we analyze the differences in positive engagement only, and we will refer to it simply as "engagement".

We find that there are significant differences in **positive user engagement** (i.e., the sum of *manual accept* and *explicit accept*) across the eight experimental conditions ($\chi^2(7) = 97.987$, $p <$ .001). We divide our exploration of the differences in positive user engagement into four subsections: In subsection 4.3.1.1 we compare the levels of engagement in each adaptation condition (C2-C8) against the condition where no adaptions were applied (C1). We subsequently compare the levels of engagement among the conditions where all features were adapted (C2-C5, subsection 4.3.1.2) and among the conditions where awkward/irreversible features were avoided (C6-C8, subsection 4.3.1.3). We then compare the pairwise differences between the indiscriminate ("all") and selective ("some") versions of Automation, Suggestions, and Tailor in subsection 4.3.1.4, and conclude with a summary of the findings in subsection 4.3.1.5.

4.3.1.1 Suggestions and tailored adaptations increase engagement

On average, participants who interacted with the prototype that did not make any adaptations (C1) positively engaged with 39% of the privacy features. Comparing the level of engagement in all other conditions against C1, (positive) engagement is significantly higher for participants in the ${}_{all}$Suggestions (C4, 68%, $\beta = 1.30$, $p < .001$), ${}_{all}$Tailor (C5, 55%, $\beta = 0.70$, $p < .001$), ${}_{some}$Suggestions (C7, 48%, $\beta = 0.36$, $p < .001$), and ${}_{some}$Tailor (C8, 47%, $\beta = 0.39$, $p < .001$) conditions. Using the logistic regression βs to calculate odds ratios[6] ($e^{\beta} = OR$), we find that the odds of engaging with the privacy features are 3.67 times higher for participants in the ${}_{all}$Suggestions condition, 2.01 times higher for participants in the ${}_{all}$Tailor condition, 1.43 times higher for participants in the ${}_{some}$Suggestions condition, and 1.48 times higher for participants in the ${}_{some}$Tailor condition. These are small to medium-sized effects.

The differences in engagement between the None condition and the ${}_{all}$Automation (C2, 41%, $p = 0.291$), ${}_{all}$Highlight (C3, 40%, $p = 0.916$), and ${}_{some}$Automation (C6, 33%, $p = 0.94$) conditions are not significant.

These findings indicate that the level of user engagement with the available privacy features can be increased by providing privacy suggestions or by tailoring the adaptation method of the features to users' prior awareness and usage.

[6] In the remainder of the paper we skip the β-coefficients and directly report the odds ratios. Odds ratios translate to effect sizes, with the values 1.68, 3.47 and 6.71 translating to small, medium and large effects.

4.3.1.2 Among the "all" conditions, Suggestions lead to the highest engagement, followed by Tailor

In this subsection, we present pairwise comparisons of the level of engagement among the adaptation conditions where *all* privacy features were adapted (i.e., $_{all}$Automation (C2), $_{all}$Highlight (C3), $_{all}$Suggestions (C4), $_{all}$Tailor (C5)).

On average, participants in the $_{all}$Suggestions condition positively engaged with 68% of the privacy features. Their odds of engaging with the features are 3.56 times higher than in the $_{all}$Automation condition (41%, $p < .001$), 3.77 times higher than in the $_{all}$Highlight condition (40%, $p < .001$), and 1.82 times higher than in the $_{all}$Tailor condition (55%, $p < .001$). Moreover, for participants in the $_{all}$Tailor condition, the odds of engaging with the features are 1.92 times higher than in the $_{all}$Automation condition ($p < .001$) and 2.03 times higher than in the $_{all}$Highlight condition ($p < .001$). The difference in engagement between the $_{all}$Automation and $_{all}$Highlight conditions is not significant ($p = .916$).

These findings indicate that the $_{all}$Suggestions adaptation resulted in a significantly higher level of engagement than any of the other conditions in which all privacy features were adapted, with the $_{all}$Tailor condition taking second place with a significantly higher level of engagement than the remaining two conditions.

4.3.1.3 Among the "some" conditions, Suggestions and Tailor lead to the highest engagement

Namara et al. [170] recommended avoidance of the Suggestion adaptation for features that would be awkward to suggest or the Automation adaptation for features that would lead to seemingly irreversible consequences if automated. In this section, we present pairwise comparisons of the level of engagement among the adaptation conditions that avoided making such awkward/irreversible adaptations (i.e, $_{some}$Automation (C6), $_{some}$Suggestions (C7), $_{some}$Tailor(C8)).

Engagement is significantly higher in the $_{some}$Suggestions (48%) and $_{some}$Tailor (47%) conditions than in the $_{some}$Automation (33%) condition (see Fig 4.4). The odds of participants in the $_{some}$Suggestions condition in engaging with a privacy feature are 1.89 times higher than in $_{some}$Automation ($p < .001$) and the odds of participants in the $_{some}$Tailor condition ($p < .001$) engaging with a privacy feature are 1.99 times higher than in the $_{some}$Automation condition ($p < .001$).

The difference between the $_{some}$Suggestions and $_{some}$Tailor conditions is not significant (p = .913).

These findings indicate that if awkward/irreversible adaptations are avoided, Suggestions and Tailoring both significantly increase engagement over Automation.

4.3.1.4 The “all” conditions generally lead to higher levels of engagement

In this subsection, we present pairwise comparisons of the level of engagement between the indiscriminate (“all”) and selective (”some”) versions of the Automation, Suggestions, and Tailor conditions (i.e., $_{all}$Suggestion(C4) Vs $_{some}$Suggestion (C7), $_{all}$Automation (C2) Vs. $_{some}$Automation (C6), and $_{all}$Tailor (C5) Vs $_{some}$Tailor (C8)).

The odds of engagement with the privacy features are 2.46 times higher in the $_{all}$Suggestions condition (68%) than in the $_{some}$Suggestions condition (48%, p < .001). Likewise, the odds of engagement are 1.52 times higher in the $_{all}$Automation condition (41%) than in the $_{some}$Automation condition (33%, p < .001). There is however no significant difference between the $_{all}$Tailor (55%) and $_{some}$Tailor (47%, p = .0584) conditions.

These findings indicate that the “all” conditions generally lead to higher levels of engagement than the “some” conditions—the awkward/irreversible adaptations did not discourage participants from positively engaging with the privacy features.

4.3.1.5 Summary of engagement findings

To summarize the findings regarding engagement:

- At 68%, the $_{all}$Suggestions condition leads to the highest levels of engagement—higher than the other “all” conditions and its “some” variant.
- The $_{all}$Tailor (55%), $_{some}$Tailor (47%), and $_{some}$Suggestions (48%) conditions also increase engagement compared to no adaptations—these are not significantly different from one another.
- Given that the Automation adaptation operates completely outside of the user’s awareness, we are not surprised that the $_{all}$Automation and $_{some}$Automation conditions do not increase engagement compared to no adaptations (39%)—$_{all}$Automation (41%) leads to significantly higher engagement than $_{some}$Automation (33%), though.
- Surprisingly, Highlight (40%) did not increase engagement either, despite the visual prominence of the adaptations in this condition.

4.3.2 Privacy Protection Outcomes

While positive user engagement results in higher levels of privacy protection, some of the experimental conditions (e.g., the Automation conditions) result in protection even when the user ignores the privacy features. In this subsection we analyze the differences in the average amounts of privacy protection participants end up with in each of the eight experimental conditions.

We find that there are indeed significant differences in the amounts of **privacy protection** (i.e., the sum of *manual accept*, *explicit accept*, and *implicit accept*) achieved across the eight experimental conditions ($\chi^2(7) = 391.45$, $p < .001$). We divide our exploration of these differences similarly to the engagement section: In subsection 4.3.2.1 we compare the level of privacy protection achieved in each adaptation condition (C2-C8) against the condition where no adaptions were applied (C1). We subsequently compare the level of privacy protection achieved in the conditions where all features were adapted (C2-C5, subsection 4.3.2.2) and among the conditions where awkward/irreversible features were avoided (C6-C8, subsection 4.3.2.3). We then compare the pairwise differences between the indiscriminate ("all") and selective ("some") versions of Automation, Suggestions, and Tailor in subsection 4.3.2.4, and conclude with a summary of the findings in section 4.3.2.5.

4.3.2.1 Apart from Highlight, all adaptation methods improve privacy protection

In the prototype without adaptations (C1) participants are only protected if they engage with a feature. Hence, their protection is equal to their level of engagement: 39%. In contrast, protection is enabled-by-default in the $_{all}$Automation condition (C2), unless the user intervenes through an *explicit reject* action. Such actions are rare, hence the privacy protection in the $_{all}$Automation condition is virtually perfect, at 98%. Notably, although some of the privacy features are not adapted in $_{some}$Automation condition (C6), users seem to manually engage with those privacy features anyway, leading to virtually perfect privacy protection (99%) in this condition as well. Unsurprisingly, the pairwise differences between these conditions and the None condition are strongly significant ($p < .001$).

Further comparisons with C1 reveal that the odds for achieving privacy protections are 3.67 times higher for participants in the $_{all}$Suggestions condition (C4, 68%, $p < .001$), 2.01 times higher for participants in the $_{all}$Tailor condition (C5, 58%, $p < .001$), 1.42 times higher for participants in the $_{some}$Suggestions condition (C7, 48%, $p < .001$), and 1.48 times higher for participants in the

${}_{some}$Tailor condition (C8, 51%, $p < .001$). The privacy protection outcomes for the participants in ${}_{all}$Highlight condition (C3, 40%) are not significantly different ($p = 0.916$).

These findings indicate that all adaptation methods lead to better privacy protection outcomes, except for the Highlight adaptation.

4.3.2.2 A clear privacy protection hierarchy exists among the "all" conditions

In this subsection, we present pairwise comparisons of privacy protection outcomes among the adaptation conditions where *all* privacy features were adapted.

As mentioned before, the protection in the ${}_{all}$Automation condition (98%) is virtually perfect—strongly significantly higher than all other "all"conditions. Among the remaining "all" conditions, the protection odds in the ${}_{all}$Suggestions condition (68%) are 3.78 times higher than in the ${}_{all}$Highlight condition (40%, $p < .001$) and 1.82 times higher than in the ${}_{all}$Tailor condition (58%, $p < .001$). Moreover, the protection odds in the ${}_{all}$Tailor condition are 2.03 times higher than in the ${}_{all}$Highlight condition ($p < .001$).

These findings show a clear hierarchy in privacy protection, with ${}_{all}$Automation providing the highest level of protection, followed by ${}_{all}$Suggestions, then ${}_{all}$Tailor, and finally ${}_{all}$Highlight.

4.3.2.3 Among the "some" conditions, Automation leads to the highest level of protection

In this subsection, we present pairwise comparisons of privacy protection outcomes among the adaptation conditions that avoided making awkward/irreversible adaptations. The privacy protection outcomes in the ${}_{some}$Automation condition (99%) is virtually perfect and hence strongly significantly higher than the ${}_{some}$Suggestion (48%) and ${}_{some}$Tailor (51%) conditions. The privacy protection odds between the latter two did not differ significantly ($p = .913$).

These findings indicate that even when features that are awkward/irreversible to adapt are avoided, automation still affords the best privacy protection outcomes.

4.3.2.4 Some differences exist between the "some" and "all" conditions

Pairwise comparisons between the indiscriminate ("all") and selective ("some") conditions reveal that the privacy protection odds are 2.47 times higher in the ${}_{all}$Suggestion condition (68%)

than the $_{some}$Suggestion condition (48%, $p < .001$). This result mirrors the engagement results, as the Suggestion conditions do not contain an "implicit accept" option.

The privacy protection odds are 1.87 times higher in the $_{some}$Automation condition (99%) than in the $_{all}$Automation condition (98%, $p < .001$). This is is surprising: even though the $_{some}$Automation foregoes automating certain features, the overall level of protection is higher than in the $_{all}$Automation, arguably because explicit rejections are lower in the former condition and because participants manually engage with the features that were not adapted.

Finally, there was no significant difference in privacy protection between the $_{all}$Tailor (58%) and $_{some}$Tailor (51%) conditions ($p = .0584$).

4.3.2.5 Summary of privacy protection findings

To summarize the findings regarding privacy protection:

- At 98% and 99% respectively, the $_{all}$Automation and $_{some}$Automation clearly lead to the highest levels of privacy protection—this is evident by the relatively low incidence of explicit rejections.

- The fact that $_{some}$Automation outperforms $_{all}$Automation in terms of privacy protection speaks to the apparent superiority of this more prudent approach. Users seem to implement the avoided adaptations anyway, while at the same time issuing fewer explicit rejections.

- The $_{all}$Suggestions condition (68%) follows in third place, with a higher level of protection than $_{all}$Tailor (58%) and $_{some}$Tailor (51%) as well as $_{some}$Suggestions (48%).

- The $_{all}$Highlight condition (40%) performs worst, offering no significant protection benefits over no adaptations at all (39%).

4.3.3 Subjective Evaluations

In the assessment of the user subjective evaluations of the platform, we find that the **perceived decision help** and **perceived usefulness** measurement scales were highly correlated ($r = 0.858$). For the sake of completeness we include results from both scales (see Figure 4.5).

Compared to the condition where no adaptations were applied (C1), participants in the $_{some}$Suggestion condition (C7) deemed the platform more helpful ($\beta = 0.677$, $p < .001$) and more useful ($\beta = 0.677$, $p < .001$). While all other adaptation conditions were also deemed more helpful and useful than C1, none of these differences were significant.

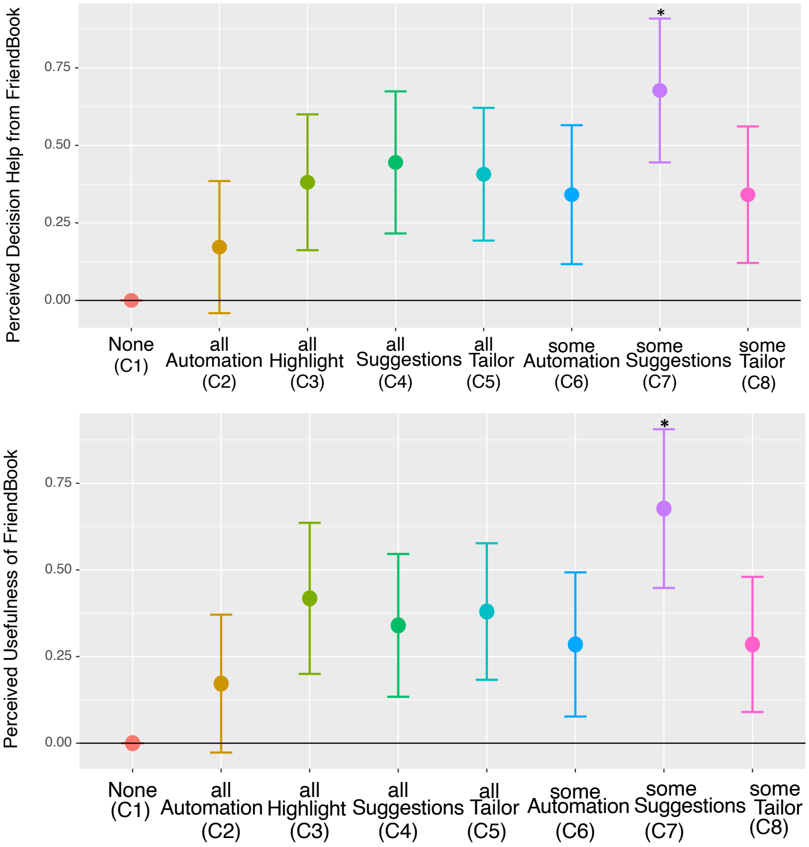


Figure 4.5: The effect of the experimental conditions on perceived decision help and perceived usefulness. Factors have no inherent scale, so their values are fixed to zero for C1, and scaled in sample standard deviations of the measured factor. Error bars are ±1 standard error of the comparison with C1. $*$: $p < .001$

4.4 Discussion

A predominance of existing work in the area of adaptive privacy has focused on accurately predicting user preferences and behaviors [121, 23, 148, 234, 230], without devoting enough effort to how privacy adaptations could ultimately be *presented*. Studying adaptation methods is particularly important in contexts where users do not expect a system to provide privacy advice or make decision on their behalf during the course of use.

In our study we used three adaptation methods identified by Namara et al [170]—Automation, Highlight, and Suggestions—and examined their effectiveness in helping users better manage their privacy on an SNS platform. In this discussion section we reflect on the effect of each of these adaptation methods on users' engagement with the privacy features, their privacy protection, and their subjective evaluations.

4.4.1 The effectiveness of the Adaptation Methods

Our results suggest that the **automation** of adaptations to privacy features towards stricter settings considerably increases the level of privacy protection afforded by the system and does not seem to negatively affect the level of user engagement with the privacy features.

Namara et al. [170] found that users were worried about the accuracy of the automation of the privacy features, and that automation would reduce their ability to make their own privacy decisions. They therefore suggested avoiding automatic adaptations that users' thought to be "irreversible". One interesting finding is that protection is high even when the automatic adaptations of such "irreversible" features is avoided: users seem to implement the avoided adaptations anyway, and may even issue fewer explicit rejections than if all features are automatically adapted.

Although the automatic adaptations somewhat improve users' perceived decision help and usefulness over the baseline system with no adaptations, this difference is not significant—perhaps because much of the protection happens outside of users' awareness. Another reason could be that some users still fear that the system might not be able to accurately capture their privacy preferences [170]. Indeed, Page et al. [181] assert that even when not adapted, some users are very concerned about how the use of privacy features (e.g.,untagging, unsubscribing or unfollowing a friend) hurts their relationships with others. Automation would only exacerbate the concerns of these users.

In contrast to Namara et al.'s [170] assertion that **highlights** might be able to unobtrusively raise users' awareness about privacy features, we found that this adaptation method improved neither users' level of engagement nor the overall privacy protection compared to the baseline system with no adaptations. The observed increase in subjective ratings were also not significant. This finding aligns with Warherberg et al.'s [227] assertion about the effectiveness of privacy nudges (e.g., the use of highlights) in influencing privacy decisions: they argue that the effects of some of nudges are fragile and potentially impractical for many applications. Perhaps, then, highlights should rather be used to convey and serve as indicators of new changes to an interface (e.g., to indicate a new notification or as chat/online status indicators [49]) rather than a privacy nudge or adaptation method.

Presenting adaptations to privacy features as **suggestions** results in the highest levels of engagement and relatively a high level of privacy protection. Users also found suggestions significantly more useful and helpful, but only in the condition where awkward suggestions were avoided. Namara et al. [170] assert that users appreciate suggestions as a means to increase their awareness about a privacy feature, or as a convenient shortcut to apply an adaptation without having to navigate to the feature. Our work shows that suggestions are indeed effective at increasing user engagement with privacy features, which in turn improves their privacy protection.

Namara et al.'s [170] key recommendation was that adaptation methods should be **tailored** to users' awareness and prior use of the privacy features. We find that the tailored conditions increase users' engagement (but not as much as suggestions) and protection (but not as much as automation). The tailored conditions do provide an interesting blend of *manual accept*, *explicit accept*, *implicit accept* and *implicit reject* outcomes, with very small incidences of *explicit reject*. Perhaps tailoring the adaptation methods could help strike a balance between the convenience of automation and the engagement of suggestions while avoiding their potential threats of loss control and undue burden, respectively.

4.4.2 Design Implications

Our results show how a variety of privacy adaptation methods can significantly improve upon the traditional SNS privacy features in different ways. Hence, which adaptation method is "best" for a certain SNS platform depends on what the designers of the platform want to accomplish? We argue that one important goal of providing privacy adaptations is **to improve users' privacy protection without causing undue burden**. In this light, we find that the **automation** of privacy feature adaptations affords users the most privacy protection without increasing or decreasing their engagement.

Whereas automations are inevitability executed by the system and can occur without explicit notification of the user, Markus and Reinhardt [215] assert that restrictive default privacy settings do not change users' perception and enjoyment of a system (e.g., social media platform). This suggests that once users realize that an automated privacy action was executed by system on their behalf, this is not likely to change their perception about the platform. Instead, the increased privacy protection outcome is likely to alleviate their privacy concerns [215]. Thus, we recommend that if the system's objective is to drastically increase user privacy, automation or restrictive default settings should be adopted.

To decide on what features to automate, we recommend that developers automate features that would not result into unintended consequences for the user [170, 181]. We observe that avoiding the automation of certain seemingly "irreversible" privacy features does not reduce privacy protection (i.e., users will simply engage with those features manually), and may even increase protection as it makes users less likely to reject any of the adaptations.

Another important goal of providing privacy adaptations is to **encourage active ownership over one's privacy** by increasing user engagement with the available privacy features. Liu et al.'s[150] show that there tends to be a mismatch between SNS users' desired privacy settings and their actual settings, with 36% of content on social media being shared with the default settings. Our results suggest that the provision of well-timed **suggestions** can help remedy this mismatch and provide an opportunity for users to learn about the available privacy features. Under these circumstances, suggestions could be considered as a way to inform or remind users about the available privacy features in a system and the possible actions users can undertake to achieve their desired privacy setting/level. By proactively guiding users on how to appropriately safeguard their privacy, suggestions ultimately help users improve their own privacy whilst using the platform (cf. [225])—even though the protection improvements of suggestions are not as substantial as those of automation.

In line with Namara et al [170], we find it beneficial not to make suggestions for features that users would consider awkward. Although this did somewhat reduce protection and engagement (from 68% to 48%), this strategy did result in improvements in perceived decision help and perceived helpfulness—in fact, it was the only condition in which these improvements were significant.

Suggestions should also be well designed and timed. In a computer security context, Vance et al. [217] warn that constant provision of notifications is prone to habituation, which suggests that over time, users would likely stop paying attention to the suggestions. One solution would be to make the privacy suggestions stand out (with a different look and feel) from other suggestions/notifications furnished by the platform [217]. In our context, we used a virtual character ("the privacy dinosaur") to increase the salience of the suggestion and to make it more endearing.

Finally, our results show that **tailoring** adaptations to users' privacy preferences can help **strike a balance between user engagement and privacy protection**. The effect of tailoring is dependent on a wide range of parameters, so future research should further investigate how this can pragmatically be achieved.

4.5 Limitations and Future Work

This research was primarily motivated by the earlier works of Namara et al. [170] and Colnago et al. [50]. We leveraged their insights in the development of adaptive privacy features within a working prototype of an SNS platform and examined the effects of their adaptation methods on the level of user engagement and overall privacy protection outcomes.

For experimental control purposes, we put people in the same scenario (i.e., the specific task of having participants "clean" their social media page), having the same goal towards managing their profile (i.e., in preparation for a job search). Thus, we developed a semi-functional working prototype of an SNS platform with a fictitious profile to create an experience that was the same for all participants (safe for the adaptation method). We are cognizant that participants interactions, decisions, and subjective experiences are susceptible to the design of the site [208]. Indeed, participants may have behaved differently in our prototype with another person's profile than they would on their preferred SNS using their own profile. We made the interaction with our prototype as realistic as possible to mitigate this reduction of ecological validity needed to create a feasible and carefully controlled experimental setup.

SNS platforms typically contain a plethora of privacy features. To make our study more manageable, we adopted 13 privacy features that support some of the most common privacy behaviors on SNS platforms as catalogued by Wisniewski et al. [242]. We ensured that these features kept the same core functionality as those on Facebook. As one of the goals of privacy adaptations is to support users in navigating a deluge of privacy features, we conjecture that an increase in the implemented privacy features would only strengthen our findings regarding the positive effects of the proposed adaptations.

Additionally, privacy features on social media platforms are used over time and in different contexts. In our study, we used a job scenario to motivate users to explore, engage, and review "their" profile. Whereas the scenario helped implore and provide rationale for users to partake in our study; users may have acted differently if this was their real profile and had used it overtime. As such, we are cognizant that this scenario (i.e., having participants "clean" their social media page in preparation for a job search) was a very particular context, so the behaviors may not be 100% representative of participants' day-to-day social media use.

Future work should investigate some of our surprising results, such as why highlights did not increase user engagement, despite their visual prominence. One could argue that the highlight color or size were not prominent enough to incur curiosity among users. Alternatively, users could have ignored the highlights due to a lack of explanation as to why certain privacy features were highlighted.

Finally, the design teams of social networking sites like Facebook can replicate our findings in a real-world setting, thereby investigating the feasibility and effectiveness of using the proposed adaptation methods to improve the privacy of their own social media profiles.

4.6 Conclusion

This chapter highlights the effectiveness of three adaptation methods—Automation, Highlights and Suggestions—in improving user engagement and overall privacy protection on a modern online technologies such as an SNS platform. Our findings reveal that automation of privacy features affords users the most privacy protection, while giving privacy suggestions significantly increases their level of engagement with privacy features and improves their perceptions of helpfulness and usefulness (as long as awkward suggestions are avoided). Similarly, my work in Chapter 3 demonstrates that privacy suggestions are preferred as a means of teaching users privacy features that they were unfamiliar with. These findings point to the opportunity of leveraging "privacy suggestions" to communicate recommended privacy behaviors when users need to review their privacy settings or are not aware of existing privacy features. In the next chapter 5, I examine the appropriate framing (i.e., tone) that privacy suggestions should embody to encourage users to "better" manage their social media privacy with regard to how they feel about their privacy (i.e., privacy-related privacy).

Chapter 5

The Influence of Privacy Suggestion Tone & Privacy-Related Affect on Adaptation-Supported Privacy Decision-Making.

The studies described in Chapter 3 & 4 find that privacy suggestions are the preferred adaptation presentation method and significantly increase social media users' level of awareness and engagement with privacy features. More importantly, privacy suggestions encourage users to take active ownership of their privacy. Depending on the recommended action, such privacy suggestions would serve as intelligent and convenient means to inform, remind or educate the user about existing privacy features and afford them control to make "better" privacy decisions [170]. In essence, privacy suggestions serve as personalized shortcuts and means to implore users to take privacy actions without necessarily navigating through layers of hidden menus [170]. However, for the successful provision and acceptance of such privacy suggestions, research related to intelligent agents [46, 94, 50] and privacy decision-making [4, 14, 24] reveals that the message framing (i.e., the way an option or information is presented to the user) ought to be carefully considered. Otherwise, users are more likely to feel resigned or detached from the recommended actions [50]. This resignation and psychological reluctance to follow recommended privacy actions arise because privacy risks do not always seem tangible, their negative consequences are not immediate, and as a result are not always at the forefront of social media interactions [78, 8]. Research shows that users can instead refuse to follow through on the recommended action based on the way the message is conveyed (e,g., based on the tone used to describe the recommended action) and its' relatedness to the decision context [94, 57, 46]. Within the context of presenting privacy adaptations, this raises an important question: **what framing (i.e., tone) should privacy suggestions embody if they are to encourage users to "better" manage their social media privacy?** Understanding the appropriate tone to use will help improve the effectiveness of privacy suggestions in enhancing user engagement with privacy features, urgency, and reaction to helpful privacy tips or information and improve trust in the platform. Consequently, such "appropriately" toned privacy suggestions will alleviate the burden inherent with privacy decision-making and enable users to set their desired level(s) of privacy [160, 231, 174, 162].

Prior work[1] reveals that privacy decision-making is a highly complex process affected by several factors that range from the "privacy "calculus“ to [affect]; from asymmetric information to bounded rationality; and from resignation and learned helplessness to cognitive and behavioral biases." [6, p.741]. Together, these factors help explain and influence the decision-making techniques users employ to make privacy-related decisions [4, 162]. In the making of privacy-related decisions, several scholars reveal that users rely on heuristic rather than analytical/ systematic assessment of the availed privacy choices [73, 4]. Heuristics are "automated cognitive processes that circumvent the conscious deliberation of information" [108, p.564]. However, research shows that heuristics are susceptible to elements such as message framing and user affect—the emotions or feelings that a user might experience/display [143, 142, 194], especially in the context of new information or evolving situations [108, 65, 73]. For example, in the making of social media posting decisions, when users were presented with persuasive cues (i.e., a justified reason for the user to disclose information by giving a reason why it would be better to disclose or appeal to the social norm by displaying what others had done) as decision aid heuristics, Ferwerda et al. [65] found that such cues affected users' decisions based on their framing. More specifically, users were less sensitive towards positive guidance for posting, and more sensitive towards negative guidance. In instances when there were still in doubt of hurting their self-presentation, they erred on the safe side of their posting behaviors

[1] see Chapter 2;Section 2.4 for an extended review

by not posting content at all [65]. Additionally, Anaraky et al [73] also found that when it came to revealing personal information to a financial application, young adults tended to rely on their affect heuristic of trust (i.e., positive or negative feelings based on the trust they had in the app provider) primarily as a way to infer about the privacy sensitivity of their data and subsequently the privacy risks associated with its disclosure. These research works show the importance of examining the impact of user affect, trust, and framing (i.e., linguistic tone in the case of our study) on users' privacy decision outcomes, especially if we are to understand and aid users' in their privacy decision-making process. More significantly, if social media platforms are to provide personalized privacy suggestions that effectively encourage users to exact more control over their social media privacy, an "appropriate" tone is essential to modulate between a users' privacy-related affective state and motivation to make privacy decisions, even when in the instances when decision heuristics that users rely on break down or are incorrectly applied [97].

Therefore, in this chapter, I conducted an experimental study to systematically understand the impact of privacy-related affect and what tone privacy suggestions should embody if they are to more effectively encourage users to "better" manage their social media privacy. The primary goal was to examine the privacy suggestion tone style that would "better" encourage users to engage with privacy features to achieve their desired privacy, considerate of users' feelings about social media privacy (i.e., privacy-related affect). Furthermore, I wanted to assess the impact on users' experience with the platform (i.e., perceived decision helpfulness and trust in the platform) based on the engagement with the provided privacy suggestions. I manipulate the privacy suggestion tones and use priming to put participants into various (privacy-)affective states to evaluate the most appropriate framing (i.e., tone) for each and ultimate privacy decision outcomes.

We found that the examined three varying privacy suggestion tone styles (i.e.,*neutral*, *passive*, *assertive*) indeed influence users' privacy decision outcomes. However, the nature of the effect significantly differs based on users' pre-existing privacy (or lack thereof) induced affective states, i.e., the mood a user is in before the actual privacy protection decisive situation occurs. In particular, we find that the "appropriate" privacy suggestion tone style to use largely depends on the users' privacy-related affective state. For instance, when users are in a positive privacy-related affective state, the neutral tone tends to work better at encouraging them to make "better" privacy decisions. In contrast, the assertive tone tended to work best when users were in a negative privacy-related affective state. Furthermore, we observe that the tone that privacy suggestions embody not only influences users' behavior regarding these suggestions and/or the privacy actions they recommend; they also impact users' other privacy actions (i.e., actions that are not subject to suggestions by the platform), indicating that tone has a robust, system-wide effect. These findings suggest that considering users' privacy-related affect (i.e., how users feel about their social media privacy) is crucial in determining the tone style that system designers can use to craft and present personalized privacy suggestions. Overall, these results provide a better understanding of the impact of privacy suggestion tone styles on users' privacy decision outcomes in light of their privacy-related affect. This study provides several insights towards the advancement of the *presentation* of user-tailored privacy adaptations and or personalized privacy systems.

5.1 Chapter Background

In the qualitative study detailed in Chapter 3, participants indicated that privacy suggestions provide an opportunity to educate, persuade, and encourage them to take active ownership or control of their social media privacy. In particular, participant H stated that *"I think [privacy suggestions] would be helpful if they gave a [clear] reason"* for the suggestion. However, in the same line, he pointed out that *"the tough part [with providing privacy suggestions] would be counting on the person to follow through"* [170]. In a different IOT context, Colnago et al [50] found similar user sentiments and perceptions about the potential pitfalls of personalized privacy assistant recommendations. So, how should social media users be implored to follow through and make appropriate privacy decisions when they are provided with suggestions or recommendations?

Prior work on privacy decision-making, adaptive or intelligent systems assert that a primary influence on users' privacy-based decisions depends on the structure or wording of the privacy choice (termed the "suggestion" or "recommended action" in this chapter) [4, 36, 14, 235]. One of the important language features in the wording of such privacy choices is tone [153, 96, 212]. Tone helps modulate the voice of the recommended action and has been found to significantly affect users' experience with intelligent or adaptive computing systems [153, 94, 46]. More specifically, the tone embodied by a message can encourage or de-motivate users from taking the recommended action [212].

In our studies described in the Chapters 3 and 4, we did not account for the effects of tone on users' privacy decisions, when presented with a suggestion. Instead, we presented users with a privacy "suggestion" adaptation (see Chapter 3, Figure 3.3) that proactively recommended a privacy action using a positive framing that nudges them towards taking action (i.e.,"I think you should...[take the disposed action]"), coupled with the options to accept ("*Ok*") or reject ("*Rather Not*") the recommended action. While, in the context of these studies, this type of wording/framing was appropriate, it failed to account for tone which would have helped to accurately convey the value and urgency in taking such recommended action [212]. Hence, it is against this background that in this chapter, we seek to identify and understand what privacy suggestion tone style would effectively encourage users to manage their social media privacy. As such, the objective of the study is to answer the the following research questions:

RQ1: How do three different privacy suggestion tone styles (i.e., neutral, passive, assertive) affect users' privacy decisions?

RQ2: How do the three different privacy suggestion tone styles influence users' experience with the platform (i.e, perceived decision helpfulness, and trust in the platform)?

RQ3: Does the effect of privacy suggestion tone styles depend on a user's privacy-related affective state?

5.2 Related Work

In the following subsections, we review the related work on the impact of tone styles on behavioral change and decision-making, first in general and then explicitly within intelligent and personalized privacy systems. We also review the impact of user affect and trust on social media disclosure behaviors.

5.2.1 The impact of Tone on Decision-Making

Tone—defined in this work as the language style used within privacy suggestions to encourage users to take recommended actions towards the safeguarding of their social media privacy—influences peoples actions, attitudes, and opinions [174, 94, 96, 155]. While less studied within the privacy field, research in the psychology, persuasive communication, and intelligent computing disciplines has repeatedly shown that the tone embodied by the message (or information) used to persuade people to take certain actions significantly influence whether an individual is receptive and takes steps towards achieving the desired effect [243, 167, 153, 131]. More specifically, the tone of a recommended action can increase users' trust and motivation to take action or make changes [212, 243, 160, 55]. Literature within the field of persuasive computing, asserts that using an appropriate tone would help users' feel like the computing system cares about them and wants to help them understand it (i.e., is willing to meet them where they are) [174, 46]. In health and conservation communication, tone has been shown to affect users' attention, motivation, and inspiration [107, 167]. These influences are the basis for the ability of the given message to change people's actions, attitudes, opinions and ultimately behaviors [174].

Several researchers have examined the different effects of tone styles on users' behaviors and actions, especially within the contexts of environmental messaging and public health [107, 167, 243, 212]. The majority of this research work reveals an inconsistency regarding the effects of particular linguistic tone styles (e.g., neutral vs. passive/suggestive vs. assertive tones) on users' receptivity, preference, and ultimately behaviors [212, 167, 107]. For example, within the context of mental and public health, Muench et al. [167] found that individuals were sensitive to the variation in the linguistic tone of mobile delivered health-related messages designed to help them achieve a personal goal (e.g., reduce alcohol consumption). In some cases, participants had clear preferences for one type of tone over another. More specifically, 75% of the participants were found to prefer messages that were "grammatically correct, free of textese, polite, nonaggressive, and directive as opposed to passive" [167, p.1]. Thus, subtle manipulations of the tone style of the message, such as changing ("Try to...") to ("You might want to try to...") were found to have significant effect on the user preference for the message and their intention to act. Within the context of running successful water conversation campaigns, Kartz et al. [107] compared the impact of assertive (e.g., "You must conserve water") and suggestive (e.g., "Please consider conserving water") linguistic tone style variations on individuals' residential water conservation behaviors. They found that suggestive messages were better than assertively phrased messages at encouraging users to take action (i.e., changing their residential water conservation behaviors). In the context of prompting behavioral change (e.g., reducing alcohol consumption) among university students, Thomas et al. [212] found that students were more respective and likely to engage with messages that were neutral and clear (i.e., based on facts and balanced in both their positive or negative framing). Overall, the majority of intervention-based messaging research work in these areas shows that variations in linguistic tone styles of the message content can affect message receptivity, preference and ultimately behavior.

More importantly, this research work shows that there is no global or universal preference for one tone over the other. Instead, there are circumstances under which the tone style used can either successfully motivate or fail to produce the desired effect, and lead to results that are at odds with the intent of the message [55]. As such, researchers have recommended tailoring message tone styles based on the user context and desired effects [212]. To this end, little is known about the ways in which tone would affect privacy-related decision-making, especially within the context of personalized privacy systems.

To uncover the potential impacts of tone on users' privacy-related decisions, we review research on the effect of linguistic tone styles within intelligent systems (such as chat-bots, virtual avatars, self-driving cars) — which is also somewhat limited [50, 94, 39, 160]. For example, in examining interface presentations of chat-bots within the context of social media customer care, Hu et al. [94] emphasized the importance of considering the tone that the chat-bot embodies, given their far-reaching effects. In particular, the authors found that different tone styles that the chat-bot might embody can have significant influences on the user experience, attitude towards the system, and their assessment of the quality service of social media customer care they receive. For example, an empathetic or passionate tone can lead to more user trust in the chat-bot, and reduce user stress [94]. Wilkinson et al. [235] go deeper and demonstrate that even the justification style a chat-bot uses to explain or justify its actions, can impact users' perception of system transparency, perception of control, trust and willingness to depend on the system's advice. Similarly, in the context of online intelligent agents (i.e., online avatars), Brave et al. [39] further assert that an agent that uses an empathetic tone can lead to greater likability and trustworthiness. In the context of using virtual agents to promote psychical fitness, Lucas et al. [153] found that agents that embodied an affective tone were able to successfully motivate users to embark on maintaining pyhsical fitness. Indeed, Martelaro et al. [160] found that an intelligent agent that embodies a tone that engenders its vulnerability and expressivity, can encourage user trust, disclosure, and feeling of companionship. Additionally, within the context of self-driving cars, Wong et al. [243] found that assertive voices were more effective in grabbing people's attention, especially when engaged in an immersive task. In this context, people pay attention not only to *what* the message says, but also *how* the message is said. In a similar context, Neirbuhr and Michalsky al. [174] also found that a persuasive charismatic (i.e, more empathetic) tone style has far reaching influence on people's opinions and actions than a less charismatic one. Taking a more nuanced look at the insights from the above examples reveals that it is not super universal which tone is best for use even within intelligent systems. Instead, it appears that the receptivity, preference, and effectiveness of the linguistic tone style of messages, even within intelligent systems, depends on the decision context/situation.

Why is it important to understand the impact of linguistic tone styles within the context of personalized privacy systems? In their work examining appropriate implementations of supportive designs for personalized intelligent privacy assistants, Colnago et al [50] assert that recommended actions should be presented in a clear and informative way. Otherwise, users are likely to feel resigned or detached from taking the recommended action. Similarly, Liao et al [145] assert that such intelligent systems/agents have to be straightforward in how they assist users, otherwise users are less likely to be persuaded or to trust them. The authors note that this is very important, especially when the intelligent agent relies on users' personal information to foster informed privacy decision-making. Prior work on privacy decision-making also asserts that the framing of the privacy choice (i.e., the suggested actions) impacts user privacy-related decision outcomes [24, 14, 4]. Whereas the framing in this realm mainly refers to the (positive vs. negative) structure of privacy decision choices instead of the linguistic tone style, it reveals that framing significantly influences users' decision-making processes and outcomes [24, 14]. For example, within the context of privacy decision-making in IOT, Bahirat et al [24] found that privacy message framing can reduce the amount of deliberation users expend to arrive at a decision that appropriately reflects their evaluation of the context of the decision. The authors found that a positive framing of decision choices reduced the likelihood of information disclosure. Herein, the positive framing of a decision choice was likely to help users focus on the particular aspects of the decision context. More specifically, a positive framing was more likely to help users focus on the expectedness, appropriateness, and usefulness of the decision context (or lack thereof), and less likely on whether they were comfortable with the decision context and/or whether they found it risky. In the context of social media, Anaraky et al. [14] found that different framing and default settings of a photo tagging request on Facebook could influence the tagging rate among participants. More specifically, a "positive normative cue could boost tagging rate in combination with positive framing." [14, p.5]. Furthermore, the authors found that providing a rationale or justification for the decision choice was not effective, echoing Knijnenburg et al.'s [123] findings. In the context of virtual agents, Lucas et al [153] found that agents that embodied a positively framed motivation message were able to successfully motivate users to embark on physical fitness than those that used negatively framed messages. Taken together, it appears that if personalized privacy systems are to effectively aid users' in their privacy decision-making process, the presentation and framing (i.e., the linguistic tone in the case of our study) of the privacy choices has to be carefully studied.

5.2.2 The Effect of User Affect on Social Media Disclosure

Privacy decision making is a dynamic process prone to external factors such as affect that can alter user perceptions, physiology and ability [142]. User affect is the emotional mental state of activation that "arises from appraisals of events or one's thoughts" [22, p.1]. Affect is elicited by relevant external events that affect a person but can also emerge from the "interaction of an event's actual or anticipated consequences and the subject's concerns."[70, p.6]. Affect has a profound effect on user behavioral responses to privacy risks by mediating users' cognitive evaluations of the risk [15, 17]. The limited set of studies that have examined the role of affect on user's privacy behaviors, find that user affect can serve as an input in the decision making process or can serve as shortcut to decision-making [15, 109, 110, 47]. More specifically, the affect-as-information theorem asserts that when making an evaluative judgement, individuals tend to ask themselves about "how they feel about the object/action?" before subsequently tapping into their present feelings to form judgement [48]. Herein, if the present feeling happens to be positive, then the decision maker's evaluations of specific options are likely to be relatively positive, and vice versa for negative feelings [48, 47]. In most cases, affect evokes immediate responses towards taking meaningful action, and can serve as an underlying basis for motivation to undertake action [60, 77, 56]. Therefore, user affect can also be thought of as one's state of feeling or how one feels when performing some task, action or activity [72]. Overall, user affect is an important factor to consider when examining users' privacy decision-making process, as it plays an essential role users' decision outcomes and perception among other functions [159, 51].

In trying to understand the different ways in which affect (i.e., emotions) influence the decision making process, Loewenstein & Lerner [152] affirmed that there are two main ways in which affect enter the process: 1)*as expected emotions* or 2) *as immediate emotions*. Expected emotions are experienced as a result of the anticipated or absolute consequence of the decision itself. These emotions might be reflected through changes in the visceral influence on behavior [26]. Examples include regret and disappointment. Loewenstein & Lerner [152] note that a major shortcoming of factoring *expected emotions* into the decision-making process is that many, if not most, of the consequences of the decision occur in the future. Hence, the expected emotions that an individual experiences might differ from those that prevailed when the actual decision was made. On the other hand, *immediate emotions* are experienced at the moment of decision making. Loewenstein & Lerner [152] note that immediate emotions are capable of (direct or indirectly) propelling decision-making. However, since the sources of these emotions can be present in the environmental stimuli (such as good smell, beautiful sights), or can be encompassed if a person's mood or temperamental disposition is perturbed [26], they are capable of propelling behavior in directions that are counter to self interest [152]. Nevertheless, they are essential in the examination of the underlying influence of user affect (i.e., emotion) on behavior or decision making [152]. For this reason, in this study, we examine the influence of affect on users' privacy decisions as *immediate* rather than *expected* emotions.

Researchers that have attempted to understand the influence of affect on decision making, have treated affect as a uni-dimensional (e.g., positive/negative) or bipolar (e.g., joy/fear) construct [16, 109, 142]. In their review of the different conceptualizations of affect, Nathienal et al. [175] found that the—uni-dimensional approach—is the most predominant way affect is categorized. In particular, the dimensional approach caters to the notion that individuals can be highly activated and be pleasantly ("positive activation") or unpleasantly ("negative activation") engaged in an experience. Watson and Tellegen's affirmed and termed this uni-dimensional conceptualization of affect as "positive " Vs "negative" [210]. They assert that positive affect and negative affects are relatively independent and hence can be assessed separately. Positive affect includes emotions such happiness, awe, desire, joy while negative affect includes emotions such as anger, sadness, fear, disgust [99, 75]. Additionally, Barclay and Kiefer [27] strongly encourage the inclusion of both positive and negative affects within the same study, as they may have different effects on users' behavior [27]. For this study, we use the uni-dimensional categorization of affect to differentiate between users' positive and negative affect about social media privacy (or lackthereof).

Additionally, user affect has also been shown to play a role in how individuals make decisions, mainly when presented with persuasive messages (i.e., meant to encourage them to change their attitude or adopt new behaviors concerning a particular issue) [54, 219]. In particular, affect provides information that can influence the acceptance of a persuasive message [54]. Herein, individuals' affective states "exert powerful information influences on the kind of information people selectively access and use when constructing a response to a [certain] situation" [69, p.514]. For instance, individuals in a positive affective state (e.g., happy people) may recall and use more positive information to enact a response (or make decisions). In contrast, those in a negative affective state might rely on more negative information [69]. Forgas [69] asserts that negative affect can promote systematic and elaborate processing of the received persuasive message, which in turn, can result in a more accurate judgment of the message. In constract, positive affect can promote a more abstract and constructive processing style that increased the incidence of message judgment distortions. Nonetheless, these affect influences are susceptible to the framing and structuring of the persuasive message [219]. For example, Riet et al. [219] found that a positive(gain)-framed information resulted in higher levels of information acceptance and attitude change than negative(loss)-framed information, an effect that was mediated by positive affect. In other words, participants in a positive affective state were more likely to accept and change their attitude when the message was positive than negatively framed, suggesting that affect does play a role in the persuasion process. Affect can help explain the underlying mechanisms of message framing effects. Dillard et al. [54] also found that when individuals were shown public service announcements (PSA's), their affective states influenced their judgement in terms of the perceived effectiveness of the PSA message (i.e., based on the assessment of the information), which, in turn influenced their attitude and behavior towards the particular issue at hand. Drawing from this research work, we examine the effects of user affect on user perception (of decision help and trust in the platform) and privacy behavior, given the varying tones styles embodied by the presented privacy suggestions in our study. The message tone is known to help modulate a system's voice to account for the affective state of people using it, but work on the underlying mechanisms of affect effects remains limited [46].

Finally, rather than examine for effects of affect in general, in this study, we focus on its effects as specifically related to privacy (i.e., *privacy-related affect*). This focus is important because prior work reveals that the effects of affect are also dependent on the personal relevance of the decision or task at hand [68, 71]. In particular, Forgas et al. [68] found that people in negative affective states exhibited more efficient decision strategies when dealing with information or situations that were personally relevant to them. Similarly, Garg [71] also found individuals in these negative states tended to engage in thoughtful and detail-oriented processing of cognitive tasks if the task at hand was relevant to them [71]. Whereas the research that has examined the specific role of privacy-related affect remains limited, few prior works have attempted to highlight the potential role of affect in privacy-related decision-making. For example, Pengnate & Antonenko [185] and Kehr et al. [109] have respectively shown the influence of momentary affective states on privacy assessment in the health and mobile app environment: consumers underestimate the risks of information disclosure when confronted with a user interface that elicits positive affect. Jung and Park [104] found that affect aroused by privacy concerns falls into three major forms: anxiety, anger, and disappointment, which further lead to different types of coping behavior in response to privacy threats. Kehr et al. [111] also found that individuals in positive affective states (e.g., happy) perceived lower situation-specific privacy risks compared to those in negative affective states (e.g., fear). Lerner & Keltener [137] found that individuals in negative affective states tend to choose the "sure thing" as their affective state (e.g., fear) activated higher estimates of the likelihood of risky events occurring. Nonetheless, Johnson and Tversky et al. [101] also highlight that people can still make judgments and decisions congruent with their affective state, even when the subject matter is unrelated to the cause of that state. As such, we also include general affect (i.e., *non-privacy-related affect*) as a control condition.

5.2.3 The Effect of User Trust in the Platform on Social Media Disclosure

Research on computer-mediated communications asserts that social media is a contextual based media, thus the effects of tone and message relatedness can influence user attitudes and perceptions (e.g., user trust) of platforms [57, 146]. Furthermore, prior work in privacy-decision making that has examined online social exchanges, boundary management, and self disclosure finds that privacy-related trust in the platform is a principal antecedent to privacy decision-making (i.e., information disclosure) [59, 103, 157]. According to Malhotra et al. [157], trust helps to bridge the tension between the platform's providers' need/use of personal information and the users' privacy concerns about disclosure of information. Consequently, trust increases the confidence users have in the platform, which lowers their perceived risk of disclosing personal information, and ultimately increases the likelihood of users engaging in information exchanges [100, 209]. In this light, trust encompasses individuals' willingness to depend on or be vulnerable to the specific online technology provider, especially when the technology is essential or needed to complete tasks [134, 113]. When users do not have enough trust in the platform, they either refrain from use or are reluctant to share information [91]. For example, Krasnova et al. [130] found that while the perception of privacy risks can create a barrier to information disclosure on SNS, such perceptions can be mitigated by users' trust in the service provider and availability of control options. Bergström [32] investigated how socio-demography, internet experience, trust and political orientation altogether influence online information disclosure: the results showed that trust (in the technology) among others is the single most important factor explaining privacy concerns toward using online modern technologies and applications; the higher the trust in the technology, the less concern individuals had about the potential misuse of their personal information.

Research also reveals that user trust can be composed of an affective component that is based on one's positive and negative feelings [139, 140]. In other words, positive and negative affective states can shape user trust [140]. Lewicki & Brinsfield [139] assert that such affective states tend to dictate the levels of trust, while Scholz & Lubell [200] argue that this kind of trust helps streamline the disclosure process. Anaraky et al [73] also reveal that trust as an "affect heuristic" can shape individuals' privacy risk perceptions and guide decision-making. Based on this premise, it remains unknown how *privacy-related affect* (i.e., users' feelings about social media privacy) would affect user trust in the platform, and in turn, disclosure. Therefore, in this study, we also examine ffor the effect of privacy-related affect on user trust in the platform.

Finally, although a great deal of research has established trust in the platform is an essential antecedent of online disclosure [157, 100, 113, 134], research that examines the effect of language features such as tone style on users' trust remains rather limited [160, 94]. In this study, we expound on the effect of user trust in the platform on users' privacy protection decisions, in light of privacy suggestions with three varying tone styles.

5.3 Research Framework & Hypotheses

Drawing from the literature summarized above, privacy suggestions may be beneficial in situations where social media users' need a great deal of awareness and assistance in the management of their privacy [169]. Such adapted privacy suggestions can help convey to users that they can depend on or trust the system to behave in their best privacy interests. However, users are likely to engage in a privacy calculus where they independently weigh the benefits Vs. costs of information disclosure before making a decision [4, 127, 14]. Herein, the tone style that privacy suggestions embody could influence whether users' find the provided suggestions compelling or helpful enough for them to follow and make ‘appropriate” privacy decisions using the available privacy features [50, 145, 94]. Furthermore, the impact of such privacy suggestion tone styles on users’ privacy decision outcomes might also differ based on pre-existing pre-existing user affect (e.g., an individual’s feeling about the decisive situation at hand) [142, 97, 111]. Hence, to better understand the impacts of different privacy suggestion tone styles on users’ privacy decision making process, we also examined the influence how one feels about social media privacy—termed as the *privacy-related affect* in this study.

Figure 5.1 depicts our research framework that shows the proposed hypotheses and summarizes the core constructs underlying our research. More specifically, our research framework proposes that: (a) depending on users’ pre-existing privacy-related affect, the three varying privacy suggestion tone styles can have different impact on users’ perceived decision help and trust in the platform; (b) depending on users’ pre-existing privacy-related affect and users’ sated general informational privacy concerns can have direct impact on their privacy protection decision outcomes; and (c) the direct effect of the three privacy suggestion tone styles could be mediated by user perception of the decision help and trust in the platform.

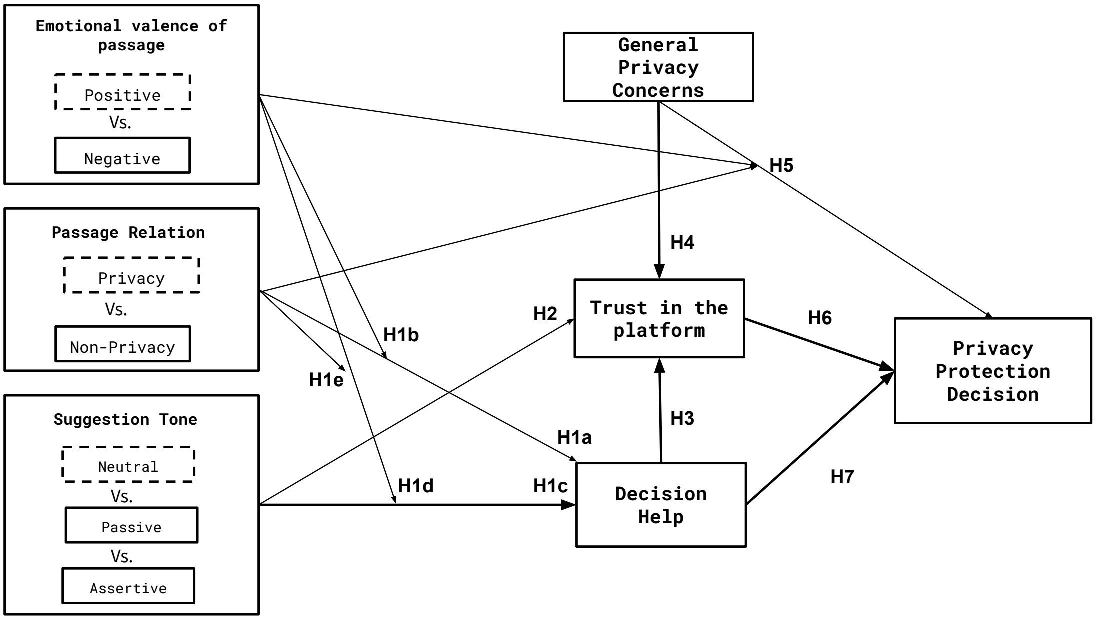


Figure 5.1: The research framework to examine the influence of pre-existing privacy-related affect, privacy suggestion tone styles, on decision help, trust in the platform, and privacy protection decision outcomes.

5.3.1 Operationalization of the Research Framework

For the experimental set up of our study and operationalization of the above mentioned framework (see Figure 5.1), a story or “passage” was used to prime and induce either a *positive* or *negative* affect that was either *privacy* or *non-privacy* related amongst participants (see Figure 5.2 for more details) [63, 71]. More specifically, to induce *privacy-related affect* amongst participants, we ensured that the passage was *related* to social media privacy (e.g., the unwanted access or use of disclosed social media information in the hiring process) and differed in *valence* by either highlighting the advantages (i.e., upsides) or disadvantages (i.e., downsides) related to the use of such personal information in making hiring decisions (see Section 5.4.3 for further details). On the other-hand, *non-privacy related affect* was also induced to serve as a control. This non-privacy related affect control condition would ensure that any observed effects were not as a result of affect (in general) but rather unique to privacy-related affect. As such, the passage used in this instance did not make any reference to the use of social media personal information in the hiring process. Taken together, the *emotional valence of the passage* served as a primary dimension to help compare between the effects of positive Vs. negative affective states [71]. While the the *passage's relation* to social media privacy helped serve as a primary dimension compare between the effects of privacy Vs. non-privacy affects.

Additionally, a common approach to vary or differentiate tone styles involves changing the linguistic content to showcase the urgency or the forcefulness of the message to persuade people to alter their behaviors [76]. For instance, assertive messages typically employ a commanding tone such as “You must do [X]” whereas non-assertive messages use a gentler approach, as in “Please consider doing [X]”[107]. Assertive messages use imperatives such as “should”, “must”, ‘ought” rather than the gentler imperatives “could”, and “might” to indicate the degree to which an individual is obligated or has the option to refuse to take action, [114, 41]. As an example, a message with a gentler non-urgent tone could read as (e.g., “It could be helpful to think about what you will lose if you give up on your goals.”) [167]. Therefore, for our study, we identified and defined three possible privacy suggestion tone styles as: 1) **Neutral**: A more general system suggestion for a user to action (e.g., “Hey Alex, do you want to change the audience of this post”); 2) **Passive**: A more gentle system suggestion for a user to take action (e.g.,“ Hey Alex, perhaps you might think about changing the audience of your post”); 3) **Assertative**: A more commanding/forceful system suggestion for a user

Privacy Related	Non-Privacy Related
Positive Affective State: Employers use social media to screen candidates during the hiring process. More specifically, **employers state that viewing someone's profile gives them a glimpse into their personality and online behavior beyond their resumes. As a result, it helps them recruit candidates who are a good fit for their company.** This means that if you search for a job in the future, you could be eligible for a higher starting pay because of what employers learn about you from your social media..	**Positive Affective State:** Given the growth and reach of social media sites, many businesses continue to rely on social media to market and sell their products or services. As a result, people use their social media sites to buy products online. This means that it is easier for you to discover small businesses and products that you might enjoy on social media.
Negative Affective State: Employers use social media to screen candidates during the hiring process. More specifically,**employers report rejecting job applicants based on their social media posts that reflect poorly on the applicant. As part of the interview process, some employers go to the extent of asking applicants to share their login details to their social media accounts in order to view all their posts.** This means that if you search for a job in the future, you might be disqualified from a job before you interview based on what employers learn about you from your social media.	**Negative Affective State:** Given the growth and reach of social media sites, many businesses continue to rely on social media to market and sell their products or services. As a result, it has become very expensive for businesses to market their products online. This means that products may be more expensive for you as businesses need to pay more for social media ads.

Figure 5.2: The positive and negative *privacy* or *non-privacy* related passages used to induce user affect.

to take action (e.g.,"Hey Alex, you should absolutely change the audience of this post").

In summary to examine for the impact of privacy-related affect and privacy suggestion tone styles on user privacy protection decisions and experiences with the platform (i.e., perceived decision help and trust), we operationalized the experimental conditions into three main core parts for our research framework: 1) the emotional valence of the passage (i.e., *positive* Vs. *negative*); 2) the passage relation (i.e., *privacy* Vs. *non-privacy*); and 3) the privacy suggestion tone styles (i.e., *neutral* Vs. *passive* Vs. *assertive*). Next, we further describe the constructs and hypothesize the relationships that constitute this research framework.

5.3.2 Perceived Decision Help as a Benefit of Privacy Suggestions

As detailed in Chapter 3 and 4, we find that users would appreciate privacy suggestions as an adaptation presentation method where the system/platform proactively guides, educates and helps encourage them to make the “right”’ privacy decisions. More specifically, users discern that privacy suggestions would enable them to proactively take ownership over their privacy [169]. Hence, a platform that provides privacy suggestions that embody the “appropriate” tone style to modulate how users’ react would likely be perceived to be very helpful. However, prior work reveals that this perception is likely to differ based on users’ pre-existing attitudes (in our study these are manipulated as privacy-related affective states) [111]. For example, Bahirat et al. [24] suggest a “privacy” focused decision-making context matters as it can influence users’ privacy decision outcomes, especially for users who may be vocal about privacy (i.e., those with high concerns). Such contextualization is likely to focus users and highlight the importance of privacy suggestions, especially if they are vocal or highly concerned about their privacy [24]. Kehr et al. [111] further asserts that users’ benefit considerations (e.g., of intelligent components such as privacy suggestions) might differ based on their pre-existing affect towards situation-specific privacy risk. Compared to users in a negative affect state, users in a positive affect state tend to perceive lower situation-specific privacy risks [111]. In essence, users in a negative affect state might perceive a system that provides privacy suggestions to be more helpful than those in a positive affect state, since users in a positive affect state tend to have more positive beliefs that can lead to an underestimation of the privacy risks inherent with system use [142]. Therefore,we hypothesize:

H1a: *In contrast to participants primed using a non-privacy-related passage, participants primed using a privacy-related passage will perceive the platform to be more helpful.*

H1b: *However, compared to participants primed using a positive privacy-related passage, the effect observed in (H1a) will be stronger among participants primed using a negative privacy-related passage*

Furthermore, in regards to the impact of privacy suggestion tone styles on users’ perceived decision helpfulness of a system, prior work by Muench et al.[167] reveals that individual’s have clear tone style preferences, especially if the message is designed to help them achieve their personal goals. More specifically, the authors found that individuals clearly prefer a more directive (i.e., assertive) message tone over a more suggestive (i.e., passive or neutral tone [167]. In other words, individuals tended to prefer more directive (i.e., assertive) than passive or neutral toned messages. Thus, subtle manipulations of the tone style of a message could affect user preference for the message, and as such the perceived helpfulness of the platform [167]. Thus, we hypothesize:

> **H1c**: *Participants' provided with privacy suggestions that embody an assertive tone will perceive the the platform to be more helpful in helping them make privacy protection decisions, than participants provided with suggestions that are either neutral or passive.*

Research also further reveals that user message tone style preferences might be susceptible to user affective states [131, 55, 111]. Li et al. [142] assert that while users in a positive affect state tend to care more about the protection of their privacy, they also tend to underestimate the inherent privacy risks they face. Bless et al. [34] also found that, depending on the content of the message, happy individuals (i.e., individuals in a positive affect state) can be persuaded by both strong(e.g., a more direct assertive) and weak (e.g., a more suggestive neutral or passive) message, while sad individuals (i.e., individuals in a negative affect state) are more persuaded by the strong than weak message. Li et al. [142] suggests that this is likely because individuals in a negative affect state tend to focus more on the privacy risks involved in a situation, reach a quick decision regarding the potential downsides of the risk, and then act accordingly. Thus, we hypothesize that:

> **H1d**: *Compared to participants primed using a positive passage, the effect in (H1c) will be much stronger among participants' primed using a negative passage. This effect will only be true for participants primed with a positive or negative privacy-related passage.*

With regard to privacy concerns, users are likely to respond to more assertive tones (i.e., pushy or more forceful messages) in domains that they view as important while more suggestive (i.e., neutral or passive tones) are likely to work best when they lack initial conviction [131]. In other words, users' are likely to prefer an assertive tone style and perceive a system to be helpful, if they are very concerned about their privacy. As such, we hypothesize that:

> **H1e**: *The effect observed in (H1d) will not be true for participants primed using a positive or negative non-privacy-related passage.*

5.3.3 Trust in the platform as an antecedent for Privacy Decision Making

As highlighted in section 5.2.3, user trust is a principal antecedent to their privacy decisions (e.g., whether to disclosure or not to disclose information) [73, 103, 139]. User trust in the platform—conceptualized in this study as an individual's confidence that the platform will not misuse his or her data [112]—shapes users' judgements and privacy decision outcomes [139]. According to Malhotra et al. [157], trust helps to bridge the tension between the platform's providers' privacy practices and the users' privacy concerns about disclosure of information. Consequently, trust increases the confidence users have in the platform, which lowers their perceived risk of disclosing personal information, and ultimately increases the likelihood of users engaging in information exchanges [100, 209]. From my prior work in Chapter 3, we learn that users' are likely to perceive a platform that provides privacy suggestions as one that deeply cares and values their privacy. As such, the provision of privacy suggestions is likely to influence the level of trust in the platform. However, keeping in line with past findings [39, 94], the level of user trust in the platform could vary based on the tone embodied by the suggestion. As such, we hypothesize that:

H2: *Participants' provided with privacy suggestions that embody an assertive tone, will have higher trust in the platform than participants provided with suggestions that are either a neutral or passive.*

H3: *Given the provided privacy suggestions, perceived decision help will be positively associated with trust in the platform.*

H4: *Users' general informational privacy concerns will be negatively associated with perceived user trust in the platform.*

5.3.4 Dependent Variable: Privacy Protection Decision Outcome

Social media users often have to make privacy decisions pertaining to the management of their online identity and access to their personal information (i.e., reputation management) [180, 240]. Reputation management involves careful curation of the type of content (or posts) that can be shared, viewed or accessed by others in order to project oneself in a way that suits specific audiences [19, 61]. As part of this privacy decision making process, users can engage with the available privacy controls or features to delete shared content they do not want others to see, select an audience that can access the content, self-censor (i.e., decide not to share any content at all), among many other actions [240].

As such, in this study, the actual privacy-related decision pertains to reputation management as a common privacy management practice on social media [240]. Herein, participants were implored to review "their" fictitious social media profile and on their own accord take privacy protective actions (i.e., reputation management decisions) based on the review of ten posts that varied in privacy sensitivity [144, 29]. The actual permitted privacy actions that participants could undertake included either deletion or edit or changing audience of the posts that participants' deemed to be highly sensitive or poorly reflective of them. We treated the total number of posts whose privacy settings were changed as the outcome variable of interest. We termed this "total" as the *privacy protection decision.*

Thus, keeping in line with past literature on the influence of privacy concerns, user trust, and perceived decision help on users' privacy decision outcomes [207, 157, 209, 169], we hypothesize:

H5: *Participants with high general information privacy concerns will make a higher number of privacy protection decisions, depending on their positive or negative privacy-related affective state.*

H6: *Perceived trust in the platform will be negatively associated with the number of privacy protection decision outcomes.*

H7: *Perceived decision help will be positively associated with the number of privacy protection decision outcomes.*

Finally, several privacy scholars suggest that in examining users' privacy decision-making processes, researchers should take into account the effects of a dual-route or ("hybrid") decision-making approach—privacy calculus (i.e., cost-benefit analysis of disclosure) integrated with heuristic considerations [73, 97, 11, 142, 110]. More specifically, the scholars highlight the risks associated with only relying on either a privacy calculus or a heuristic-based approach. For example, Anaraky et al. [73] reveal that in making privacy decisions, users can "employ a hybrid process that integrates heuristics, such as taking into account the perceived trust in the [platform] along with making calculated assessment of the benefits and costs of disclosure". The authors strongly assert that relying on either a privacy calculus or a heuristic-based approach can result into obscured or diluted effects as it might not sufficiently capture the true effects or ways in which all users make privacy-related decisions [73]. Kehr et al. [110] also reveal that users can employ different thinking styles; employing either a rational way (i.e., a cognitive process where they thoroughly assess the anticipated benefits and risks associated with the decision) or intuitive way (i.e., relying on their hunches rather than employing a cost-benefit analysis) in making disclosure decisions. Al-Maidani & Al-Jabri [11] find that social media users' affective states can bias their privacy calculus process. To that end, in this study we account for the possible effects of a dual-route decision-making approach by not only examining how the three different privacy suggestion tone styles directly affect users' privacy decision as moderated by privacy-related affect, but also inspect the moderating effect of decision help and trust in the platform in this process. By taking into account the different ways users might make decisions and studying the combination of these variables, we are able to comprehensively explore the boundary conditions of the effects of privacy suggestion tones and privacy-related affect. Thus, we answer the research question:

> **RQ4:** Are the effects of the experimental manipulations (i.e., privacy-related affect and privacy suggestion tone styles) on privacy protection decision, mediated by decision help and trust in the platform?

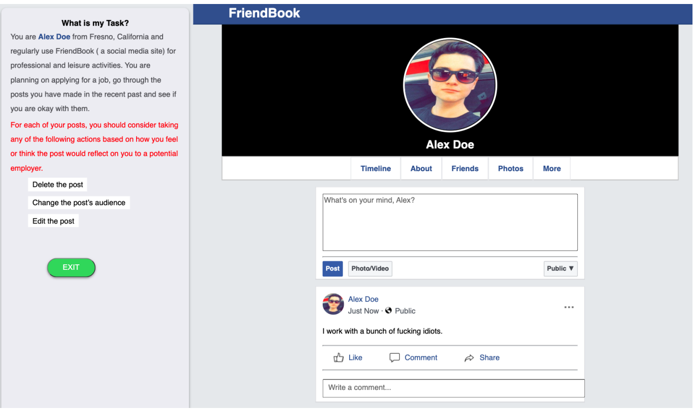
5.4 Methodology

The goal of this research work was to better understand how privacy-related affect and the different privacy suggestion tone styles impact users' experience (i.e, perceived decision help and trust in the platform) and privacy decision outcomes. As such, we conducted an online user experiment exploring the impact of three privacy suggestion tone styles (*neutral*, *passive*, *assertive*) and privacy-related affect on users' users' experience and privacy decision outcomes. The Clemson University Institutional Review Board (IRB) approved our study.

5.4.1 Study Design: The SNS User Interface Mockup

One of the objectives of this study was to overcome the shortcomings of studies with hypothetical scenarios and obtain increased ecological validity. Therefore, participants interacted with a carefully controlled working prototype SNS platform [2] ("FriendBook", see Figure 5.3) that could purportedly provide its users with "privacy suggestions"(e.g., Figure 5.4) to help inform their privacy decisions. To avoid a cluttered user interface and profile, FriendBook was populated with only ten posts based on a Tweet corpus collected by Cachola et al. [42], with each post containing settings (or features) for the three plausible post privacy actions that users could undertake (i.e., edit a post, delete a posts, change the post audience). Furthermore, these posts varied in content privacy sensitivity (i.e., low, medium, high) based on a taxonomy by Li et al. [144]. In particular, based on Li et al's taxonomy, the posts were comprised of four high (i.e., two posts each containing vulgar text, one photo post of a medical condition, and one showing a photo of a bong filled with cannabis), three medium (i.e., two posts containing negative text, and one with a photo of a disorganized home), three low (i.e.., two posts each containing positive text celebrating an event such as a birthday and get-to-together party, and one post photo of a vacation) privacy sensitive content (see Figure 4, Appendix B.4; for the full gallery of posts used in the study) . Hence, each user saw the exact same posts, friends, etc., thereby guaranteeing that all users had the same opportunities to engage with the same profile.

[2] Developed based on the User Interface (UI) of the Facebook web application to increase the realism and ecological validity of the experiment.



What is my Task?

You are Alex Doe from Fresno, California and regularly use FriendBook (a social media site) for professional and leisure activities. You are planning on applying for a job, go through the posts you have made in the recent past and see if you are okay with them.

For each of your posts, you should consider taking any of the following actions based on how you feel or think the post would reflect on you to a potential employer.

Delete the post

Change the post's audience

Edit the post

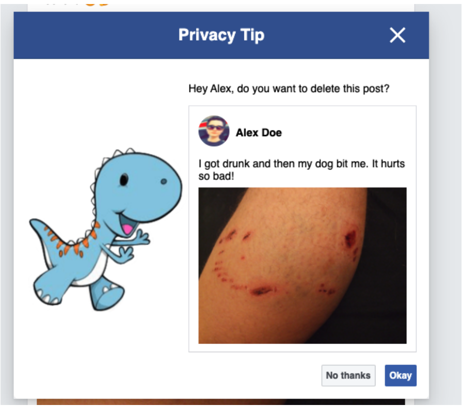
What's on your mind, Alex?

Alex Doe
Just Now · Public

I work with a bunch of fucking idiots.

Write a comment...

Figure 5.3: The semi-functional social media platform ("FriendBook") used to provide users with "privacy suggestions" to help inform their privacy decisions. Free public images accessed from the internet (under a (CC0) commons creative license) and fictitious names were used in the creation of "Alex Doe's" profile.



Privacy Tip

Hey Alex, do you want to delete this post?

Alex Doe

I got drunk and then my dog bit me. It hurts so bad!

Figure 5.4: A sample post privacy suggestion (termed "privacy tip" within the study) encouraging participants to delete one of the highly privacy sensitive posts on the FriendBook platform.

Using FriendBook allowed us to manipulate how we presented the "privacy suggestions" and aptly examine the impact of tone styles on users' privacy decision making process. A privacy suggestion was presented for each of the three most "high" privacy sensitive posts and in alignment with each of three recommended privacy actions (see Section 5.5 for the examined three privacy actions, and Section 5.4.3 for a description of the 12 experimental conditions). The other one "high" privacy sensitive post did not get a suggestion. We chose to present only a total of three privacy suggestions under each condition so as to not overwhelm participants. For each privacy suggestion, participants were asked to consider taking a particular privacy protection action, and could respond by pressing the "Reject" or "Accept" options. The privacy suggestions were designed to make a single point in a straightforward manner, and therefore be clear and informative [50].

We ran a small pilot study with 10 participants to ensure there were no usability issues and get a good timing pattern for the appearance of the privacy suggestions. Based on their feedback, the three privacy suggestions were updated to appear at the time intervals of 40s, 60s, 80s, for a maximum appearance time of 10 seconds respectively.

All user interactions with the posts were recorded and used to access overall engagement patterns and privacy protection decision outcomes (see Section 5.8.3).

5.4.2 Study Setup & Procedure

After reading a brief description of the study's purpose and providing consent, participants completed a pre-survey (see Appendix B.1,Table 2). This pre-survey asked participants to indicate their current (Facebook) usage (based on scale adopted from Ernala et al. [62]), awareness and past usage for each of the three post privacy features (i.e., "edit post","edit audience", "delete post") used to enact the examined privacy actions in this study (see Section 5.5). This was done by showing the participant an image of the privacy feature under examination and asking them 1) "Are you familiar with the [Name of Feature] Facebook post feature?"(response options: Yes, No) and 2) "How often do you use this feature?" (response options: Never Used, Used Once, Occasionally Use, Frequently Use). The response to these questions enabled us get a clear understanding of how often our participants used social, and their level of familiarity and usage of the post features used to make to the privacy decisions within this study (see Section 5.8.1.1 on the resultant descriptive statistics of the responses to these questions). Thereafter, participants were randomly assigned into one of the 12 experimental conditions to interact with FriendBook where they could partake in the privacy decision-making process (see Section 5.4.3 for details on the experimental setup).

Prior work reveals that it has become increasingly common for potential employers to "canvass social media sites for information on potential employees and candidates, and act on the basis of the information found therein" to make hiring decisions [1, p.95]. For job seekers, objectionable social media posts (such as those that include inappropriate photographs or information, evidence of alcohol or drug use, and information revealing that the applicant might have lied on the job application) may decrease their chances of getting a job based on a potential employer's judgements about their character or reputation [1, 188]. Thus, one of the common privacy behaviors that social media users—especially those seeking employment opportunities—exhibit is managing their reputation [240]. Herein, users seek to manage their online identity and access to their personal information that they might not want to be seen, shared or made available to others [240]. As such, a job search scenario was used as a motivating context in which participants could explore and manipulate the FriendBook profile used in the study. Participants were implored to review and make privacy protection decisions based on the careful curation of the posts if they were to be viewed by a potential employer. More specifically, participants were invited to imagine that:

> *"You are Alex Doe from Fresno, California and regularly use FriendBook (a social media site) for professional and leisure activities. You are planning on applying for a job, go through the posts you have made in the recent past and see if you are okay with them."*

Together, the scenario and the post privacy-feature related questions helped participants learn, navigate, engage, explore and review "their" profile on FriendBook. For easy recollection of the use context, the scenario and list of possible post privacy actions was also presented as a persistent sidebar throughout the user interaction process with FriendBook (see Figure 5.3). These were carefully pilot-tested with the study target sample population (N = 10) to make sure that participants were properly motivated to manage their profile without explicitly demanding that they would engage in specific privacy management practices. Responses in our pilot-test debriefing interviews convinced us that participants would interact with their profile and make privacy choices that *they themselves* thought to be the most appropriate ones to undertake.

Participants were subsequently asked to explore and interact with their profile, with the goal of ensuring that they were okay with what is on it, given the imagined upcoming job interview. In this phase, participants reviewed the various recently shared posts on their timeline and—where appropriate—made changes using the available post privacy features on their own accord or with the aid of privacy suggestions [3]. Depending on the experimental condition, the tone style of the privacy suggestions was varied.

Upon completing the review of the posts on FriendBook, participants were directed to complete the post-stimulus survey. Based on their interactions and privacy decisions, participants were asked to evaluate the overall trust they had in the FriendBook platform (based on a scale

[3] Participants who spent too little time (< 1 minute) interacting with FriendBook were removed from the analysis so as to ensure that all the participants in the study had seen or interacted with at-least one of the privacy suggestions. The remaining participants spent an average of 2 minutes on FriendBook

adopted from Krasnova et al. [130]), general informational privacy concerns (based on a scale adopted from Malhotra et al. [157], and the usefulness of the FriendBook platform (based on a scale adopted from Knijnenburg et al. [124]). For each of the ten posts, participants were asked about what privacy action they took and a corresponding reason. Each participant was compensated with $2 for participating in the study [82].

5.4.3 Experimental Conditions

To address our research questions, we employed a 2 X 2 X 3 between-subjects experimental study, that relied on users' reading a story or "passage" before interacting with FriendBook so as to ensure all participants had a uniform induction of affect. For each passage, we combined these factors to ensure that the elicited affect and ultimate judgements were based on the same decision context with differences in relation to privacy and framing only [63, 35]. More specifically, the passage was either directly related to privacy (*privacy-related*) and positively or negatively framed— based on the emphasises of the advantages (i.e., *positive privacy framing*) or disadvantages (i.e., *negative privacy framing*) related to the access and use of user information on social media; or directly unrelated to privacy(*non-privacy related*)— where we intentional emphasized the advantage (i.e., *positive non-privacy framing*) or disadvantages (i.e., *negative non-privacy framing*) related to the general use of social media without specific mention of the benefit or loss related to social media use of users' personal information (see Figure 5.2 for further details).

Immediately after reading the passage, participants were asked to indicate how it made them feel. We used a Positive and Negative Affect Schedule (PANAS-SF) scale with 20 items as a manipulation check to ensure the right affective states were elicited [229]. Thereafter, participants were shown the study scenario (described in Section 5.4.2) and subsequently directed to the FriendBook platform where the presented privacy suggestions varied in tone (see Figure 5.1). As such, we developed a total of 12 experimental conditions: passage relation (*privacy* versus *non-privacy*), the emotional valence of the passage (*negative* versus *positive*), and on-platform privacy suggestions in tone styles (*neutral* versus *passive* versus *assertive*) (see Table 5.1 for the exact sample privacy suggestion tone styles)

Tone	Description	Example
Neutral Tone	The message was framed using a general neutral tone	“Hey Alex, do you want to delete this post?”
Passive Tone	The message was framed using a passive tone	“Hey Alex, perhaps you might think about deleting this post”
Assertive Tone	The message was framed using an assertive tone	“Hey Alex, you should absolutely delete this post”

Table 5.1: The three different suggestion tone styles that were used to offer on-platform decision support.

5.5 Measurement

We recorded all user interactions with the post privacy features to measure user engagement with the privacy suggestions and capture their ultimate privacy protection decision outcomes:

Explicit accept: The participant explicitly accepted the privacy suggestion, either by approving the suggestion (by clicking “Ok”).

Implicit ignore: The participant ignored the privacy suggestion or the suggestion disappeared before they were able to interact with it, thereby implicitly ignoring it. .

Explicit reject: The participant explicitly rejected the privacy suggestion (by clicking “Rather Not”).

Privacy Protection Decision Outcome: Based on their perception of the post, the participant on their own accord or implored by the privacy suggestion either: 1) **deleted the post**: this action provided the privacy benefit of completely getting rid of the entire post; or 2) **selected an appropriate audience for the post**: this action provided the privacy benefit of categorically controlling *who* could view the post without necessarily having to delete or edit the post; or 3) **edited the post**: this action provided the privacy benefit of altering the post content without necessarily getting rid of it entirely or changing its’ audience. Thus, we define the “*privacy protection decision*” as the total number of posts where any one of these privacy actions was undertaken.

5.6 Participant Recruitment

Participants were recruited between January and March 2022 via Prolific [4], a participant recruitment platform where people complete short tasks and receive automatic payments. A total of 993 adult participants who were users of social media (e.g., Facebook) were recruited. We restricted participants to people within the United States with a high "worker reputation" (i.e., those with a HIT approval rate greater than 95% with at least 50 approved past tasks) to ensure satisfactory response quality. We also included several attention check questions and quality checks to remove participants who spent little time (less than 1 minute) within the study environment or who did not carefully read/respond to the pre- and post-survey questions [132]. We discarded 243 pariticipants who did not meet our participant requirements and data quality checks, the valid data used in the analysis was from 750 participants [5] : (169 Men, 573 Women), with aged between 18 and 60 (average age 34). We summarize the distribution of participants across the 12 experimental conditions in Table 5.2.

Privacy Suggestion Tone Styles	**Privacy-Related (N = 366)**		**Non-Privacy Related (N = 384)**		**Total (N = 750)**
	Negative Framing	***Positive Framing***	***Negative Framing***	***Positive Framing***	
Neutral	61	59	71	63	254
Passive	66	57	54	68	245
Assertive	60	63	60	68	251

Table 5.2: The Distribution of Participants across the 12 Different Experimental Conditions

5.7 Data Analysis Approach

Our data analysis approach was three-fold: 1) we first assessed the reliability and validity of the pre-validated post-study survey items assessing the constructs of perceived decision help, perceived trust in the platform, general information privacy concerns using confirmatory factor analysis (see Section 5.7.1) [125]; 2) we examined the research model (Figure 5.1) and tested the hypotheses using structural equation modeling (SEM) (see Section 5.7.2) [125]; and 3) in a behavioral analysis (see Section 5.8.3), we analyze for significant differences in privacy protection decision outcomes across the varying privacy suggestion tone style conditions based on the *passage relation*

[4] https://www.prolific.co/

[5] A power analysis, with α=.05, power=.95, df=11, and twelve groups revealed that the suggested sample size (N=413) of a factorial ANOVA test was sufficient for detecting a medium effect (f=0.25)

condition and congruent *emotional valence of the passage*. Subsequently, we examine the particular privacy actions undertaken and privacy suggestion engagement patterns.

5.7.1 Confirmatory Factor Analysis

The validity of our constructs (i.e.,perceived decision help, perceived trust in the platform, general information privacy concerns) was assessed using confirmatory factor analysis (CFA) in R Studio. CFA helps establish convergent and discriminant validity to ensure that the survey items are a valid measurement of the constructs [125]. More specifically, convergent validity helps "determine whether the items of a scale measure a single construct (i.e., that the scale is not a combination of multiple constructs, or simply a collection of items with no common ground), while discriminant validity determines whether two scales indeed measure two separate constructs (i.e., that two scales are not so similar that they actually measure the same construct)" [125, p.25]. In CFA, survey items that belong to the same scale are represented by a latent factor. The analysis determines to what extent the item serves as an adequate indicator of the factor (loading). We iteratively removed items with high cross-loadings and items with low (< .70) loadings on their own factor; these items have no loading in Table 3, Appendix B.3. Overall, the results in (Table 3, Appendix B.3) show adequate convergent (AVE > 0.50) and discriminant validity ($\sqrt{}$(AVE) > largest correlation) for each factor, and a substantial loading for each item (i.e., each item loading exceeded 0.70), with a good[6] overall model fit [117]: χ^2 (183) = 869.977, p < 0.001; RMSEA = 0.071, CFI = 0.985, TLI = 0.982, albeit with a high RMSEA.

5.7.2 Structural Equation Modelling

In conducting an SEM, we examined the research model (Figure 5.1) and determined the statistical significance of the hypothesized relationships between the constructs. SEM is an "integrative statistical procedure that tests the measurement model and all hypotheses (known as the structural model) at the same time. Therefore, our research model with the experimental conditions, constructs (i.e., perceived decision help, perceived trust in the platform, general information

[6] A good model has χ^2 that is not statistically different from a saturated model (p >.05), but this statistic is considered too sensitive. Researchers have considered other fit indices [31]. Hu and Bentler [93] propose cutoff values for other fit indices to be: CFI > .96, TLI > .95, and RMSEA < .05, with the upper bound of its 90% CI below 0.10.

privacy concerns), and privacy protection decision outcomes, was examined for both hypothesized and potential non-hypothesized effects using a "saturated" path model of these core factors [125]. We iteratively pruned the non-significant effects of the model and examined the sign and significance of the path coefficients; in our resulting model (see Figure 5.6). The solid incoming arrows (→) between constructs represent significant relationships while the broken line arrows (⇢) represent tested relationships that were found to be non-significant. Each regression contains a regression coefficient (indicated by the number on the arrow), the standard error of the regression effect (in parenthesis) and the significance level denoted by asterisks (*) or "ns" for non-significant effects). The latent constructs were scaled to have a standard deviation (SD) of one, so that one SD difference in a construct (e.g. perceived decision help) causes a β SD difference in another construct (e.g. perceived trust in the platform).

5.7.3 Behaviorial Analyses

Next, we conducted a behaviorial analyses to better understand how the various privacy suggestion tone styles affected participants' privacy protection decisions depending on the passage_relation and emotional valence of the passage conditions. In particular, we ran a fit analysis of variance (*ANOVA*) to evaluate and examine the overall effects of the three experimental conditions on participants' privacy protection decisions (based on a total count of posts for which privacy actions where undertaken). Additionally, we also ran a *generalized linear mixed effects regression model (glmer)* to examine for similar overall effects based on particular post level decisions. Herein, we included a logit link function to account for the binary post decision outcome variable (logistic regression) and created a random intercept to account for the within subjects (multiple post decisions per participant) design of the study. We first created a baseline model, which only comprised of a random intercept. Next, we added the passage_relation, emotional valence of the passage, suggestion tone style, and respective interaction effects as additional variables to the baseline model. We tested whether there was a significant improvement upon adding the new variables using a $\chi 2$-based model comparison.

A series of similar *generalized linear mixed effects regression models (glmer)* were also conducted to examine for the impact of post and privacy suggestion attributes (e.g., post content privacy sensitivity and whether a suggestion was provided or not) on participants' privacy protection decisions, particularly among participants who were primed using a privacy-related passage.

Furthermore, to better understand the influence of privacy suggestions, we also assessed the particular privacy actions undertaken and overall user engagement patterns. We report on the post-hoc findings in Section 5.8.3 based on the $\chi2$-based model outcomes.

5.8 Results

Below, we describe our study's findings. We first provide the descriptive statistics (see Section 5.8.1) regarding the participants daily use of social media, level of privacy feature awareness and rate of use (Section 5.8.1.1), and the effectiveness of our experimental passage(s) affect manipulations (Section 5.8.1.2). Then, we present our hypotheses test results (Section 5.8.2), followed by a post hoc analysis to further unpack additional nuances from the participant data (Section 5.8.3).

5.8.1 Descriptive Statistics

In the following subsections, we provide the descriptive statistics for social media use, privacy feature awareness, and usage, and manipulation check outcomes.

5.8.1.1 Social Media Use, Privacy Feature Awareness and Usage

For each participant, we inquired about their social media activity, timeline privacy feature awareness and use, especially for features that were meant to support the three main privacy actions that they could consider undertaking to safeguard their privacy while on "FriendBook" (i.e., edit a post, delete a posts, change the post audience). Out of the 750 participants, 635 (84.6%) indicated being active social media (i.e., Facebook) users that used the SNS for at-least 10 minutes to more than three hours per day. A majority of participants reported that they were unfamiliar with some of the examined post privacy features: Edit Post (61, 8.13%), Delete Post (391, 52.1%), Edit Post Audience (425, 56.7%). Subsequently, they also reported never using them: Edit Post: (143, 19.07%), Delete Post: (457, 60.93%), Edit Post Audience: (491, 65.47%). This finding confirms prior work that reveals that many social media users' are often unaware of the privacy feature controls available to them via their respective settings [92]. This finding also suggests that indeed the application of "privacy suggestions" as plausible privacy adaptation presentation method could benefit social media users and encourage them to take active control over their privacy [169].

5.8.1.2 Manipulation Check

Immediately after reading the respective passage, participants were asked to rate their current affect state on the Positive and Negative Affect Schedule (PANAS-SF) scale as a means to examine the exact valence of the elicited affect (see Appendix B.2, Figure 3) [229]. As internal consistency of the two scales was sufficient in the passages that varied based on their relation to privacy: Privacy-Related passages (Positive affect: 10 items, $\alpha = 0.93$; Negative affect: 10 items, $\alpha = 0.85$) and Non-Privacy-Related passages (Positive affect: 10 items, $\alpha = 0.88$; Negative affect: 10 items, $\alpha = 0.86$), we conducted an independent sample t-tests to assess the effectiveness of the manipulations across the privacy and non-privacy related affect experimental conditions. Results indicated that affect elicitation was successful, with participants feeling more negative than positive if they read a negative framed passage, and vice versa. More specifically, for negative affect elicitation (i.e., where participants read a negatively framed passage), participants were more likely to report significantly feeling more negative than positive in the non-privacy-related condition ($t(327.03) = 4.35, p < .001$) and privacy-related condition ($t(322.04) = 5.92, p < .001$). On the other hand, for positive affect elicitation (i.e., where participants read a positively framed passage), participants were more likely to report significantly feeling more positive than negative in the non-privacy-related condition ($t(235.76) = 11.07, p < .001$) and no significant differences were observed for the privacy-related condition ($t(368.84) = 0.36, p > .05$). Figure 5.5 provides an overview of these manipulation check findings. These results reveal that privacy-related affect from an emotional perspective generally tends to be different from non-privacy-related affect . More specifically, privacy-related affect tends to elicit more negative than positive affect in comparison to non-privacy-related affect.

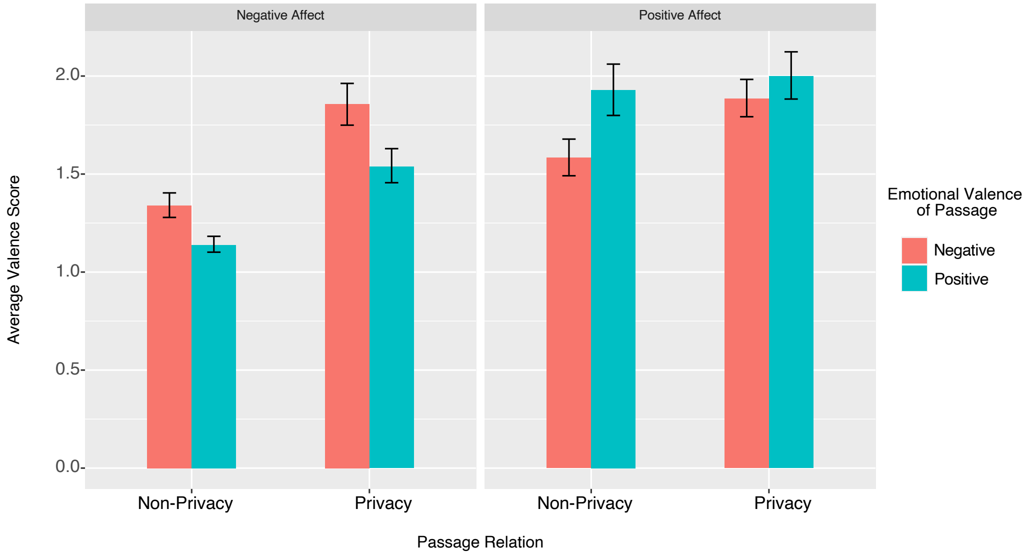


Figure 5.5: Manipulation Check: The average valence scores for the elicited negative and positive affective states across the privacy and non-privacy related affect experimental conditions. Negative and Positive affects are assessed based on a summation of the particular items in the PANAS-SF scale (see Figure 3, Appendix B.2) [229].

5.8.2 Structural Model

Using structural equation modelling (SEM), we examined the research model (see Figure 5.1) to examine for the statistical significance of the hypothesized relationships between the constructs. Having removed the non-significant effects in the research model, we analyze the significant effects. The resulting model (Figure 5.6) had a good model fit ($\chi2$ (211) = 359.181, p < 0.000; RMSEA = 0.031 (which is well below the suggested maximum of .05, CFI = 0.997, TLI = 0.999 [93]). Below, we discuss (from left to right) the individual hypothesized effects.

Effect of passage relation (H1a - H1b): The results indicate that there was no direct main effect of passage-relation on the perceived helpfulness of the platform (**H1a** not supported). Furthermore, there was no significant interaction effects between passage relation and affect on the perceived decision help (**H1b** not supported).

Effects of privacy suggestion tone styles (H1c - H1e): The results indicate that contrary to our hypotheses, there was no significant direct effect of privacy suggestion tone styles on participants' perception of the helpfulness of the platform (**H1c** not supported) or even a moderating effect

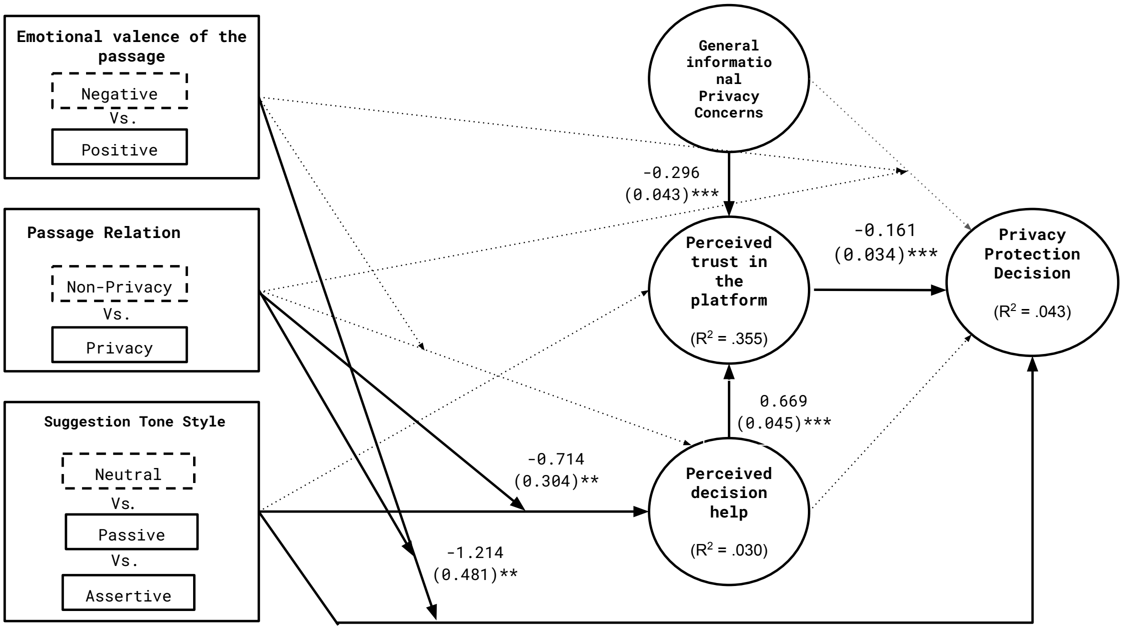


Figure 5.6: The Structural Equation Model (SEM) of the research framework used to examine the influence of pre-existing privacy-related affect and privacy suggestion tone styles on decision help, trust in the platform, and privacy protection decision outcomes. The model shows the direct effects of the hypothesized and non-hypothesized determinants (Significance levels: ***p <.001, **p <.01, 'ns' p > .05, R^2 is the proportion of variance explained by the model. Numbers on the arrows represent the β coefficients and standard errors (in parenthesis) of the effect). Factors are scaled to have an SD of 1

of user affect (**H1d** not supported). Instead, The model shows there was a significant two way interaction effect between passage relation and privacy suggestion tone styles on users' perceived decision helpfulness of the platform (β = -0.714, p < .001; **H1e** supported). Participants primed using a privacy-related story and provided with privacy suggestions that embodied a neutral tone, perceived the platform to be more helpful than those primed using a non-privacy passage and provided with a suggestion that embodies a neutral tone (see Figure 5.7).

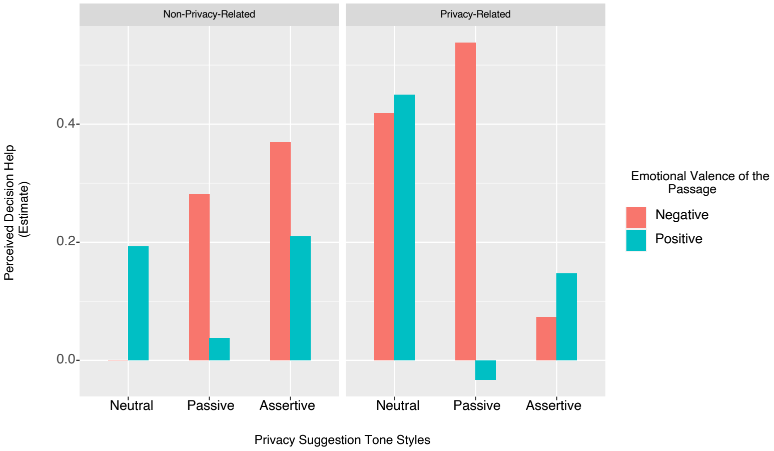


Figure 5.7: The effect of the passage relation (non-privacy Vs. privacy), emotional valence of passage (negative Vs. positive) on perceived decision helpfulness of the platform.

Effect on trust in the platform (H2-H4): There was no significant direct main effect of privacy suggestion style on the perceived trust in the platform (**H2** not supported). Perceived helpfulness of the platform had a positive direct main effect on trust in the platform (β = 0.669, p < .001; **H3** supported). Additionally, general informational privacy concerns had a negative direct main effect on trust in the platform (β = -0.296, p < .001; **H4** supported). This suggests that users with high privacy concerns are more likely to have less trust in the platform [112].

Effect on privacy protection decision outcomes (H5-H7): There was no significant three way interaction effect of passage relation, the emotional valence of the passage, and general information privacy concerns on privacy protection decision outcomes (**H5** not supported) neither was there a significant direct main effect of decision help (**H7** not supported). Instead, there was a negative direct main effect of trust in the platform on privacy protection decision outcomes (β = -0.161, p < .001; **H6** supported). Additionally, we found a non-hypothesized three way interaction effect of passage relation, emotional valence of the passage, and privacy suggestion tone style on privacy protection decision outcomes (β = -1.214, p < .001).

Overall, our resulting model findings (see Figure 5.6) show marginal effects of passage relation, emotional valence of the passage, and privacy suggestion tone styles on the subjective constructs (i.e., perceived decision help and trust in the platform). However, an examination of the user actions or actual behavior exhibited on the platform show inimitable effects of the experimental conditions on participants' ultimate privacy protection decision outcomes. Taken together, these findings suggest that the effects of the experimental manipulations (i.e., privacy-related affect and privacy suggestion tone styles) on participants' privacy protection decisions are partially mediated by perceived decision help and trust in the platform (**RQ4**). Otherwise, the experimental manipulations also have direct ramifications on users' actual privacy decision outcomes.

Therefore, in a post-hoc analysis, we investigate these behaviors in much greater detail to uncover the true impact of the experimental conditions on participant's privacy decision outcomes. Herein, we also examine the particular privacy actions undertaken, impact of privacy sensitivity of the posts on participants' decisions, and impact of privacy suggestions the observed engagement patterns. The resultant findings from this analysis makes up the bulk of our discussion.

5.8.3 Privacy Protection Decision Outcomes

In this subsection, we present the differences in privacy protection decision outcomes across the varying privacy suggestion tone style conditions based on the passage_relation condition and congruent emotional valence of the passage.

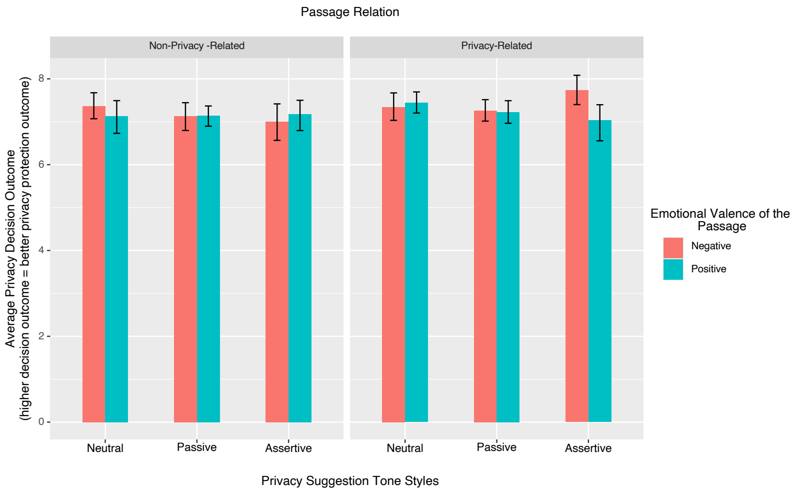


Figure 5.8: The effect of the passage relation (non-privacy Vs. privacy), emotional valence of passage (negative Vs. positive) on participants' privacy protection decision outcomes per the privacy suggestion tone style used.

Figure 5.8 shows the effects on privacy protection decision outcomes (i.e., the total number of posts where a privacy action was undertaken) across the three varying privacy suggestion tone style conditions differed based on the passage_relation and emotional valence of the passage. In particular, a linear ANOVA model based on privacy protection decision outcomes (i.e., the total count of the posts for which privacy actions where undertaken), revealed that the passage relation, emotional valence of the passage, and privacy suggestion tone style independently did not reach significance ($p > .05$). Instead, a three-way interaction between these three experimental conditions was significant ($F_{(1,2)}$ = 3.417, p = .0333), showing that their effects were largely interdependent (see Table 5.3). A *glmer* model based on particular post-level decisions as a binary outcome, showed relatively similar results ($\chi^2(2)$ = 5.440, p = .06). Therefore, in subsection 5.8.3.1, we further unpack the interaction effects of experimental conditions on participants' privacy protection decision outcomes. In subsection 5.8.3.2 and 5.8.3.3, we analyse the impact of post content privacy sensitiveness, and presence of privacy suggestions, in the privacy-related condition only.

ANOVA Table	*df*	$F - Value$	*p*-value
privacy_protection_decision_outcomes			
+Passage Relation	1	0.0089	.925
+Emotional Valence of the Passage	1	1.0805	.298
+Suggestion Tone Style	2	1.285	.277
Two way Interactions			
+Passage Relation:Affect	1	1.0067	.316
+Passage Relation:Suggestion Tone Style	2	2.782	.063
+Emotional Valence:Suggestion Tone Style	2	0.814	.443
Three way Interactions			
+**Passage Relation: Emotional Valence: Suggestion Tone Style**	**2**	**3.417**	**.033**

Table 5.3: Direct main and interaction effects of passage_relation, emotional valence of passage, suggestion tone styles on privacy protection decisions (significant effects are boldfaced).

5.8.3.1 The Optimal Suggestion Tone Style depends on Privacy-Related Affect but not on General Non-Privacy-Related Affect

Within the privacy-related passage condition, we find that there are significant differences in privacy decision outcomes across the three varying privacy suggestion tone style conditions based on the emotional valence of the passage ($\chi^2(2)$ = 6.374, p =. 0413). Comparisons between the two-way interactions of *emotional valence of the passage* condition (i.e., negative Vs privacy) and *suggestion tone style* (neutral Vs assertive), revealed that the privacy protection decision outcome was significantly higher in the assertive than neutral suggestion tone style condition, especially when the passage was negatively framed ($\beta = -1.6057$, $p < .01$) (see Figure 5.9). On average, participants within the assertive privacy suggestion tone style condition, made more privacy protection decisions when the privacy-related passage was negatively framed (M = 7.73, SD = 1.38) than positively framed (M = 7.03, SD = 1.73). Thus, the odds for taking a privacy protection action were 4.98 times higher for participants in the assertive than neutral privacy suggestion tone style condition. There were no similar observable differences between privacy-related passage framing and tone style amongst participants in the neutral and passive suggestion tone style conditions (p = .775).

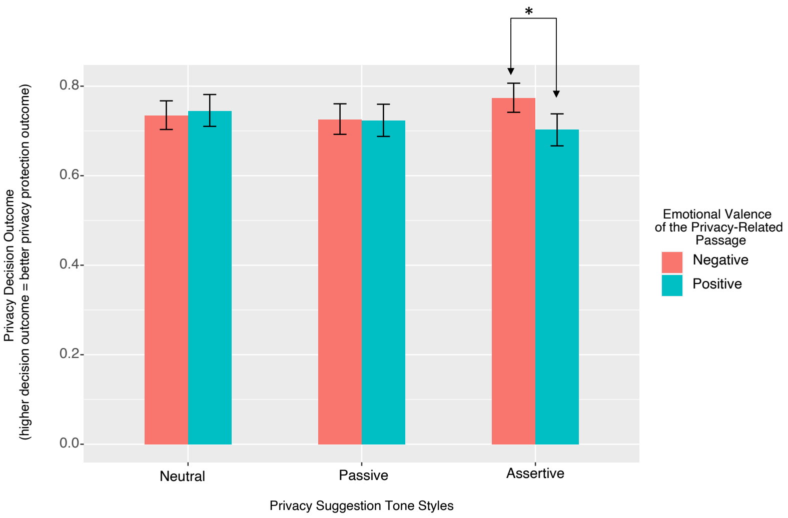


Figure 5.9: The two-way interaction effect of emotional valence of passages on participants' privacy protection decision outcomes per the privacy suggestion tone style used, specifically when primed using a privacy-related passage.

These findings indicate that a neutral tone does better when participants are in a positive privacy-related affective state. In contrast, an assertive tone does better when participants are in a negative privacy-related affective state.

For participants primed using a non-privacy-related passage, the effect of the varying privacy suggestion tone styles on participants' privacy protection decisions did not significantly differ based on the emotional valence of the passage ($\chi^2(2)$ = 0.8063, p = .668). Instead, on average, we find that participants' across the three privacy suggestion tone style conditions tended to make the same number of privacy protection decisions (M = 7.16, SD = 1.38). Nevertheless, we observe that participants in the neutral privacy suggestion tone style condition, made somewhat more privacy protection decisions, irrespective of whether the non-privacy related passage was negatively or positively framed (p = 0.332). This finding suggests that generally a neutral privacy suggestion tone style would work better than passive or assertive tones when people are in a more general affective state unrelated to social media privacy.

In the remainder of the subsections, we focus on the analysis outcomes of the differences in privacy protection decision outcomes within the privacy-related passage condition—as a primary focus of this work. These examinations helped us better understand the effects of particular post related attributes such as the privacy sensitivity of the post and presence of privacy suggestions on participants' privacy ultimate decisions. We also examine participants' specific privacy actions, user engagement, and the respective privacy actions undertaken using privacy suggestions. Overall, this analysis gives us a better understanding of the differences in decision-making when people are particularly concerned about their privacy in a permissive (more positive) or restrictive (more negative) way.

5.8.3.2 No Differences in the Moderating Effect of Privacy Related Affect based on the Privacy Sensitivity of the Post Content

In their work developing a post-content privacy sensitivity taxonomy, Li et al. [144] revealed that the privacy sensitivity of the post content could affect users' privacy decisions. For example, the authors found that many social media users did not want to share or post content they perceived to be highly sensitive [144]. Thus, we added the privacy sensitivity of the post content as a factor to test for moderations of the two-way interaction effects of the emotional valence of the passage and privacy suggestion tone styles on participants' privacy protection decisions discussed in the previous section (5.8.3.1). This subsection presents the three-way interaction results (Emotional Valence of the Passage X Suggestion Tone Style X Post Privacy Sensitivity → Privacy Protection Decision).

We found that the privacy sensitivity of the post content did not significantly moderate the tow-way interaction of emotional valence of the passage and privacy suggestion tone styles on participants' privacy protection decisions ($\chi^2(4)$ = 3.7980, p =. 434). In other words, the observed two-way interactions on participants' privacy decisions outlined in section (5.8.3.1) was the same irrespective of the privacy sensitivity of the post (i.e., low, mid, high) (see Figure 5.10). This finding suggests that participants exhibited the same decision making mechanisms irrespective of whether the content of a post was deemed to be of (low or medium or high) privacy sensitivity.

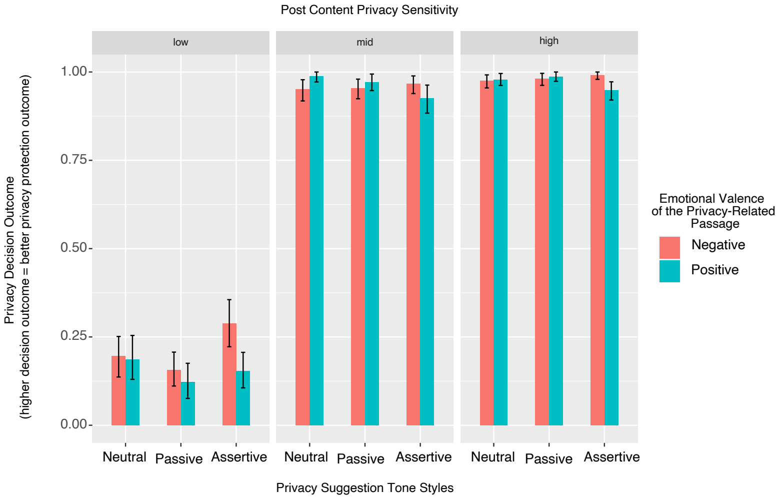


Figure 5.10: The three-way interaction between emotional valence of the passage, privacy suggestion tone style, and post content privacy sensitivity on participants' privacy protection decision outcomes, specifically when primed using a privacy-related passage.

5.8.3.3 No Differences in the Moderating Effect of Privacy Related Affect based on the Provision of Privacy Suggestions

We wanted to examine if the two-way interaction effects of emotional valence of the passage and privacy suggestion tone styles on participants' privacy protection decisions discussed in section (5.8.3.1) would only hold for posts that received a suggestion, or whether the effect would spill over to other posts as well. Privacy suggestions were only provided for three of the highly privacy sensitive posts appearing at 40, 60, 80 second time intervals, for a brief 10 seconds at a time. Thus, in this subsection, we present the three-way interaction results of (Emotional Valence of the Passage X Suggestion Tone X Privacy Suggestion_Provision→ Privacy Protection Decision).

We found that the presence of the privacy suggestions did not significantly moderate the two-way interaction of emotional valence of the passage and privacy suggestion tone styles on participants' privacy protection decisions ($\chi^2(2)$ = 5.241, p =. 073). In other words, the effect of the observed two-way-relationship between emotional valence of the passage and suggestion tone style was relatively the same irrespective of whether a post received a privacy suggestion or not (see Figure 5.11). Nonetheless, the marginally significant effect (p = .073) indicates that the moderating effect of privacy-related affect on participants' privacy protection decision outcomes was more substantial when the post received a suggestion than when it did not. For example, within the assertive privacy suggestion tone style condition, the odds of taking a privacy protection action were 1.50 times higher for participants in a negative than positive privacy-related affective state when the posts did not receive a privacy suggestion, and 1.61 times higher when the posts did receive a privacy suggestion.

Nonetheless, while the odds of taking a privacy protection action based on one's pre-existing privacy-related affective state seem to be higher for posts that did receive a privacy suggestion than those that did not, the findings suggests that the observed two-way interaction effects on privacy suggestion decision outcomes did not come about as a plain consequence of participants indiscriminately following the provided suggestions, but instead as a result of people fundamentally changing their behaviors beyond their encounter or interaction with a privacy suggestion.

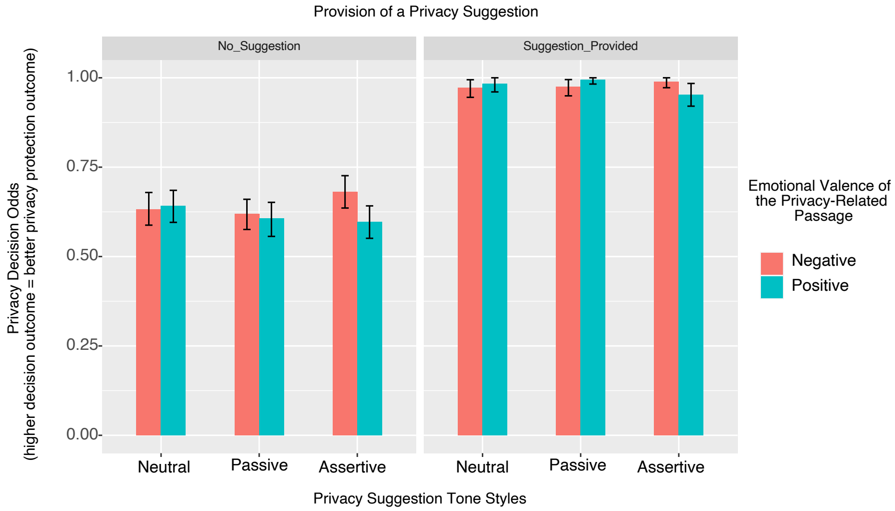


Figure 5.11: The three-way interaction between emotional valence of the passage, privacy suggestion tone style, and provision of a privacy suggestion on participants' privacy protection decision outcomes, specifically when primed using a privacy-related passage.

5.8.3.4 Summary of Effects on Privacy Protection Decision Outcomes

To summarize the findings regarding the effects of the experimental conditions (i.e., passage relation, emotional valence of the passage, and privacy suggestion tone styles) on participants' privacy protection decision outcomes:

- An Assertive privacy suggestion tone style worked better in the negative privacy-related condition, while a neutral privacy suggestion tone style worked better in the positive privacy-related condition.

- Generally, a neutral privacy suggestion tone style worked better in the non-privacy related condition, irrespective of the emotional valence of the passage.
- There were no observable differences between the neutral and passive privacy suggestion tone style on privacy protection decision outcomes.
- In the privacy-related passage condition, the observed effect of the two-way interaction between the emotional valence of the passage and privacy suggestion tone style on participants' privacy decisions did not differ based on the post content privacy sensitivity (i.e., low, medium, high). This finding indicates that the observed two-way interaction effect on participants' privacy behavior applied to all the available posts, irrespective of their level of privacy sensitivity.
- In the privacy-related passage condition, the observed effect of the two-way interaction between the emotional valence of the passage and privacy suggestion tone style on participant's privacy decisions did not differ based on the provision of a privacy suggestion. This finding indicates that the observed two-way interaction effect on participants' privacy behavior applied more broadly to all the available posts beyond the three posts that received privacy suggestions.

5.8.4 Privacy Actions and Suggestion Engagement Patterns

As discussed in Section 5.5, participants had a range of privacy actions that they could undertake to make privacy protection decisions primarily based on how privacy sensitive and reflective they found the posts to be of *them.* The particular post privacy actions ranged from doing nothing and thus leaving the post as is, editing the post's content, changing the post's audience, or deleting the post. Thus, this subsection examines the specific privacy actions undertaken and privacy suggestion engagement patterns of the participants within the privacy-related passage condition. We first report on the prevalent privacy actions.

5.8.4.1 Most of the posts were deleted, especially when deemed inappropriate or poorly reflective of the participant.

Overall, we find that a majority of participants deleted the posts they were uncomfortable with or found offensive and thus not appropriate for a potential employer or others online to see. For example, in the negative privacy-related affect condition, with an assertive privacy suggestion tone style, out of the ten available posts: 55% were deleted, 17% had their audience changed, 6% were edited, and 23% had no action taken (see Figure 5.12). Similarly, in the positive privacy-related passage, with a neutral privacy suggestion tone style; 46% of the posts were deleted, 20% had their audience changed, 8% were edited, and 26% with no action taken. Additionally, we observe that the deletion of posts was lower in general for a positive privacy-related passage while the change in the audience tended to be generally higher.

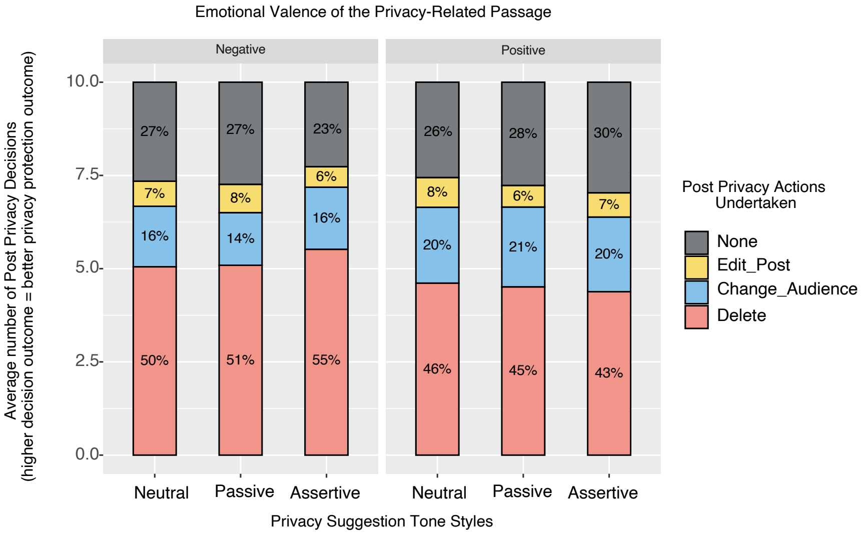


Figure 5.12: The post privacy actions undertaken across the different privacy suggestion tone style and emotional valence conditions, specifically when the participants where primed using a privacy-related passage.

An inquiry into the reasoning behind the specific privacy actions revealed that participants had varying justifications. For example, for a highly privacy-sensitive vulgar text post that read *"I work with a bunch of fucking idiots"*, participants reported finding it inappropriate, negative, and offensive to coworkers (See Figure 4 (c),Appendix B.4). Therefore, most participants reported deleting it due to being wary about presenting a wrong image or perception about themselves to a potential employer. Consequently, participants were concerned that with such a poor, tainted image of themselves, the potential employer would not hire them. The few participants who reported changing the audience stated that they would be comfortable with the shared post as it would be a true testament to how they felt in the moment. However, it is not something they think a potential employer should see, thus changing the audience to limit it to themselves or close friends. For the other few participants who reported editing the post, they stated editing it to make it more positive or friendly, as they otherwise would not be comfortable with a potential employer seeing it in its unedited state. The very few who left it as is stated that it was a spur of the moment and, therefore, a true reflection of how they felt.

5.8.5 When encountered, post privacy suggestions were more likely to be accepted than rejected

On average, if participants' encountered any of the three privacy suggestions that recommended one of the three privacy actions (change post audience, delete post, edit post), they were more likely to explicitly accept than reject the suggestion (β = 16.4921, p < .001) (see Figure 5.13). For example in the negative privacy-related passage condition, where assertive privacy suggestions were provided: among participants who encountered the change audience privacy suggestion (28.3% explicitly accepted it, 16.7% explicitly rejected it), delete post suggestion (23.3% explicitly accepted it, 18.3% explicitly rejected it), edit post suggestion (18.3% explicitly accepted it, 6.7% explicitly rejected it).

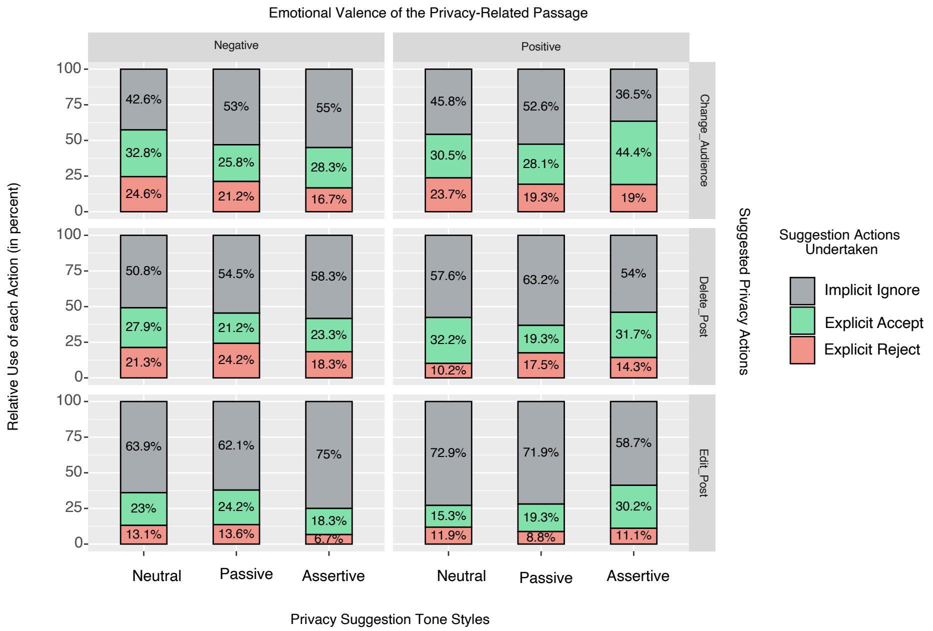


Figure 5.13: The particular actions (i.e., reactions) to privacy suggestion when participants' encountered them .

An inquiry into the reasoning behind the privacy actions undertaken revealed that indeed some participants followed the privacy suggestion when they encountered it. For example, a privacy suggestion was provided encouraging participants' to delete a post that contained a highly privacy sensitive photo that contained a photo of a dog bite (see Figure 5.4 for the details of the sample suggestion) with the text *("I got drunk and then my dog bit me. It hurts so bad!")* (see Figure 4 (a), Appendix B.4 for the corresponding post). Participants who encountered the suggestion reported following the suggestion. For instance, two participants in the positive privacy-related, neutral privacy suggestion tone style condition, stated that they deleted the post because *"Friendbook recommended deleting the post"*, *"Followed the advice of the bot on the platform."*, while two others in the positive privacy-related, assertive privacy condition stated that they deleted the post because it *"was suggested it be removed"*, *"i think it was suggested by the popup, but it speaks poorly to him [Alex Doe] by him getting drunk and being irresponsible"*.

5.8.6 In instances when post privacy suggestions were implicitly ignored or explicitly rejected , the posts were more likely to be deleted manually.

In the instances that privacy suggestions were explicitly rejected or ignored, we observe that most of the posts were likely to ultimately be deleted despite the recommended privacy action (see Figure 5.14). For example, in the negative privacy-related affect condition, with assertive privacy suggestions: among participants who encountered the “edit post” privacy suggestion (18.3% explicitly accepted, 6.7% explicitly rejected, and 75% implicitly ignored the advice) (see Figure 5.13). However, an analysis of the ultimate privacy actions undertaken for posts whose privacy suggestions were either explicitly rejected or implicitly ignored showed that 81.7% of the participants selected to delete rather than edit the post independently. A similar trend was observed among participants who encountered the “delete” and “change audience” privacy suggestions. More specifically, among the participants who encountered the “delete post” privacy suggestion (18.3% implicitly rejected and 58.3% implicitly ignored it), 51.7% of these participants ultimately decided to delete the post. Likewise, among the participants who encountered the “change audience” privacy audience suggestion (16.7% explicitly rejected and 55% implicitly ignored it). Herein, 11.7% more of the participants manually changed the audience while 56.7% chose to delete the post instead. These findings suggest that in the instances where participants either ignored or were not in agreement with the recommended privacy action, they ended up taking a more restrictive privacy action (i.e., deletion of the post) at times a privacy action not recommended by the system.

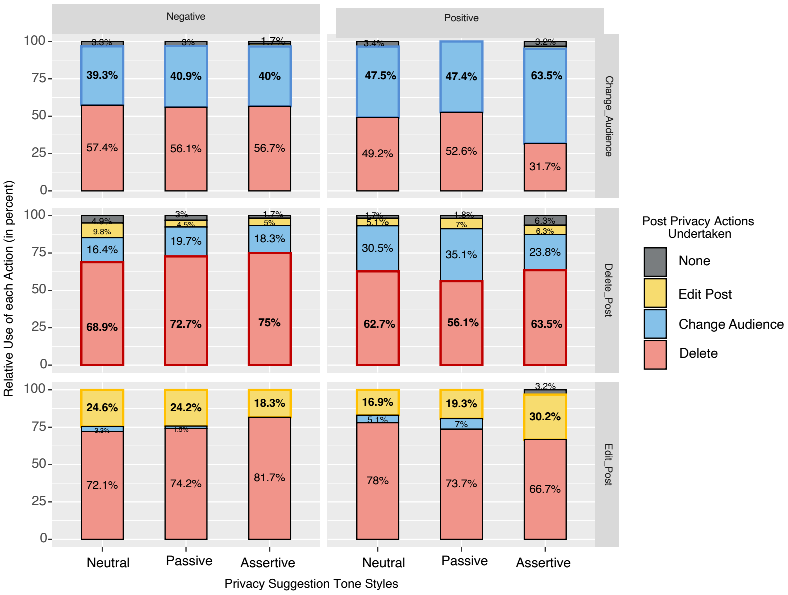


Figure 5.14: The post privacy protection actions undertaken when privacy suggestions were encountered or actions undertaken when suggestions were either ignored or explicitly rejected. Participants were more likely to follow the recommended privacy actions, otherwise tended to delete the post later on.

5.8.6.1 Summary of Effects on Privacy Actions and Suggestion Engagement Patterns

To summarize the findings regarding the participant privacy action undertaken and privacy suggestion engagement patterns:

- Majority of the posts were deleted, especially if participants deemed them to be inappropriate or poor reflective of them.
- When encountered, post privacy suggestions were more likely to be accepted than rejected.
- When post privacy suggestions were either rejected or ignored, users were more likely to delete the post independently.

5.9 Discussion

Below, we describe the impact of privacy suggestion tone styles on social media users' privacy decision-making processes depending on users' pre-existing privacy-related affect (or lack thereof), primarily based on the post-hoc behavioral analysis results in section 5.8.3. More specifically, we offer insight into the appropriate use of (*"neutral"*, *"Passive"*, *"assertive"*) tone styles in assisting users in making privacy protection decisions. We also discuss the consequences of these findings for personalized social media system designs if privacy suggestions are to be used as privacy adaptation presentation methods.

Our results show that the three varying privacy suggestion tone styles (i.e.,*neutral*, *passive*, *assertive*) indeed influenced users' privacy protection decision outcomes (RQ1). However, this effect was significantly dependent on users' *pre-existing privacy-related affective states* (RQ3), i.e., positive or negative states induced before the actual privacy protection decisive situation occurred. Furthermore, we found differences in impact that were unique to privacy-related affect that did not extend to general non-privacy-related affect. For instance, we found that the impact of the observed moderating effect of privacy-related affect on users' privacy protection decisions was not only limited to the posts that received a privacy suggestion or posts that were particularly sensitive but rather went beyond. Participants fundamentally changed their behavior throughout their interaction with the "FriendBook" system according to the privacy suggestion tone style and their affective states. We believe that once the participants received a privacy suggestion, they internalized the recommended message or action and subsequently applied it more broadly. As a result, participants did not just blindly follow the recommended privacy action but also altered their privacy behavior to match or supersede the suggested privacy actions throughout their entire interaction(s) with the platform.

Furthermore, we find the three different privacy suggestion tone styles influence users' perceived decision helpfulness of the platform, depending on how agitated they are about privacy (or not) (RQ2). However, we find that the users' experience (i.e., perceived decision helpfulness, and trust in the platform) partially mediate their ultimate privacy protection decisions (RQ4). Otherwise, as mentioned above, there is a strong direct effect of privacy-related suggestion tone styles on users' privacy protection decisions, moderated by on their privacy-related affect (RQ1). These findings are essential because we believe that message features (e.g., tone) that define privacy suggestions align well with the conditions that are prone to heuristic processing [54]. For example, privacy suggestions tend to be brief, straight to the point, and appear for a short time length (e.g., between 10-60 seconds). Thereby leaving no room for extensive detail and thus offering "little grist for the mill of systematic processing" [54, p.463]. Therefore, we observe that users who encounter privacy suggestions "may have little choice but to gravitate toward heuristic processing, " which can be susceptible to their pre-existing privacy-related affect states.

5.9.1 Impact of Privacy Suggestion Tone Styles

In this work, we demonstrate that the effect of privacy suggestion tone styles on users' privacy protection decision outcomes varies based on their positive and negative feelings, associated with the privacy decision at hand. While we are not the first to demonstrate such effects of pre-existing affective states on user's privacy decisions [111, 109, 142, 54], our work it takes a step further by investing the effect of specifically *privacy-related affect* (in comparison to general non-privacy related affect) in concert with privacy suggestion tone styles on users' privacy protection decision outcomes. Our novel contribution is that we demonstrate that this effect of linguistic tone styles differs based on users' privacy-related affect. This work advances our knowledge of the impact and relationship between privacy-related affect and privacy suggestion tone styles on multiple fronts.

More specifically, our results show that an **assertive** privacy suggestion tone style—which was phrased as a commanding request ("You should absolutely do [X]")—may lead to higher privacy-protective decision outcomes. Otherwise, (e.g. when the user things positively about privacy, or when the user does not think about privacy at all), a neutral tone is more effective in getting users to protect their privacy. As such, the impact of an assertive tone style is heavily dependent on a users' pre-existing situation-specific privacy concerns and affective state (see Figure 5.9). The fact that an assertive privacy suggestion tone can persuade people to make more privacy-protective decisions, (i.e., those whose with negative privacy concerns) who may strongly be thinking about the loss or violation of their privacy, is particularly *important* because it highlights ways in which the impact of tone on users' privacy decisions can differ simply based on their privacy-related affect (i.e., how they feel about social media privacy). As such, our work reveals that an assertive privacy suggestion tone style can be most helpful for social media users who might strongly care about their privacy but feel apprehensive (i.e., negative) or are resigned about its management [6].

Prior work in domains such as the structuring of environmental messages also provides clear examples of when an assertive tone style can be successfully employed to either persuade or encourage individuals to take action, especially when they are concerned or care about the issue at hand [131]. In particular, Kronrod et al. [131] asserts that individuals tend to respond better to pushy or commanding (i.e., assertive) requests in domains that they view as important but might need more suggestive (i.e., neutral or passive) appeals when they lack initial conviction. This suggests that if users have a strong attitude about their privacy (i.e.., feel negative about it), then an assertive rather than a neutral privacy suggestion might work best to motivate them to manage their privacy. In other words, social media users who care or are vocal about their privacy but feel helpless and resigned to managing it can be provided with privacy suggestions that embody an assertive tone as means to motivate them to take meaningful action(s) [147, 197].

Although we expected to find a difference between the effects of the **passive** privacy suggestion tone style—which was phrased as a suggestive appeal (i.e.,"Perhaps you might think about doing [X]")—on participants' privacy protection decision outcomes (compared to the neutral privacy suggestion tone style), we did not find any significant or observable differences (see Figure 5.8). This finding could suggest that from a user-centric perspective, there are no observable clear-cut distinctions between the passive and neutral privacy suggestion tone styles. Instead, both privacy suggestion tone styles could be perceived to embody a more suggestive appeal, making it difficult for users to distinguish between the phrasing of the two tone styles. Compared to the more commanding or pushy nature of an assertive tone style, the passive and neutral could be viewed together as non-assertive because they are more suggestive, polite or non-urgent. The commanding nature of an assertive tone implies that the action cannot be avoided, yet the suggestive, non-urgent nature means that the user has an option on whether to follow or ignore the recommended action [131, 167]. Thus, for future studies, these two tone styles (i.e., passive and neutral) can be perceived as having the same effects on user privacy decisions. Consequently, comparisons can be made between 'assertive" versus "non-assertive" tone styles in future work.

Furthermore, our results demonstrate that a **neutral** privacy suggestion tone style—which was phrased as a general appeal (i.e., "Do you want to do [X]")—may lead to higher privacy-protective decision outcomes; irrespective of users' pre-existing positive or negative privacy or non-privacy related affective states (although the effect is somewhat stronger for users' in a negative than a positive non-privacy related affective state. The reverse is true for when the pre-existing affect state is connected to privacy, see Figure 5.8). In other words, when social media users' pre-existing privacy-related affective states are taken into consideration, we find that a neutral tone generally functions better at motivating them to protect their privacy, especially when in a positive affective state. However, it is essential to note that this effect occurs in spite of rather than because of the privacy related affect state. In terms of the benefits of a neutral tone style, prior work in health and well-being reveals that 1) people tend to value and like neutral tone messages, and (2) such messages need to be clear, supportive, and positive enough for people to engage or partake in the recommended actions [212]. In this light, we suggest that in the *presentation* of privacy adaptations, it may be most appropriate to use a privacy suggestion that embodies a neutral tone. Such a privacy suggestion is more likely to motivate and engage users, irrespective of their pre-existing privacy related or general non-privacy-related affective state.

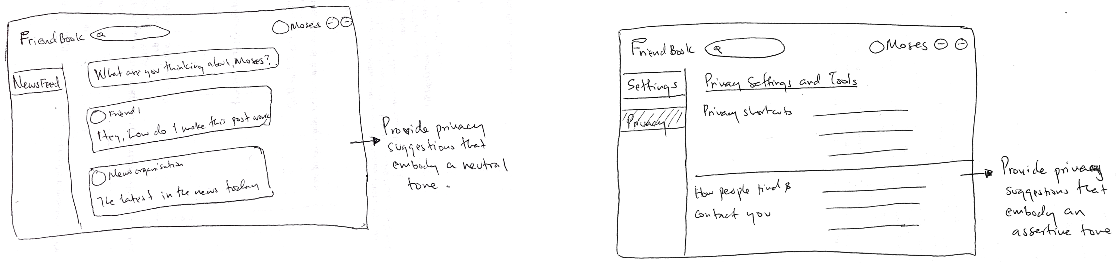
What kind of privacy suggestion tone style(s) should be used in presenting privacy adaptations? This is an important design element that has to be considered and crafted in alignment with users privacy-related affect as far as feasible. More specifically, we highlight the unique ways in which linguistic tone can impact users' privacy decision-making process/outcomes. Thus, we argue that for personalized privacy systems, linguistic tone styles are an integral component of the privacy choice structure and influence user motivation to engage or follow a recommended privacy action(s). This assertion is line with the findings of Muench et al [167], who found that in crafting messages directed at informing goal-directed behavioral interventions, individuals tend to be "sensitive to variations in the linguistic content of [the] messages designed to help them achieve a personal goal, and in some cases, have clear preferences for one type of message over another." [167, p.1]. Our work goes a step further and demonstrates that the impact of privacy suggestion tone styles on users' privacy protection decisions depends on their pre-existing privacy-related affect states. In particular, we find, in general, a neutral privacy suggestion tone style could motivate users to follow recommended privacy actions, when their privacy affect is not triggered, or when they feel positive about privacy. An assertive tone style could work best when users' are vocal about their privacy and feeling apprehensive (or negative) about it.

5.9.2 Design Implications

Our work has practical implications for the design and provision of privacy suggestions on social media platforms. As our results show, there is a variation in the effects of privacy suggestion tone styles on users' privacy protection decisions, depending on their pre-existing privacy-related affect states. Therefore, it is essential to understand and aptly use a privacy suggestion tone style that is more likely to implore people to follow the recommended action.

Determining the appropriate privacy suggestion tone styles also depends on what the platform designers want to communicate and enable the users to accomplish. Suppose a system seeks to proactively guide and support users on how to safeguard their privacy appropriately. In that case, our results suggest that the privacy suggestions should embody a tone that aligns with the users' privacy-related affective state to improve users' chances of engaging in such action. In this regard, prior work recommends tailoring the tone of the message to align with the privacy-related affect states of users [212, 66, 187]. However, tailoring privacy suggestion tone styles to align with user privacy-related affect involves providing suggestions that embody the right tone at precisely the point where they matter most (i.e., under the proper context) [24, 176].For example, on social media, Whiting and Williams [233] list ten primary objectives for use: social interaction, information seeking, passing the time, entertainment, relaxation, communicatory utility, convenience utility, expression of opinion, information sharing, and surveillance/knowledge about others. Herein, the authors point out that "managing privacy"—is seldom a primary end-user goal [7, 233]. Based on these social media use behaviors and contexts, how should designers tailor privacy suggestion tone styles if social media users are not always thinking about privacy?

The set up of our experimental conditions was meant to mirror one such prevalent behavior (i.e., using social media for network and employment opportunities) that is susceptible to context collapse [7, 233]. As such, the non-privacy versus privacy passage relation and the emotional valence of the passage experimental conditions mimic users' level of privacy concerns and related feelings. The non-privacy-related passage aspect maps onto users' regular use of social media, where they are not always thinking about privacy. Our results suggest that a neutral privacy suggestion tone would suffice in proactively encouraging or reminding people to safeguard their privacy, in situations when their privacy affect is not triggered, or when they feel positive about privacy (see Figure 5.15). The privacy-related aspect maps onto the context under which users are concerned about privacy and presumably motivated to visit the social media platforms' privacy interface/center/settings. Our results suggest that in this context, the tone style to adopt depends on how users' feel about their privacy: either in a more permissive (i.e., positive) or restrictive (i.e., negative) way. If a system can accurately predict users' affect state, then a neutral tone would suffice when they feel positive about privacy. Otherwise, then an assertive tone style would work best if they feel negative about it.


FriendBook
Moses
NewsFeed
What are you thinking about, Moses?
Friend 1
Hey, how do I make this post work
News organisation
The latest in the news today
Provide privacy suggestions that embody a neutral tone.
FriendBook
Moses
Settings
Privacy
Privacy Settings and Tools
Privacy shortcuts
How people find & contact you
Provide privacy suggestions that embody an assertive tone

(a) System designers could provide privacy suggestions that embody a neutral tone in situations when users' privacy-related affect is not triggered, or when they feel positive about privacy (e.g., when scrolling through their timeline or NewsFeed).

(b) System designers could provide privacy suggestions that embody an assertive tone in situations when users' privacy-related affect is triggered, or when they are agitated about privacy (e.g., when they visit the privacy settings or center pages).

Figure 5.15: Selectively apply neutral or assertive privacy suggestion tone styles depending on the pages social media users visit

Privacy scholars, like Vishwanath et al. [222] reveal that in most circumstances, social media users like those that use the Facebook platform tend to think about their privacy more negatively. For example, the authors state that users fear social losses stemming from inaccurate self-presentation and perceive it as a significant privacy threat. Hinds et al. [89] also report that Facebook users find privacy confusing, lack knowledge about the true privacy risks they face, and, as such are typically reluctant to update their settings due to "endless" data breaches and updates. Under such negative privacy-related circumstances, our results suggest that an assertive privacy suggestion tone can be used as a simple way to effectively communicate about the benefits of the various privacy features/settings, emphasize the relatedness of all privacy features/settings, connect the social implications of the loss of any of them, and explain how coping with one requires monitoring all the other settings [222].

There are circumstances where users can take steps to alleviate their privacy concerns and thus have a positive feeling about it (e.g., by only sharing posts with people they know) [102]. Under these circumstances, our results suggest that a neutral privacy suggestion tone can work best in motivating them to safeguard their privacy further. When people are in a positive affective state, they tend to underestimate their privacy risks [109, 142]. Therefore, a privacy suggestion tone that embodies a neutral tone style would likely be fundamental in altering their privacy behavior beyond the interaction with the suggestion. Consequently, this is likely to lead to substantial privacy protection outcomes.

5.10 Limitations and Future Work

For experimental control purposes, we put people in the scenario, having the same goal towards managing their shared posts on their timeline. Thus, we developed a semi-functional working prototype of an SNS platform with a fictitious profile to create an experience that was the same for all participants. We are cognizant that participants' interactions, decisions, and subjective experiences are susceptible to the design of the site [208] and context of use [176]. Indeed, participants may have behaved differently in our prototype with another person's profile than they would on their preferred SNS using their own profile. We made the interaction with our prototype as realistic as possible to mitigate this reduction of ecological validity needed to create a feasible and carefully controlled experimental setup.

We observed that users exhibited similar privacy decision-making approaches irrespective of the privacy sensitivity of the post (i.e., low or medium, or high) (see Section 5.8.3.2). While users must engage in privacy management strategies that they are most comfortable with to minimize privacy risk and regrets [226], this finding could also be indicative of a failure of our participants to properly discern between low, mid, and high privacy-sensitive posts. In other words, while it is desirable for people to moderate their high privacy-sensitive posts, it might not be desirable for them to take similar drastic privacy measures (e.g., deletion) for their low privacy-sensitive posts. Otherwise, this failure in discernment can easily lead to self-censorship [65, 205]. Although self-censorship "is an effective strategy to prevent regret, it also increases the chances that content that would have been safe is left unshared." [65, p.20]. Sleeper et al. [205] reveal that one of the primary reasons users self-censor is to control their self-presentation. Given that we used a job search-related scenario to motivate users to explore, engage, and review the posts on "their" timeline in our study, this scenario might have heightened the focus on self-presentation. Thus, future work should examine if the same findings are exhibited under different contexts or when a different scenario is used. Otherwise, future work should examine how users can make privacy decisions, especially related to their sensitive private posts while leaving their less sensitive posts untouched. One potential approach could entail marking or suggesting posts that they can leave or feel free to share.

SNS platforms typically contain a number of posts and related privacy features that are used over time and in different contexts. To make our study more manageable, we populated the SNS profile with only ten posts with the necessary privacy features to take related privacy actions [29]. We ensured that these privacy features had kept the same core functionality as those on Facebook.

The provided privacy suggestions only appeared for a short period of time, and once for each examined privacy action per participant. This could have affected user perception and interaction with the provided suggestions, with some participants missing or not fully comprehending the suggestion(s) due to the time-constraint. Future work can examine the appropriate appearance timing for such suggestions. The number of privacy suggestions were also limited to three so as to not overcrowd the interface and overwhelm the participants.

Furthermore, all experimental conditions had three privacy suggestions adapted for identical posts and recommended the same privacy-protective actions. Whereas this helped examine the true impact of suggestion tone styles, future work can include a state where no privacy suggestions are provided, or different posts are adapted for other privacy-protective actions.

Lastly, we recruited participants from Prolific and restricted the participant pool to only adult social media users in the U.S. We acknowledge that the demographics of such a sample participant pool may deviate from the general population of social media users. Additionally, the offered messages used to assess the impact of tone styles were written in English. Future work, can investigate the generalizability of our findings to other populations and languages.

5.11 Conclusion

This empirical study examines the impact of varying privacy suggestion tone styles—*neutral*, *passive*, and *assertive*—on users' privacy decision outcomes. We also consider users' pre-existing privacy-related affect. We find that the three varying privacy suggestion tone styles (i.e.,*neutral*, *passive*, *assertive*) indeed influence users' privacy decision outcomes. However, the nature of the effect significantly differs based on users' affective states, i.e., the mood a user is in before the actual privacy protection decisive situation occurs. In particular, we find that an assertive tone style works best when a user feels negative about privacy. Otherwise (i.e., when they feel positive about privacy or when they do not think about privacy at all) a neutral tone style will be more effective in increasing users' privacy protection behaviors. Furthermore, we observe that the impact of these effects transcends interactions with privacy suggestions alone and instead fundamentally alters users' privacy behaviors throughout their whole interaction with the social media platform. We encourage privacy researchers, designers, and developers to consider tailoring the privacy suggestions' tone to align with users' privacy-related affective states. Overall, these findings advance our knowledge of the relationship between privacy-related affect and privacy suggestion tone styles on several fronts.

Chapter 6

General Conclusion and Discussion

6.1 Summary

Motivated by the need to relieve the user burden inherent in privacy decision making and guided by the User-Tailored Privacy (UTP) adaptive privacy approach [122, 129], the work in this dissertation examined the potential of three adaptive privacy presentation methodologies—*Automation*, *Suggestion*, *Highlight*—in supporting social media users' in their privacy decision-making processe(s). More specifically, it takes an important step towards understanding the appropriate means through which a system can present personalized privacy adaptations to the user, to effectively educate, inform and support them in their privacy decision making process(es) [50, 236]. In a series of studies, we first learn about the user preferences for the trio of adaptation presentation methods that could be used to adaptively assist users with their privacy management practices on a social media site like Facebook (Chapter 3). Then, we developed a semi-working social media platform and systematically examined the effectiveness of these adaptation methods in improving user engagement and overall privacy protection (Chapter 4). Finally, in Chapter 5, we examined the "appropriate" privacy adaptation linguistic tone style in consideration of users' pre-existing concerns and affect (i.e., feelings) that could be used to motivate and support users' in their privacy decision-making process. This work was necessary since the "optimal" adaptation method needs to provide useful information and control that help users meaningfully engage with the available privacy features without overwhelming or misleading them [120].

In their work, Sheridan and Verplanks [203] proposed 10 continuum levels of automation in human-computer decision making, with lower levels representing increased autonomy of humans over computer action and higher levels representing increased autonomy of the computer over human action (i.e., fully manual performance – full automation of decision and action selection). Drawing from this 10 level scale, we identified three potential adaptation presentation methodologies that varied in the level of autonomy and control they afford to users in the privacy decision making. Namely: 1) *automation* involves the automatic application of the privacy settings by the system without user input; 2) *highlights* emphasize certain privacy features to guide users to apply the settings themselves; and 3) *suggestions* explicitly inform users about the availability of certain settings that can then be applied directly by the user [170]. Thus, in chapter 3, we examined what types of privacy features could be adapted using these adaptive presentation methods and the subsequent user perceptions towards the actual implementation of these methods. Our results revealed that the user preference for an optimal adaptation presentation method depended on the users' familiarity with the privacy feature, how they use them, and their judgement of the awkwardness and irreversibility of the implemented privacy functionality. More specifically, we found that participants generally disliked the full Automation method, except for privacy features they use frequently and perceive as inconsequential where it can alleviate some of the behavioral effort involved in the management of one's social media privacy. The Highlight method was appreciated for its ability to unobtrusively raise users' awareness about a privacy feature, and thus most suitable for features users occasionally use. The Suggestion method was preferred as a means to teach users privacy features they are unfamiliar with, unless this results in awkward suggestions or behaviors with negative social connotations. These findings not only provided concrete insight into ways in which privacy adaptations could be applied to help give users the privacy they desire [239], but also showed variation in preference for privacy feature adaptation.

Given the varying opinions on the three adaptations present methods and preference for their implementation for different privacy features, it remained unclear how exactly privacy adaptations could be applied in the development of user-tailored privacy interfaces, and if applied, how useful and effective they would be at supporting users in their privacy decision-making process. Therefore, the results in Chapter 3 served as building blocks for the experimental work in Chapter 4. For the work in this chapter, we systematically investigated the optimal user interface mechanism to present the privacy adaptation methods, effectiveness in improving users engagement with the adapted privacy features, and overall privacy protection. To overcome the shortcoming of studies with hypothetical scenarios and obtain increased ecological validity, we developed a semi-working prototype SNS platform ("FriendBook", see Figure 1, Chapter 4) through which we could alter and vary the adaptations of privacy feature using the corresponding presentation methods. Our findings revealed that the automation of privacy features afforded users the most privacy protection, while giving privacy suggestions significantly increases the level of engagement with privacy features and improves their perceptions of helpfulness and usefulness (as long as awkward suggestions are avoided). A key recommendation that adaptation methods should be **tailored** to users' awareness and prior use of the privacy features suggested in Chapter 3, did not fair any better than these two methods (i.e., privacy suggestions and automation). Instead, we found that the tailored conditions increase users' engagement (but not as much as suggestions) and protection (but not as much as automation).

Our results from Chapters 3 & 4 reveal that except for the few cases where the convenience of full automation is desired, privacy suggestions are a preferred adaptation presentation method to present privacy adaptations and inform, educate, support, and implore users to take action. Previous studies that have examined the usefulness and acceptability of such recommended privacy actions assert that message framing (i.e., the way an option or information is presented to the user) is very important [94, 57, 46]. Otherwise, users are more likely to feel resigned or detached from the recommended action, and thereby deem the privacy suggestions to be irrelevant or a nuisance [50]. Furthermore, prior work also reveals that in making of privacy decisions, users rely on heuristics (more than rational analysis) [4, 73, 3]. However, these heuristics can easily be influenced by external factors such as affect (i.e., how one feels) [151, 17]. Given the goal of supporting users achieve the privacy they desire, the appropriate tone would help in the communication of the importance and urgency of taking such recommended privacy action. Based on the user inference of such a suggestion, users could then make privacy decisions that align with how they feel about their social media privacy. Therefore, in Chapter 5, we sought to examine the role of privacy-related affect coupled with the message framing (i.e., tone) that could be used to support users in their privacy decision-making process. Our results reveal that the optimal tone embodied by the privacy suggestion depends on users' privacy-related affect(i.e., how people feel about privacy). More specifically, an assertive tone works best when a user feels negative about privacy. Otherwise (i.e., when they feel positive about privacy or when they do not think about privacy at all) a neutral tone style will be more effective in increasing users' privacy protection behaviors. In general, the results in Chapter 5 reveal that if privacy suggestions—as adaptation presentation methods—were to be effective, the framing and corresponding user privacy-related affect ought to be carefully assessed and considered.

6.1.1 Practical Considerations

In this dissertation, we set out to search for an "appropriate" adaptive privacy presentation method to support users in their privacy decision-making process. Our results reveal that privacy suggestions are the most preferred adaptation presentation method and are effective at helping users' in their privacy decisions, primarily due to the level of autonomy and control they afford [203]. However, we believe it is not entirely tenable for privacy suggestions to always be the most "appropriate" adaptation presentation method under all privacy decision-related circumstances. Indeed our findings reveal that suggestions, mainly when applied to private behaviors that carry a negative social perception on social media, such as deleting posts and unfollowing users, would not be deemed "appropriate" as they would break certain social norms. Hence, in implementing privacy adaptations, system designers might need to mix adaptation methods for the best user experience and privacy protection outcome.

In circumstances where suggestions are an inappropriate means to present privacy adaptations, what other adaptation presentation method could system designers consider? Contrary to our hypotheses, our findings reveal that highlights would be impractical and ineffective to use as adaptation presentation methods. Highlights neither improve users' level of engagement with adapted privacy features nor the overall privacy protection compared to the default where no adaptations are made to a system [169]. Instead, there are a few specific instances where automation could be the better adaptation presentation method. For example, social media platforms could use automation—as an adaptation presentation method—to set an SNS account to "private" at initiation, especially for vulnerable or novice/amateur users such as teenagers [191, 21], who tend to have limited abilities for self-regulation and complex decision-making [12, 238]. More specifically, Wisniewski [238] asserts that "technology should support teens in their developmental goals, including information seeking, learning about rules and boundaries, and maintaining social relationships, in addition to keeping them safe from online risks", highlighting the potential use case for automation in the management of teen privacy. Zhong et al. [245] also reveal that system designers can use automation to identify and help resolve potential privacy conflicts in photos, according to the involved stakeholders and their relationship(s). Herein, the ultimate goal for automation would be to provide an accurate early warning system to identify and resolve conflicting privacy preferences among photo stakeholders [245]. Mondal et al. [165] also reveal that system designers can use automation to resolve the mismatch between a users' shared post's active privacy setting and their desired setting, especially if the user wishes to limit the privacy of past posts. Kaur et al. [108] further reveal that system designers could use automation to reduce the cognitive effort required to craft and post content on different SNSs, primarily based on users' desired and anticipated system affordances (i.e., perceptions of the utilities of a target system). In this case, an intelligent system could be developed and used to automatically predict the appropriate SNS a user could use to share social media posts. For example, a user could input into the system their desired sharing preferences (e.g., " share this image to my friends and family and automatically delete it after some time" [108, p.564]) coupled with the affordances they anticipate the SNS to exhibit (e.g., "I can use Snapchat to share images and videos with my family and close friends, and the content automatically disappears after 24 hours" [108, p.564]). The intelligent system could then automatically direct the user to the ideal SNS that they can consider posting their content. However, automation is seldom used as an adaptation presentation method that can support users' privacy decisions. Instead, most SNS platforms use automation as a technical means to infer, detect and ban specific content (e.g., sensitive or offensive text [58, 192]), curate and customize users' NewsFeed [166], recommend potential new connections (i.e., "friends you may know") [58], among many other functions [13].

Additionally, when automation is used solely as a means to infer, predict or make decisions on behalf of the user, explanation(s) of the particular automated decisions or even notifications about the presence of automated decisions is strongly limited or non-existent [136, 223]. In our implementation of automation, while the system did not explicitly notify the user of the automatic adaptation—as an adaptation presentation method, we ensured that users could see that an automated action had occurred when they arrived at the location where they would have executed the action themselves. This implementation approach ensured that users were ardently notified and still had some autonomy and ability to either indicate comfort with (i.e., accept) or reverse the automated action. We suggest system designers adopt a similar approach in their implementation of automation–as an adaptation presentation method. This approach will not only help to notify users about the automated decision but also serve as an avenue to garner user input that can be used to correct and improve the accuracy of the automated decision [170, 50].

Inasmuch as automation could serve as an alternative adaptation presentation method to suggestions in certain circumstances, our results reveal that presenting adaptations to privacy features as suggestions generally tends to work better and is more helpful. However, prior work has also shown that people do not always make privacy decisions based on a systematic assessment of their choices [73]. Instead, they make privacy decisions based on heuristics that involve circumventing the conscious deliberation of information. Heuristics can be brittle in the context of new information or evolving situations or subjective factors such as affect (i.e., how one feels) [4]. Therefore, to align with how people often make privacy decisions and improve the effectiveness of privacy suggestions, system designers must consider the presentation of information within privacy suggestions (e.g., the tone) in the face of other factors that could influence their acceptability or usefulness [50]. To this end, our results in Chapter 5 reveal differences in the impact of tone in users' privacy decision-making process, depending on users' privacy-related affect. These findings suggest that in catering to users' privacy-related affective states, the privacy decision context is important to help determine the privacy suggestion framing (i.e., tone). For example, when users are posting or browsing their NewsFeed, we can assume that they are in a positive privacy-related affective state or do not think about privacy at all, in which case, a neutral tone would work best in imploring them to take the appropriate privacy action. On the other hand, if users navigate or browse to or through the privacy settings/center/help pages, it can be an indication that they are in a negative privacy-related affective state or apprehensive about their social media privacy. In this instance, an assertive tone would motivate them to take the appropriate privacy action.

Overall, in the efforts toward implementing User-Tailored Privacy (UTP), most of the research work has focused on the measurement and modeling aspects of the approach [122, 13]. Herein, prior work employs machine learning techniques to categorize or predict user privacy preferences and behaviors but does not try to then meaningfully aid people with their privacy decisions in a real interface [13, 33]. As a result, fewer efforts have been geared toward uncovering the human-centric interaction approaches that would aid in the adaptation aspect of UTP. Thus, this dissertation work is the first empirical effort to understand the feasibility of adaptation presentation methods in improving user engagement and privacy protection on social networks. I reveal that figuring out the "appropriate" adaptation presentation method is not trivial but a difficult task, susceptible to factors such as tone style and user affect.

6.2 Future Research Directions

In the examination of the adaptation presentation methods, I made several assumptions about the potential implementation of privacy adaptations that need to be examined in future work. Specifically:

- What other adaptation presentation methods are possible, and who will enforce their application? Although in this dissertation work, I relied on Sheridan and Verplanks' [203] human-automation model to identify the three possible adaptation presentation methods examined in this work. Other types of privacy-protective adaptation presentation methods could be analyzed, and the means through which they get applied could also differ. For example, based on current trends [154], it is not far-fetched to believe that in the near future privacy adaptations could either be user-driven or government-driven rather than system-driven. Akin to how in 1965 the U.S. Congress enacted the Federal Cigarette labelling and Advertising Act of 1965 (Public Law 89-92) to mandate for health warnings labels on the sides of cigarette packages (e.g., "Caution: Cigarette Smoking May Be Hazardous to Your Health") as a safety mechanism on the effects of tobacco use on people's health [237, 67]. A similar reproach could play out when it comes to privacy adaptations, with government entities (or other different actors) demanding for or mandating the use of specific privacy adaptation presentation methods. For example, the European Union has already proposed a regulation that bans "some uses of [automation], heavily regulates high-risk uses and lightly regulates less risky [automated] systems" [154]. Such laws could affect how intelligent systems use automation to protect user privacy. Therefore, future work should further explore other adaptation presentation methods and how they are applied across technological systems and regions in the world.
- What other social media privacy features could be adapted? Although I extensively examined *which* and *how* social media privacy features could be adapted in Chapter 3, the evidence remains rather restricted to a particular SNS (i.e., Facebook). Even then, Facebook has many privacy features meant to serve different privacy needs depending on the user and purpose of use. This study was only limited to a subset of prominent Facebook privacy features as previously identified by Wisniewski et al. [242]. I found that certain adaptations methods are suitable for certain privacy features (e.g., privacy suggestions should be avoided for behaviors with negative social connotations, and automation should be avoided for behaviors that might

result in real-world consequences). That said, other social media platforms and emerging forms of social interaction will continuously call for new forms of privacy control. Future work should further explore which other privacy features can be adapted, under what contexts, and across all SNS.

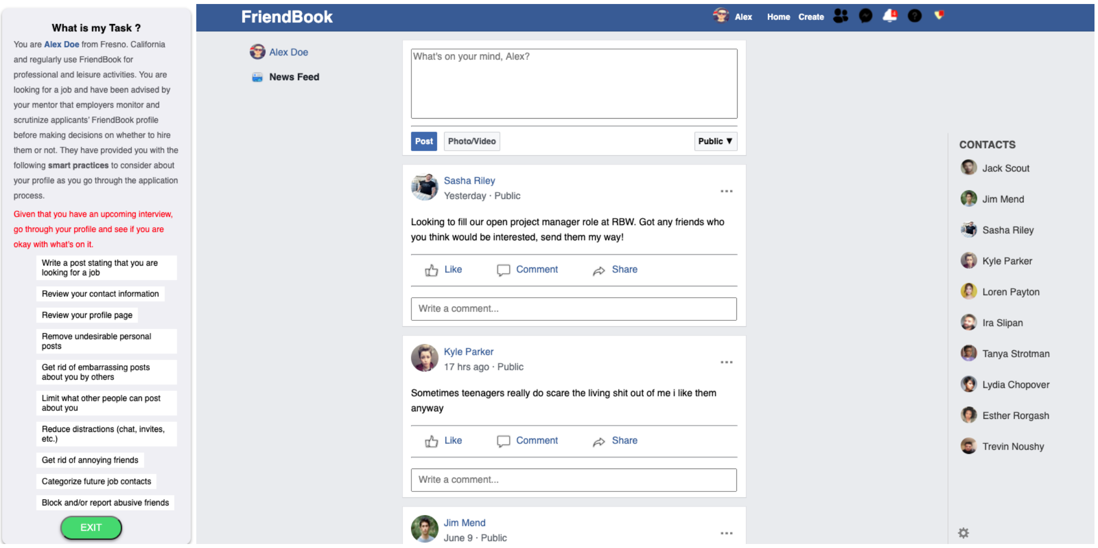
- What parameters should be considered to facilitate the proper provision of privacy adaptations presentation methods? Privacy adaptations and corresponding presentation methods can be used to proactively support users in their privacy decision-making process. However, it remains unclear which parameters should be considered for the effective implementation of privacy adaptations. My implementation of the presentation adaptation methods did not consider the additional parameters and their effects on users' management of their privacy. For example, in Chapters 3 & 4, the adaptations were by default presented or made under the assumption that they worked with 100% accuracy. In Chapter 5, there was a time interval difference between the appearances of the privacy suggestions without consideration of whether the participant needed support or not. Furthermore, we presented a limited number of privacy suggestions to avoid fatigue and monotony. A true test of the effectiveness of adaptation presentations methods would thereby require critical consideration of the contexts under which the adaptations are either made or presented to users [25]. For example, does timing of the presentation of the adaptation [163] influence users' willingness to receive privacy support or differ based on users' level of digital literacy, age, gender, etc? Is users' posting behaviors [225] a better determinant of when privacy adaptations should be made? Future work should further study the influence of these parameters on the effectiveness of the examined adaptation presentation methods.

- Is it possible to find the effects of tailoring privacy adaptations using larger sample sizes? Although, I did not find the desired effects of tailoring in Chapter 4 as I had hypothesized, this is dependent on a wide range of parameters such as sample size, access to datasets to aptly determine user preferences, among many others. In their work examining the possibility of tailoring privacy nudges to individuals' decision making and personality traits, Warberg et al. [228], assert that the effects of tailoring adaptation presentation methods might only be feasible to organizations with vast amounts of user data such as Facebook or Google. Therefore, for future work, research and design teams at social networking sites can further explore the potential of tailoring privacy adaptation to users' privacy preferences using large sample sizes.

Appendices

Appendix A Supplemental Material for Chapter 4

A.1 The Semi-Functional Social Media (“FriendBook”).



What is my Task ?

You are Alex Doe from Fresno. California and regularly use FriendBook for professional and leisure activities. You are looking for a job and have been advised by your mentor that employers monitor and scrutinize applicants' FriendBook profile before making decisions on whether to hire them or not. They have provided you with the following **smart practices** to consider about your profile as you go through the application process.

Given that you have an upcoming interview, go through your profile and see if you are okay with what's on it.

- Write a post stating that you are looking for a job
- Review your contact information
- Review your profile page
- Remove undesirable personal posts
- Get rid of embarrassing posts about you by others
- Limit what other people can post about you
- Reduce distractions (chat, invites, etc.)
- Get rid of annoying friends
- Categorize future job contacts
- Block and/or report abusive friends

Sasha Riley
Yesterday · Public

Looking to fill our open project manager role at RBW. Got any friends who you think would be interested, send them my way!

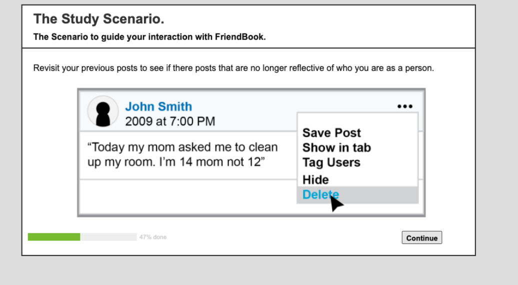
Kyle Parker
17 hrs ago · Public

Sometimes teenagers really do scare the living shit out of me i like them anyway

Jim Mend
June 9 · Public

Figure 1: The semi-functional social media platform (“FriendBook”) used in exposing participants to adapted privacy features using the adaptation methods. Free public images accessed from the internet (under a (CC0) commons creative license) and fictitious names were used in the creation of “Alex Doe's” profile. Code repo can be accessed at: https://github.com/bakman329/bakman329.github.io

A.2 Smart Practice Example



The Study Scenario.

The Scenario to guide your interaction with FriendBook.

Revisit your previous posts to see if there posts that are no longer reflective of who you are as a person.

John Smith
2009 at 7:00 PM

“Today my mom asked me to clean up my room. I'm 14 mom not 12”

Figure 2: An example of a smart practice used to orient and guide user interaction with FriendBook.

A.3 Subjective Measurement Scales

Factor	Items	Loading
Perceived Decision Help from FriendBook (based on [124]) Alpha:0.83 AVE: 0.69 Correlation: 0.858	FriendBook helped me to decide how I could use the available privacy features.	0.879
	FriendBook helped me to make a tradeoff between privacy and usefulness.	0.715
	FriendBook showed me the best ways to use the available privacy features.	0.884
Perceived Usefulness of FriendBook (based on [52]) Alpha: 0.93 AVE: 0.77 Correlation: 0.858	FriendBook enabled me to use the available privacy features more quickly.	0.824
	Using FriendBook improved the quality of the decisions I made.	0.876
	FriendBook would enhance my ability to protect my privacy online.	0.909
	Overall, I found FriendBook useful in using the available privacy features.	0.921
	FriendBook would support me in being more conscious of the things I share online.	0.851

Table 1: Items used to assess participants' subjective evaluations of the FriendBook platform, along with CFA factor loadings.

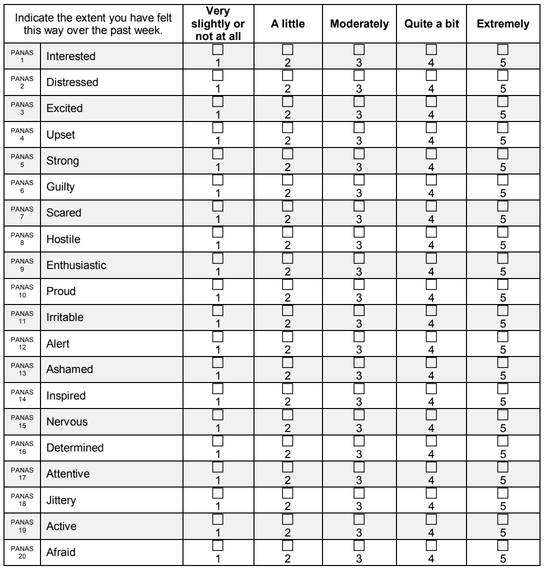
Appendix B Supplemental Material for Chapter 5

B.1 The Pre-Survey Measures

Construct	Survey Items	Scale
Social Media Use (adopted from [62])	In the past week, on average, approximately how much time **PER DAY** did you spend actively using Facebook	(5-point Likert Scale) 1) I no longer use Facebook, 2) Did not use Facebook at all within the past week, 3) Less than 10 minutes per day, 4) 10-30 minutes per day, 5) 31-60 minutes per day, 6) 1-2 hours per day, 7) 2-3 hours per day, 8) More than 3 hours per day
Privacy Feature Awareness	Are you familiar with the [Edit Audience/ Move to Trash / Edit Post] Facebook post feature?	Yes/No
Privacy Feature Usage	How often do you use this feature?	(4-point Likert Scale) 1) Never Used, 2) Used Once, 3)Occassionally Use, 4) Frequently Use

Table 2: The pre-survey items used to access social media use, privacy feature awareness and usage (see Outcomes in Section 5.8.1.1).

B.2 The Positive and Negative Affect Schedule (PANAS-SF)



Indicate the extent you have felt this way over the past week.		Very slightly or not at all	A little	Moderately	Quite a bit	Extremely
PANAS 1	Interested	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 2	Distressed	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 3	Excited	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 4	Upset	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 5	Strong	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 6	Guilty	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 7	Scared	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 8	Hostile	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 9	Enthusiastic	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 10	Proud	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 11	Irritable	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 12	Alert	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 13	Ashamed	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 14	Inspired	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 15	Nervous	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 16	Determined	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 17	Attentive	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 18	Jittery	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 19	Active	□ 1	□ 2	□ 3	□ 4	□ 5
PANAS 20	Afraid	□ 1	□ 2	□ 3	□ 4	□ 5

Figure 3: After reading the passage, participants will be asked "how the passage made them feel" to examine the emotion valence using a 5-point Likert PANAS-SF scale (ranging from Not at all - Extremely) [229].

Positive Affect Score: Add the scores on items 1, 3, 5, 9, 10, 12, 14, 16, 17, and 19. Scores can range from 10 – 50, with higher scores representing higher levels of positive affect.

Negative Affect Score: Add the scores on items 2, 4, 6, 7, 8, 11, 13, 15, 18, and 20. Scores can range from 10 – 50, with lower scores representing lower levels of negative affect.

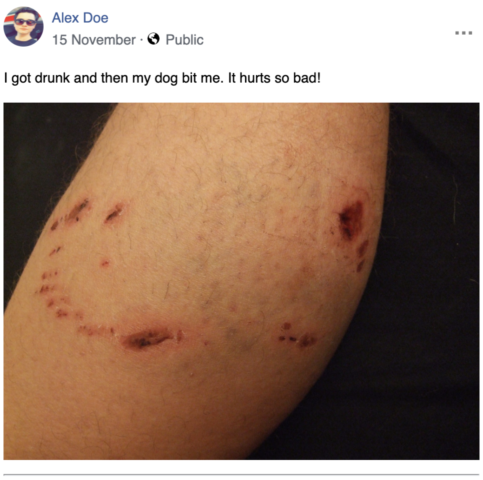
B.3 Post-Survey Subjective Measures

Construct	Survey Items	Loadings
Perceived Decision help from system(PD) (Adapted from [124]) Alpha: 0.886 AVE: 0.76 Correlation w/TR: 0.545 Correlation w/IPC: -0.087	[FriendBook's] suggestions helped me to decide how I could use the available privacy features.	0.906
	[FriendBook's] suggestions helped me to make a trade-off between privacy and usefulness.	0.836
	I felt clueless about how to use the available post-privacy features on [FriendBook].	
	[FriendBook's] suggestions showed me the best ways to use the available post-privacy features.	0.866
Perceived Trust in Platform (TR) (Adapted from [130]) Alpha: 0.946 AVE: 0.80 Correlation w/IPC: -0.242	I believe [FriendBook] would be open and receptive to the needs of its users.	0.790
	I believe [FriendBook] would make good-faith efforts to address most of its users' concerns.	0.883
	I believe [FriendBook] would be interested in the well-being of its members, not just its own.	0.904
	I believe [FriendBook] would be honest in its dealings with me.	0.917
	I believe [FriendBook] would keep its commitment to its users.	0.927
	I believe [FriendBook] would be trustworthy.	0.935
Informational Privacy Concerns (IPC) (Adapted from [157]) Alpha: 0.918 AVE: 0.81	It usually bothers me when online companies ask me for personal information.	0.878
	When online companies ask me for personal information, I sometimes think twice before providing it.	0.861
	It bothers me to give personal information to so many online companies	0.949
	I'm concerned that online companies are collecting too much personal information about me.	0.910

Table 3: Items used to assess participants' subjective evaluations of the FriendBook platform, along with CFA factor loadings. Items with no loading had a low factor loading (< .7)

B.4 A Gallery of the Ten Posts Used to populate “FriendBook”

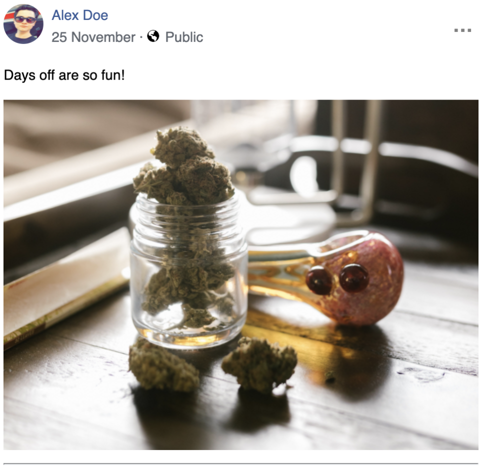
High Privacy Sensitive Posts


Alex Doe

15 November · Public

I got drunk and then my dog bit me. It hurts so bad!

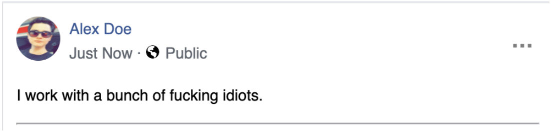
(a) Adapted with a privacy suggestion that recommended for the “deletion of the post.”


Alex Doe

25 November · Public

Days off are so fun!

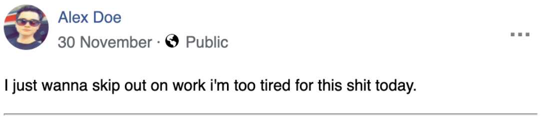
(b) Adapted with a privacy suggestion that recommended for the “change of the post audience.”


Alex Doe

Just Now · Public

I work with a bunch of fucking idiots.

(c) Adapted with a privacy suggestion that recommended for the “editing of the post.”

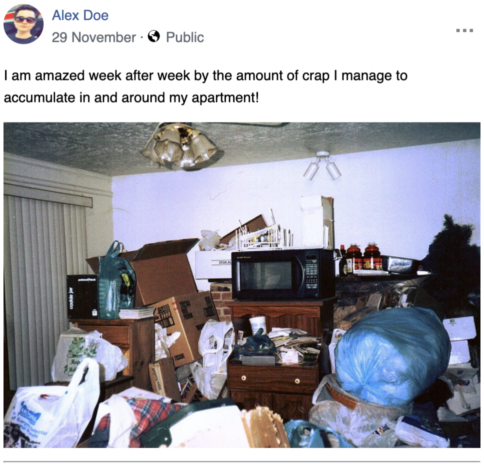

Alex Doe

30 November · Public

I just wanna skip out on work i'm too tired for this shit today.

(d) **No privacy suggestion was provided.**

Mid Privacy Sensitive Posts

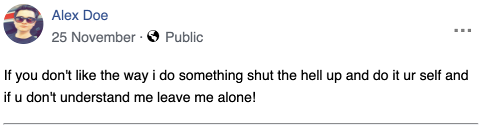


Alex Doe

29 November · Public

I am amazed week after week by the amount of crap I manage to accumulate in and around my apartment!

(e) No privacy suggestion was provided.

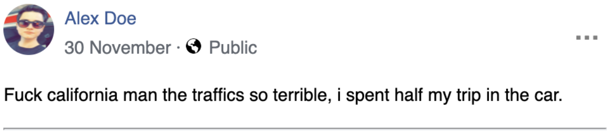


Alex Doe

25 November · Public

If you don't like the way i do something shut the hell up and do it ur self and if u don't understand me leave me alone!

(f) No privacy suggestion was provided.



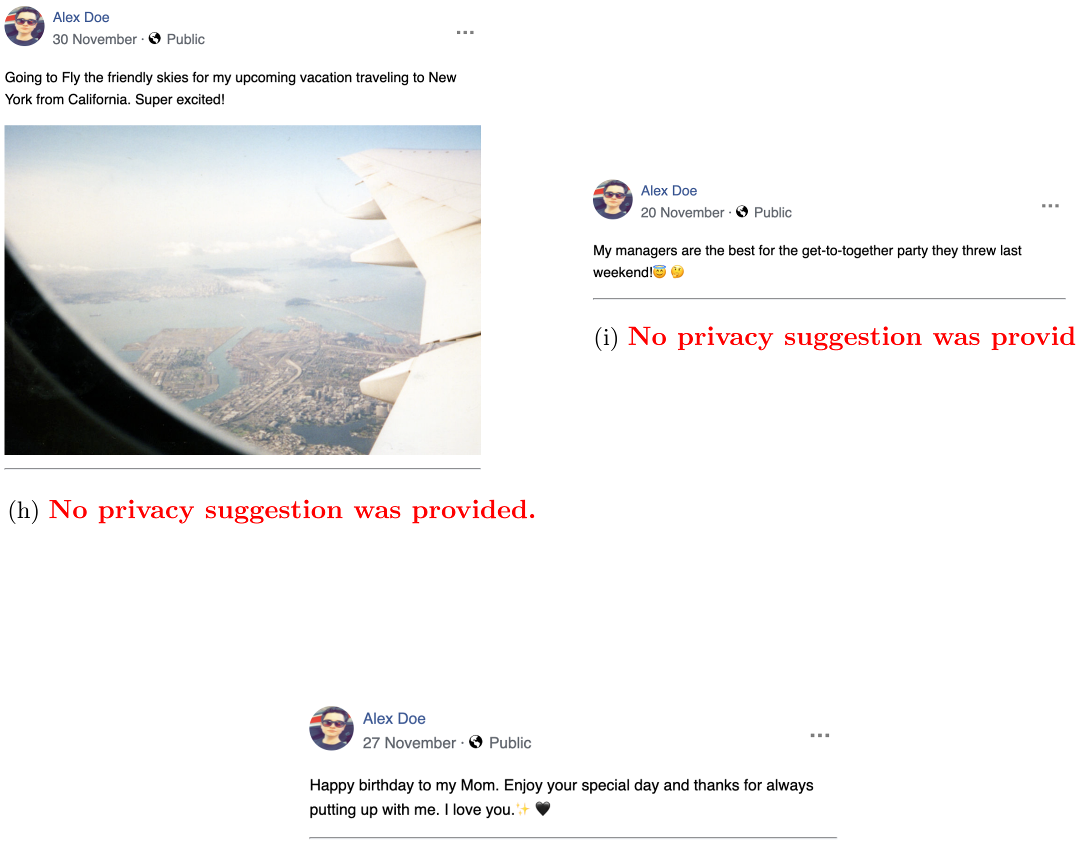
Alex Doe

30 November · Public

Fuck california man the traffics so terrible, i spent half my trip in the car.

(g) No privacy suggestion was provided.

Low Privacy Sensitive Posts



Alex Doe
30 November · Public

Going to Fly the friendly skies for my upcoming vacation traveling to New York from California. Super excited!

(h) **No privacy suggestion was provided.**

Alex Doe
20 November · Public

My managers are the best for the get-to-together party they threw last weekend!😇 🤔

(i) **No privacy suggestion was provided.**

Alex Doe
27 November · Public

Happy birthday to my Mom. Enjoy your special day and thanks for always putting up with me. I love you.✨ 🖤

(j) **No privacy suggestion was provided.**

Figure 4: The ten posts used to populate "Friendbook" (see Figure 5.3). The posts were varied in privacy sensitivity (low, mid, high) based on the photo & content sensitivity taxonomy proposed by Li et al. [144]. Note: Before participants could provide consent to partake in the study, they were forewarned about the possible encounter of post content that might be vulgar, relate to medical conditions, or express negative attitudes towards work. Code repo can be accessed at: https://github.com/henryksloan/FriendBook

B.5 For Replication Purposes

PS: For replication purposes or to access the study material of the work in referenced in chapters 3, 4, and 5, please visit (https://drive.google.com/drive/folders/1UEDwIlxrOmY1LisRiSpqnj-VJDYEtjfr?usp=sharing)

Bibliography

[1] Patricia Sánchez Abril, Avner Levin, and Alissa Del Riego. Blurred boundaries: Social media privacy and the twenty-first-century employee. *Am. Bus. LJ*, 49:63, 2012.

[2] Alessandro Acquisti. Protecting privacy with economics: Economic incentives for preventive technologies in ubiquitous computing environments. In *Workshop on Socially-informed Design of Privacy-enhancing Solutions, 4th International Conference on Ubiquitous Computing (UBICOMP 02)*, 2002.

[3] Alessandro Acquisti. Nudging privacy: The behavioral economics of personal information. *IEEE security & privacy*, 7(6):82–85, 2009.

[4] Alessandro Acquisti, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie Faith Cranor, Saranga Komanduri, Pedro Giovanni Leon, Norman Sadeh, Florian Schaub, Manya Sleeper, et al. Nudges for privacy and security: Understanding and assisting users' choices online. *ACM Computing Surveys (CSUR)*, 50(3):1–41, 2017.

[5] Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. Privacy and human behavior in the age of information. *Science*, 347(6221):509–514, 2015.

[6] Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. Secrets and likes: The drive for privacy and the difficulty of achieving it in the digital age. *Journal of Consumer Psychology*, 30(4):736–758, 2020.

[7] Alessandro Acquisti and Ralph Gross. Imagined communities: Awareness, information sharing, and privacy on the facebook. In *International workshop on privacy enhancing technologies*, pages 36–58. Springer, 2006.

[8] Alessandro Acquisti and Jens Grossklags. Privacy and rationality in individual decision making. *IEEE security & privacy*, 3(1):26–33, 2005.

[9] Alessandro Acquisti and Jens Grossklags. What can behavioral economics teach us about privacy. *Digital privacy: theory, technologies and practices*, 18:363–377, 2007.

[10] Erfan Aghasian, Saurabh Garg, Longxiang Gao, Shui Yu, and James Montgomery. Scoring users' privacy disclosure across multiple online social networks. *IEEE access*, 5:13118–13130, 2017.

[11] Tawfiq Alashoor, Nader Al-Maidani, and Ibrahim Al-Jabri. The privacy calculus under positive and negative mood states. *In Proc. ICIS*, 2018.

[12] Dustin Albert and Laurence Steinberg. Judgment and decision making in adolescence. *Journal of research on Adolescence*, 21(1):211–224, 2011.

[13] J Alemany, E Del Val, and A García-Fornes. A review of privacy decision-making mechanisms in online social networks. *ACM Computing Surveys (CSUR)*, 55(2):1–32, 2022.

[14] Reza Ghaiumy Anaraky, Tahereh Nabizadeh, Bart P Knijnenburg, and Marten Risius. Reducing default and framing effects in privacy decision-making. *Proceedings of the Special Interest Group On Humancomputer Interaction*, 2018.

[15] Catherine L Anderson and Ritu Agarwal. The digitization of healthcare: boundary risks, emotion, and consumer willingness to disclose personal health information. *Information Systems Research*, 22(3):469–490, 2011.

[16] Monica Anderson and Emily A. Vogels. Americans turn to technology during covid-19 outbreak, say an outage would be a problem, Jul 2020.

[17] Amanda D Angie, Shane Connelly, Ethan P Waples, and Vykinta Kligyte. The influence of discrete emotions on judgement and decision-making: A meta-analytic review. *Cognition & Emotion*, 25(8):1393–1422, 2011.

[18] Annie I Antón, Julia B Earp, and Jessica D Young. How internet users' privacy concerns have evolved since 2002. *IEEE Security & Privacy*, 8(1):21–27, 2010.

[19] Pekka Aula. Social media, reputation risk and ambient publicity management. *Strategy & leadership*, 2010.

[20] Brooke Auxier, Lee Rainie, Monica Anderson, Andrew Perrin, Madhu Kumar, and Erica Turner. Americans and privacy: Concerned, confused and feeling lack of control over their personal information, Aug 2020.

[21] Karla Badillo-Urquiola, Chhaya Chouhan, Stevie Chancellor, Munmun De Choudhary, and Pamela Wisniewski. Beyond parental control: designing adolescent online safety apps using value sensitive design. *Journal of adolescent research*, 35(1):147–175, 2020.

[22] Richard P Bagozzi, Mahesh Gopinath, and Prashanth U Nyer. The role of emotions in marketing. *Journal of the academy of marketing science*, 27(2):184–206, 1999.

[23] Paritosh Bahirat, Yangyang He, Abhilash Menon, and Bart Knijnenburg. A data-driven approach to developing iot privacy-setting interfaces. In *23rd International Conference on Intelligent User Interfaces*, pages 165–176, 2018.

[24] Paritosh Bahirat, Martijn Willemsen, Yangyang He, Qizhang Sun, and Bart Knijnenburg. Overlooking context: How do defaults and framing reduce deliberation in smart home privacy decision-making? In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–18, 2021.

[25] Rebecca Balebako, Florian Schaub, Idris Adjerid, Alessandro Acquisti, and Lorrie Cranor. The impact of timing on the salience of smartphone app privacy notices. In *Proceedings of the 5th Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices*, pages 63–74, 2015.

[26] Debarati Bandyopadhyay, VS Chandrasekhar Pammi, and Narayanan Srinivasan. Role of affect in decision making. *Progress in brain research*, 202:37–53, 2013.

[27] Laurie J Barclay and Tina Kiefer. Approach or avoid? exploring overall justice and the differential effects of positive and negative emotions. *Journal of management*, 40(7):1857–1898, 2014.

[28] Susanne Barth and Menno DT De Jong. The privacy paradox–investigating discrepancies between expressed privacy concerns and actual online behavior–a systematic literature review. *Telematics and informatics*, 34(7):1038–1058, 2017.

[29] Lujo Bauer, Lorrie Faith Cranor, Saranga Komanduri, Michelle L Mazurek, Michael K Reiter, Manya Sleeper, and Blase Ur. The post anachronism: The temporal dimension of facebook privacy. In *Proceedings of the 12th ACM Workshop on Workshop on Privacy in the Electronic Society*, pages 1–12, 2013.

[30] Michael Benisch, Patrick Gage Kelley, Norman Sadeh, and Lorrie Faith Cranor. Capturing location-privacy preferences: quantifying accuracy and user-burden tradeoffs. *Personal and Ubiquitous Computing*, 15(7):679–694, 2011.

[31] Peter M Bentler and Douglas G Bonett. Significance tests and goodness of fit in the analysis of covariance structures. *Psychological bulletin*, 88(3):588, 1980.

[32] Annika Bergström. Online privacy concerns: A broad approach to understanding the concerns of different groups for different uses. *Computers in Human Behavior*, 53:419–426, 2015.

[33] Reuben Binns. Algorithmic accountability and public reason. *Philosophy & technology*, 31(4):543–556, 2018.

[34] Herbert Bless, Gerd Bohner, Norbert Schwarz, and Fritz Strack. Mood and persuasion: A cognitive response analysis. *Personality and social psychology bulletin*, 16(2):331–345, 1990.

[35] Herbert Bless, Diane M Mackie, and Norbert Schwarz. Mood effects on attitude judgments: Independent effects of mood before and after message elaboration. *Journal of personality and social psychology*, 63(4):585, 1992.

[36] Vanessa Boothroyd, Ester Moher, and Khaled El Emam. Protecting personal health information: The roles of context, framing and priming in privacy-related choices.

[37] Dana Boyd and Eszter Hargittai. Facebook privacy settings: Who cares? *First Monday*, 2010.

[38] Danah M Boyd and Nicole B Ellison. Social network sites: Definition, history, and scholarship. *Journal of computer-mediated Communication*, 13(1):210–230, 2007.

[39] Scott Brave, Clifford Nass, and Kevin Hutchinson. Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. *International journal of human-computer studies*, 62(2):161–178, 2005.

[40] Laura F Bright, Susan Bardi Kleiser, and Stacy Landreth Grau. Too much facebook? an exploratory examination of social media fatigue. *Computers in Human Behavior*, 44:148–155, 2015.

[41] Penelope Brown, Stephen C Levinson, and Stephen C Levinson. *Politeness: Some universals in language usage*, volume 4. Cambridge university press, 1987.

[42] Isabel Cachola, Eric Holgate, Daniel Preoţiuc-Pietro, and Junyi Jessy Li. Expressively vulgar: The socio-dynamics of vulgarity and its effects on sentiment analysis in social media. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 2927–2938, 2018.

[43] David Charles Castano. *Affect and Online Privacy Concerns*. PhD thesis, Nova Southeastern University, 2015.

[44] Pew Research Center. 10 facts about americans and facebook. *Pew Research Center: Internet, Science & Tech*, June 2021.

[45] Pew Research Center. Demographics of social media users and adoption in the united states. *Pew Research Center: Internet, Science & Tech*, Apr 2021.

[46] Hyojin Chin, Lebogang Wame Molefi, and Mun Yong Yi. Empathy is all you need: How a conversational agent should respond to verbal abuse. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2020.

[47] Gerald L Clore. Cognitive phenomenology: Feelings and the construction of judgment. *The construction of social judgments*, 10:133–163, 1992.

[48] Gerald L Clore, Norbert Schwarz, and Michael Conway. *Affective causes and consequences of social information processing.* Lawrence Erlbaum Associates, Inc, 1994.

[49] Camille Cobb, Lucy Simko, Tadayoshi Kohno, and Alexis Hiniker. User experiences with online status indicators. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2020.

[50] Jessica Colnago, Yuanyuan Feng, Tharangini Palanivel, Sarah Pearman, Megan Ung, Alessandro Acquisti, Lorrie Faith Cranor, and Norman Sadeh. Informing the design of a personalized privacy assistant for the internet of things. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2020.

[51] Richard E Cytowic. *The neurological side of neuropsychology*. MIT Press, 1996.

[52] Fred D Davis. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, pages 319–340, 1989.

[53] Ratan Dey, Zubin Jelveh, and Keith Ross. Facebook users have become much more private: A large-scale study. In *2012 IEEE International Conference on Pervasive Computing and Communications Workshops*, pages 346–352. IEEE, 2012.

[54] James Price Dillard and Eugenia Peck. Affect and persuasion: Emotional responses to public service announcements. *Communication Research*, 27(4):461–495, 2000.

[55] James Price Dillard and Lijiang Shen. On the nature of reactance and its role in persuasive health communication. *Communication Monographs*, 72(2):144–168, 2005.

[56] Soussan Djamasbi, Diane M Strong, and Mark Dishaw. Affect and acceptance: Examining the effects of positive mood on the technology acceptance model. *Decision Support Systems*, 48(2):383–394, 2010.

[57] Evan Doyle and YoungAh Lee. Context, context, context: Priming theory and attitudes towards corporations in social media. *Public relations review*, 42(5):913–919, 2016.

[58] Natasha Duarte, Emma Llanso, and Anna Loup. Mixed messages? the limits of automated social media content analysis. 2017.

[59] William Dutton, Ginette Law, Gillian Bolsover, and Soumitra Dutta. The internet trust bubble: Global values, beliefs and practices. Technical report, World Economic Forum, 2014.

[60] Paul Ekman. An argument for basic emotions. *Cognition & emotion*, 6(3-4):169–200, 1992.

[61] Nicole B Ellison, Jessica Vitak, Charles Steinfield, Rebecca Gray, and Cliff Lampe. Negotiating privacy concerns and social capital needs in a social media environment. In *Privacy online*, pages 19–32. Springer, 2011.

[62] Sindhu Kiranmai Ernala, Moira Burke, Alex Leavitt, and Nicole B Ellison. How well do people report time spent on facebook? an evaluation of established survey questions with recommendations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2020.

[63] Seyedeh Maryam Fakhrhosseini and Myounghoon Jeon. Affect/emotion induction methods. In *Emotions and affect in human factors and human-computer interaction*, pages 235–253. Elsevier, 2017.

[64] Lujun Fang and Kristen LeFevre. Privacy wizards for social networking sites. In *Proceedings of the 19th international conference on World wide web*, pages 351–360, 2010.

[65] Bruce Ferwerda, Markus Schedl, and M Tkalčič. To post or not to post: The effects of persuasive cues and group targeting mechanisms on posting behavior. In *2014 ASE BIGDATA/SOCIALCOM/CYBERSECURITY Conference, Stanford University, May 27-31, 2014*. -, 2014.

[66] Brianna S Fjeldsoe, Alison L Marshall, and Yvette D Miller. Behavior change interventions delivered by mobile telephone short-message service. *American journal of preventive medicine*, 36(2):165–173, 2009.

[67] Food, Drug Administration, et al. Tobacco products; required warnings for cigarette packages and advertisements. *Federal Register*, 2020.

[68] Joseph P Forgas. Mood effects on decision making strategies. *Australian journal of Psychology*, 41(2):197–214, 1989.

[69] Joseph P Forgas. When sad is better than happy: Negative affect can improve the quality and effectiveness of persuasive messages and social influence strategies. *Journal of experimental social psychology*, 43(4):513–528, 2007.

[70] Nico H Frijda et al. *The emotions.* Cambridge University Press, 1986.

[71] Nitika Garg. *The role of affect in judgment and decision making.* PhD thesis, University of Pittsburgh, 2004.

[72] Jennifer M George and Gareth R Jones. The experience of work and turnover intentions: Interactive effects of value attainment, job satisfaction, and positive mood. *Journal of Applied Psychology*, 81(3):318, 1996.

[73] Reza Ghaiumy Anaraky, Kaileigh Angela Byrne, Pamela J Wisniewski, Xinru Page, and Bart Knijnenburg. To disclose or not to disclose: examining the privacy decision-making processes of older vs. younger adults. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2021.

[74] Jennifer Golbeck and Matthew Louis Mauriello. User perception of facebook app data access: A comparison of methods and privacy concerns. *Future Internet*, 8(2):9, 2016.

[75] Gian C Gonzaga, Dacher Keltner, Esme A Londahl, and Michael D Smith. Love and the commitment problem in romantic relations and friendship. *Journal of personality and social psychology*, 81(2):247, 2001.

[76] Amir Grinstein and Ann Kronrod. Does sparing the rod spoil the child? how praising, scolding, and an assertive tone can encourage desired behaviors. *Journal of Marketing Research*, 53(3):433–441, 2016.

[77] James J Gross. Emotion regulation: Current status and future prospects. *Psychological inquiry*, 26(1):1–26, 2015.

[78] Ralph Gross and Alessandro Acquisti. Information revelation and privacy in online social networks. In *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*, pages 71–80, 2005.

[79] Anatoliy Gruzd and Ángel Hernández-García. Privacy concerns and self-disclosure in private and public uses of social media. *Cyberpsychology, Behavior, and Social Networking*, 21(7):418–428, 2018.

[80] John Guare. *Six degrees of separation: A play.* Vintage, 1990.

[81] Oliver L Haimson, Albert J Carter, Shanley Corvite, Brookelyn Wheeler, Lingbo Wang, Tianxiao Liu, and Alexxus Lige. The major life events taxonomy: Social readjustment, social media information sharing, and online network separation during times of life transition. *Journal of the Association for Information Science and Technology*, 2021.

[82] Kotaro Hara, Abigail Adams, Kristy Milland, Saiph Savage, Chris Callison-Burch, and Jeffrey P Bigham. A data-driven analysis of workers' earnings on amazon mechanical turk. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pages 1–14, 2018.

[83] Eszter Hargittai. Whose space? differences among users and non-users of social network sites. *Journal of computer-mediated communication*, 13(1):276–297, 2007.

[84] Yangyang He, Paritosh Bahirat, Bart P. Knijnenburg, and Abhilash Menon. A Data-Driven Approach to Designing for Privacy in Household IoT. *ACM Trans. Interact. Intell. Syst.*, 10(1):10:1–10:47, September 2019.

[85] Justin Helper and Stacy Blasiola. Users' top-of-mind privacy concerns. Technical report, TTC Labs from Meta, 2021.

[86] Justin Helper and Maryhope Rutherford. Privacy concerns are similar across different apps.

[87] Kristen Herhold. How people view facebook after the cambridge analytica data breach.

[88] Khe Foon Hew. Students' and teachers' use of facebook. *Computers in human behavior*, 27(2):662–676, 2011.

[89] Joanne Hinds, Emma J Williams, and Adam N Joinson. "it wouldn't happen to me": Privacy concerns and perspectives following the cambridge analytica scandal. *International Journal of Human-Computer Studies*, 143:102498, 2020.

[90] Ron Hirschprung, Eran Toch, Frank Bolton, and Oded Maimon. A methodology for estimating the value of privacy in information disclosure systems. *Computers in Human Behavior*, 61:443–453, 2016.

[91] Donna L Hoffman, Thomas P Novak, and Marcos Peralta. Building consumer trust online. *Communications of the ACM*, 42(4):80–85, 1999.

[92] Silas Hsu, Kristen Vaccaro, Yin Yue, Aimee Rickman, and Karrie Karahalios. Awareness, navigation, and use of feed control settings online. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2020.

[93] Li-tze Hu and Peter M Bentler. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1):1–55, 1999.

[94] Tianran Hu, Anbang Xu, Zhe Liu, Quanzeng You, Yufan Guo, Vibha Sinha, Jiebo Luo, and Rama Akkiraju. Touch your heart: A tone-aware chatbot for customer care on social media. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pages 1–12, 2018.

[95] Jim Hudson. Consumer journeys in becoming privacy conscious. Technical report, TTC Labs from Meta, 2021.

[96] Larry M Hyman. Word-prosodic typology. *Phonology*, pages 225–257, 2006.

[97] Giovanni Iachello and Jason Hong. *End-user privacy in human-computer interaction*, volume 1. Now Publishers Inc, 2007.

[98] Jim Isaak and Mina J Hanna. User data privacy: Facebook, cambridge analytica, and privacy protection. *Computer*, 51(8):56–59, 2018.

[99] Alice M Isen. Some ways in which positive affect influences decision making and problem solving. *Handbook of emotions*, 3:548–573, 2008.

[100] Sirkka L Jarvenpaa, Noam Tractinsky, and Lauri Saarinen. Consumer trust in an internet store: A cross-cultural validation. *Journal of Computer-Mediated Communication*, 5(2):JCMC526, 1999.

[101] Eric J Johnson and Amos Tversky. Affect, generalization, and the perception of risk. *Journal of personality and social psychology*, 45(1):20, 1983.

[102] Maritza Johnson, Serge Egelman, and Steven M Bellovin. Facebook and privacy: it's complicated. In *Proceedings of the eighth symposium on usable privacy and security*, pages 1–15, 2012.

[103] Adam N Joinson, Ulf-Dietrich Reips, Tom Buchanan, and Carina B Paine Schofield. Privacy, trust, and self-disclosure online. *Human–Computer Interaction*, 25(1):1–24, 2010.

[104] Yoonhyuk Jung and Jonghwa Park. An investigation of relationships among privacy concerns, affective responses, and coping behaviors in location-based services. *International Journal of Information Management*, 43:15–24, 2018.

[105] Chris Kahn and David Ingram. Three-quarters facebook users as active or more since privacy scandal: Reuters/ipsos poll, May 2018.

[106] Frigyes Karinthy. Chain-links. *Everything is different*, pages 21–26, 1929.

[107] David Katz, Ann Kronrod, Amir Grinstein, and Udi Nisan. Still waters run deep: Comparing assertive and suggestive language in water conservation campaigns. *Water*, 10(3):275, 2018.

[108] Harmanpreet Kaur, Cliff Lampe, and Walter S Lasecki. Using affordances to improve ai support of social media posting decisions. In *Proceedings of the 25th International Conference on Intelligent User Interfaces*, pages 556–567, 2020.

[109] Flavius Kehr, Tobias Kowatsch, Daniel Wentzel, and Elgar Fleisch. Blissfully ignorant: the effects of general privacy concerns, general institutional trust, and affect in the privacy calculus. *Information Systems Journal*, 25(6):607–635, 2015.

[110] Flavius Kehr, Tobias Kowatsch, Daniel Wentzel, and Elgar Fleisch. Thinking styles and privacy decisions: need for cognition, faith into intuition, and the privacy calculus. *International Conference on Wirtschaftsinformatik (WI 2015)*, 2015.

[111] Flavius Kehr, Daniel Wentzel, Tobias Kowatsch, and Elgar Fleisch. Rethinking privacy decisions: pre-existing attitudes, pre-existing emotional states, and a situational privacy calculus. *In Proc. ECIS*, 2015.

[112] Flavius Kehr, Daniel Wentzel, and Peter Mayer. Rethinking the privacy calculus: on the role of dispositional factors and affect. 2013.

[113] Dan J Kim, Donald L Ferrin, and H Raghav Rao. A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents. *Decision support systems*, 44(2):544–564, 2008.

[114] Yeonshin Kim, Tae Hyun Baek, Sukki Yoon, Sangdo Oh, and Yung Kyun Choi. Assertive environmental advertising and reactance: Differences between south koreans and americans. *Journal of Advertising*, 46(4):550–564, 2017.

[115] Marshall Kirkpatrick. Facebook's zuckerberg says the age of privacy is over, Jan 2010.

[116] Angeliki Kitsiou, Eleni Tzortzaki, Christos Kalloniatis, and Stefanos Gritzalis. Identifying privacy related requirements for the design of self-adaptive privacy protections schemes in social networks. *Future Internet*, 13(2):23, 2021.

[117] Rex B Kline. *Principles and practice of structural equation modeling.* Guilford publications, 2015.

[118] Bart Knijnenburg and Hongxia Jin. The persuasive effect of privacy recommendations for location sharing services. *Available at SSRN 2399725*, 2013.

[119] Bart Knijnenburg, Elaine Raybourn, David Cherry, Daricia Wilkinson, Saadhika Sivakumar, and Henry Sloan. Death to the privacy calculus? *Available at SSRN 2923806*, 2017.

[120] Bart P Knijnenburg. Simplifying privacy decisions: Towards interactive and adaptive solutions. In *Decisions@ RecSys*, pages 40–41. Citeseer, 2013.

[121] Bart P Knijnenburg. Privacy? i can't even! making a case for user-tailored privacy. *IEEE Security & Privacy*, 15(4):62–67, 2017.

[122] Bart P. Knijnenburg, Reza Ghaiumy Anaraky, Daricia Wilkinson, Moses Namara, Yangyang He, David Cherry, and Erin Ash. *User-Tailored Privacy*, pages 367–393. Springer International Publishing, Cham, 2022.

[123] Bart P Knijnenburg and Alfred Kobsa. Helping users with information disclosure decisions: potential for adaptation. In *Proceedings of the 2013 international conference on Intelligent user interfaces*, pages 407–416, 2013.

[124] Bart P Knijnenburg and Alfred Kobsa. Making decisions about privacy: information disclosure in context-aware recommender systems. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 3(3):1–23, 2013.

[125] Bart P Knijnenburg and Martijn C Willemsen. Evaluating recommender systems with user experiments. In *Recommender Systems Handbook*, pages 309–352. Springer, 2015.

[126] Bart P Knijnenburg, Martijn C Willemsen, Zeno Gantner, Hakan Soncu, and Chris Newell. Explaining the user experience of recommender systems. *User Modeling and User-Adapted Interaction*, 22(4):441–504, 2012.

[127] Bart Piet Knijnenburg. *A user-tailored approach to privacy decision support*. PhD thesis, UC Irvine, 2015.

[128] Bart Piet Knijnenburg and Alfred Kobsa. Increasing sharing tendency without reducing satisfaction: Finding the best privacy-settings user interface for social networks. In *ICIS*, 2014.

[129] A Kobsa. Tailoring privacy to users' needs (invited keynote) in bauer, m., gmytrasiewicz, pj and vassileva, j.(eds), proccedings of the user modeling 2001: 8th international conference, 2001.

[130] Hanna Krasnova, Sarah Spiekermann, Ksenia Koroleva, and Thomas Hildebrand. Online social networks: Why we disclose. *Journal of information technology*, 25(2):109–125, 2010.

[131] Ann Kronrod, Amir Grinstein, and Luc Wathieu. Go green! should environmental messages be so assertive? *Journal of Marketing*, 76(1):95–102, 2012.

[132] Franki YH Kung, Navio Kwok, and Douglas J Brown. Are attention check questions a threat to scale validity? *Applied Psychology*, 67(2):264–283, 2018.

[133] Cliff Lampe, Nicole B Ellison, and Charles Steinfield. Changes in use and perception of facebook. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, pages 721–730, 2008.

[134] Nancy K Lankton, D Harrison McKnight, Ryan T Wright, and Jason Bennett Thatcher. Research note—using expectation disconfirmation theory and polynomial modeling to understand trust in technology. *Information Systems Research*, 27(1):197–213, 2016.

[135] Robert S Laufer and Maxine Wolfe. Privacy as a concept and a social issue: A multidimensional developmental theory. *Journal of social Issues*, 33(3):22–42, 1977.

[136] George Lăzăroiu, Maria Kovacova, Jana Kliestikova, Pavol Kubala, Katarina Valaskova, and Victor V Dengov. Data governance and automated individual decision-making in the digital privacy general data protection regulation. *Administratie si Management Public*, (31):132–142, 2018.

[137] Jennifer S Lerner and Dacher Keltner. Fear, anger, and risk. *Journal of personality and social psychology*, 81(1):146, 2001.

[138] Steven Levy. *Facebook: The inside story*. Penguin UK, 2020.

[139] Roy J Lewicki and Chad Brinsfield. Framing trust: trust as a heuristic. *Framing matters: Perspectives on negotiation research and practice in communication*, pages 110–135, 2011.

[140] J David Lewis and Andrew Weigert. Trust as a social reality. *Social forces*, 63(4):967–985, 1985.

[141] Kevin Lewis, Jason Kaufman, and Nicholas Christakis. The taste for privacy: An analysis of college student privacy settings in an online social network. *Journal of computer-mediated communication*, 14(1):79–100, 2008.

[142] Han Li, Rathindra Sarathy, and Heng Xu. The role of affect and cognition on online consumers' decision to disclose personal information to unfamiliar online vendors. *Decision Support Systems*, 51(3):434–445, 2011.

[143] Han Li, Rathindra Sarathy, and Jie Zhang. The role of emotions in shaping consumers' privacy beliefs about unfamiliar online vendors. *Journal of Information Privacy and Security*, 4(3):36–62, 2008.

[144] Yifang Li, Nishant Vishwamitra, Hongxin Hu, and Kelly Caine. Towards a taxonomy of content sensitivity and sharing preferences for photos. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2020.

[145] Mengqi Liao and S Shyam Sundar. How should ai systems talk to users when collecting their personal information? effects of role framing and self-referencing on human-ai interaction. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2021.

[146] Christine Liebrecht, Christina Tsaousi, and Charlotte van Hooijdonk. Linguistic elements of conversational human voice in online brand communication: Manipulations and perceptions. *Journal of Business Research*, 132:124–135, 2021.

[147] Heather Richter Lipford, Gordon Hull, Celine Latulipe, Andrew Besmer, and Jason Watson. Visible flows: Contextual integrity and the design of privacy mechanisms on social network sites. In *2009 International Conference on Computational Science and Engineering*, volume 4, pages 985–989. IEEE, 2009.

[148] Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhimedi, Shikun Aerin Zhang, Norman Sadeh, Yuvraj Agarwal, and Alessandro Acquisti. Follow my recommendations: A personalized privacy assistant for mobile app permissions. In *Twelfth Symposium on Usable Privacy and Security ({SOUPS} 2016)*, pages 27–41, 2016.

[149] Bin Liu, Jialiu Lin, and Norman Sadeh. Reconciling mobile app privacy and usability on smartphones: Could user privacy profiles help? In *Proceedings of the 23rd international conference on World wide web*, pages 201–212, 2014.

[150] Yabing Liu, Krishna P Gummadi, Balachander Krishnamurthy, and Alan Mislove. Analyzing facebook privacy settings: user expectations vs. reality. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*, pages 61–70, 2011.

[151] Yabing Liu, Krishna P Gummadi, Balachander Krishnamurthy, and Alan Mislove. Analyzing facebook privacy settings: user expectations vs. reality. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*, pages 61–70. ACM, 2011.

[152] George Loewenstein and Jennifer S Lerner. *The role of affect in decision making.* Oxford University Press, 2003.

[153] Gale M Lucas, Nicole Krämer, Clara Peters, Lisa-Sophie Taesch, Johnathan Mell, and Jonathan Gratch. Effects of perceived agency and message tone in responding to a virtual personal trainer. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, pages 247–254, 2018.

[154] Mark MacCarthy and Kenneth Propp. Machines learn that brussels writes the rules: The eu's new ai regulation. *Brookings, May*, 4:2021, 2021.

[155] Ian Maddieson. Tone. *The world atlas of language structures online*, 2013.

[156] Michelle Madejski, Maritza Johnson, and Steven M Bellovin. A study of privacy settings errors in an online social network. In *2012 IEEE International Conference on Pervasive Computing and Communications Workshops*, pages 340–345. IEEE, 2012.

[157] Naresh K Malhotra, Sung S Kim, and James Agarwal. Internet users' information privacy concerns (iuipc): The construct, the scale, and a causal model. *Information systems research*, 15(4):336–355, 2004.

[158] Huina Mao, Xin Shuai, and Apu Kapadia. Loose tweets: an analysis of privacy leaks on twitter. In *Proceedings of the 10th annual ACM workshop on Privacy in the electronic society*, pages 1–12, 2011.

[159] Elwin Marg. Descartes'error: emotion, reason, and the human brain. *Optometry and Vision Science*, 72(11):847–848, 1995.

[160] Nikolas Martelaro, Victoria C Nneji, Wendy Ju, and Pamela Hinds. Tell me more designing hri to encourage more trust, disclosure, and companionship. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 181–188. IEEE, 2016.

[161] Aleecia M McDonald and Lorrie Faith Cranor. The cost of reading privacy policies. *Isjlp*, 4:543, 2008.

[162] AKM Nuhil Mehdy, Michael D Ekstrand, Bart P Knijnenburg, and Hoda Mehrpouyan. Privacy as a planned behavior: Effects of situational factors on privacy perceptions and plans. In *Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*, pages 169–178, 2021.

[163] Tamir Mendel, Roei Schuster, Eran Tromer, and Eran Toch. Toward proactive support for older adults: Predicting the right moment for providing mobile safety help. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 6(1):1–25, 2022.

[164] Meta. Facebook reports second quarter 2021 results. Technical report, Meta Investor Relations, 2021.

[165] Mainack Mondal, Günce Su Yilmaz, Noah Hirsch, Mohammad Taha Khan, Michael Tang, Christopher Tran, Chris Kanich, Blase Ur, and Elena Zheleva. Moving beyond set-it-and-forget-it privacy settings on social media. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, pages 991–1008, 2019.

[166] Scott Monteith and Tasha Glenn. Automated decision-making and big data: concerns for people with mental illness. *Current Psychiatry Reports*, 18(12):1–12, 2016.

[167] Frederick Muench, Katherine van Stolk-Cooke, Jon Morgenstern, Alexis N Kuerbis, and Kendra Markle. Understanding messaging preferences to inform development of mobile goal-directed behavioral interventions. *Journal of Medical Internet Research*, 16(2):e14, 2014.

[168] Moses Namara and Bart P. Knijnenburg. The differential effect of privacy-related trust on groupware application adoption and use during the covid-19 pandemic. *Proceedings of the ACM on Human-Computer Interaction, 5, CSCW2*, 2021.

[169] Moses Namara, Henry Sloak, and P. Bart Kninjenburg. The effectiveness of adaptation methods in improving user engagement and privacy protection on social network sites. *To Appear In The Proceedings on Privacy Enhancing Technologies*, 2022(1), 2022.

[170] Moses Namara, Henry Sloan, Priyanka Jaiswal, and Bart P Knijnenburg. The potential for user-tailored privacy on facebook. In *2018 IEEE Symposium on Privacy-Aware Computing (PAC)*, pages 31–42. IEEE, 2018.

[171] Moses Namara, Daricia Wilkinson, Kelly Caine, and Bart P Knijnenburg. Emotional and practical considerations towards the adoption and abandonment of vpns as a privacy-enhancing technology. *Proceedings on Privacy Enhancing Technologies*, 2020(1):83–102, 2020.

[172] Moses Namara, Daricia Wilkinson, Byron M. Lowens, Bart P. Knijnenburg, Rita Orji, and Remy L. Sekou. Cross-cultural perspectives on ehealth privacy in africa. In *Proceedings of the Second African Conference for Human Computer Interaction: Thriving Communities*, AfriCHI '18, pages 7:1–7:11, New York, NY, USA, 2018. ACM.

[173] Michel Netter, Moritz Riesner, Michael Weber, and Günther Pernul. Privacy settings in online social networks–preferences, perception, and reality. In *2013 46th Hawaii International Conference on System Sciences*, pages 3219–3228. IEEE, 2013.

[174] Oliver Niebuhr and Jan Michalsky. Computer-generated speaker charisma and its effects on human actions in a car-navigation system experiment-or how steve jobs' tone of voice can take you anywhere. In *International Conference on Computational Science and Its Applications*, pages 375–390. Springer, 2019.

[175] Paula M Niedenthal. Emotion concepts. *Handbook of emotions*, 3:587–600, 2008.

[176] Helen Nissenbaum. A contextual approach to privacy online. *Daedalus*, 140(4):32–48, 2011.

[177] Helen Nissenbaum. *Privacy in context.* Stanford University Press, 2020.

[178] Thomas P Novak and Donna L Hoffman. The fit of thinking style and situation: New measures of situation-specific experiential and rational cognition. *Journal of Consumer Research*, 36(1):56–72, 2009.

[179] Deirdre O'Brien and Ann M Torres. Social networking and online privacy: Facebook users' perceptions. *Irish Journal of Management*, 31(2):63, 2012.

[180] Xinru Page, Sara Berrios, Daricia Wilkinson, and Pamela J. Wisniewski. *Social Media and Privacy*, pages 113–147. Springer International Publishing, Cham, 2022.

[181] Xinru Page, Reza Ghaiumy Anaraky, Bart P Knijnenburg, and Pamela J Wisniewski. Pragmatic tool vs. relational hindrance: Exploring why some social media users avoid privacy features. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–23, 2019.

[182] Xinru Page, Pamela Wisniewski, Bart P Knijnenburg, and Moses Namara. Social media's have-nots: an era of social disenfranchisement. *Internet Research*, 2018.

[183] Raja Parasuraman, Thomas B Sheridan, and Christopher D Wickens. A model for types and levels of human interaction with automation. *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans*, 30(3):286–297, 2000.

[184] Oleksandra Pasternak, Cleopatra Veloutsou, and Anna Morgan-Thomas. Self-presentation, privacy and electronic word-of-mouth in social media. *Journal of Product & Brand Management*, 2017.

[185] Supavich Pengnate and Pavlo Antonenko. A multimethod evaluation of online trust and its interaction with metacognitive awareness: an emotional design perspective. *International Journal of Human-Computer Interaction*, 29(9):582–593, 2013.

[186] Andrew Perrin. Americans are changing their relationship with facebook. *Pew Research Center: Internet, Science & Tech*, September 2018.

[187] Richard E Petty, John T Cacioppo, and David Schumann. Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of consumer research*, 10(2):135–146, 1983.

[188] Jacqueline C Pike, Patrick J Bateman, and Brian S Butler. Information from social networking sites: C ontext collapse and ambiguity in the hiring process. *Information Systems Journal*, 28(4):729–758, 2018.

[189] Lee Rainie and Maeve Duggan Horowitz. Privacy and information sharing, Jan 2016.

[190] Ramprasad Ravichandran, Michael Benisch, Patrick Gage Kelley, and Norman M Sadeh. Capturing social networking privacy preferences. In *International symposium on privacy enhancing technologies symposium*, pages 1–18. Springer, 2009.

[191] Afsaneh Razi, Zainab Agha, Neeraj Chatlani, and Pamela Wisniewski. Privacy challenges for adolescents as a vulnerable population. In *Networked Privacy Workshop of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020.

[192] Elissa M Redmiles, Neha Chachra, and Brian Waismeyer. Examining the demand for spam: Who clicks? In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–10, 2018.

[193] Elissa M Redmiles, Sean Kross, and Michelle L Mazurek. How well do my results generalize? comparing security and privacy survey results from mturk, web, and telephone samples. In *2019 IEEE Symposium on Security and Privacy (SP)*, pages 1326–1343. IEEE, 2019.

[194] Leslie Riopel. What is the positive and negative affect schedule? (panas), 2021.

[195] Odnan Ref Sanchez, Ilaria Torre, Yangyang He, and Bart P. Knijnenburg. A recommendation approach for user privacy preferences in the fitness domain. *User Modeling and User-Adapted Interaction*, October 2019.

[196] Denise Sauerteig and Kling Sarah. How to make privacy settings easier to find using better names and organization.

[197] Florian Schaub, Rebecca Balebako, and Lorrie Faith Cranor. Designing effective privacy notices and controls. *IEEE Internet Computing*, 21(3):70–77, 2017.

[198] Florian Schaub, Bastian Konings, Michael Weber, and Frank Kargl. Towards context adaptive privacy decisions in ubiquitous computing. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2012 IEEE International Conference on*, pages 407–410. IEEE, 2012.

[199] Florian Marcus Schaub. *Dynamic privacy adaptation in ubiquitous computing*. PhD thesis, Universität Ulm, 2014.

[200] John T Scholz and Mark Lubell. Trust and taxpaying: Testing the heuristic approach to collective action. *American Journal of Political Science*, pages 398–417, 1998.

[201] Philip Seargeant and Caroline Tagg. Introduction: The language of social media. In *The language of social media*, pages 1–20. Springer, 2014.

[202] Thomas B Sheridan. Human centered automation: oxymoron or common sense? In *1995 IEEE International Conference on Systems, Man and Cybernetics. Intelligent Systems for the 21st Century*, volume 1, pages 823–828. IEEE, 1995.

[203] Thomas B Sheridan and William L Verplank. Human and computer control of undersea teleoperators. Technical report, Massachusetts Inst of Tech Cambridge Man-Machine Systems Lab, 1978.

[204] Frank M Shipman and Catherine C Marshall. Ownership, privacy, and control in the wake of cambridge analytica: The relationship between attitudes and awareness. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2020.

[205] Manya Sleeper, Rebecca Balebako, Sauvik Das, Amber Lynn McConahy, Jason Wiese, and Lorrie Faith Cranor. The post that wasn't: exploring self-censorship on facebook. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 793–802, 2013.

[206] Katherine Strater and Heather Richter Lipford. Strategies and struggles with privacy in an online social networking community. *People and Computers XXII Culture, Creativity, Interaction 22*, pages 111–119, 2008.

[207] Frederic Stutzman, Jessica Vitak, Nicole Ellison, Rebecca Gray, and Cliff Lampe. Privacy in interaction: Exploring disclosure and social capital in facebook. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 6, 2012.

[208] S Shyam Sundar, Jinyoung Kim, Mary Beth Rosson, and Maria D Molina. Online privacy heuristics that predict information disclosure. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2020.

[209] Vanitha Swaminathan, Elzbieta Lepkowska-White, and Bharat P Rao. Browsers or buyers in cyberspace? an investigation of factors influencing electronic exchange. *Journal of Computer-Mediated Communication*, 5(2):JCMC523, 1999.

[210] Auke Tellegen, David Watson, and Lee Anna Clark. On the dimensional and hierarchical structure of affect. *Psychological science*, 10(4):297–303, 1999.

[211] David Thissen, Lynne Steinberg, and Daniel Kuang. Quick and easy implementation of the benjamini-hochberg procedure for controlling the false positive rate in multiple comparisons. *Journal of educational and behavioral statistics*, 27(1):77–83, 2002.

[212] Kristin Thomas, Catharina Linderoth, Marcus Bendtsen, Preben Bendtsen, and Ulrika Müssener. Text message-based intervention targeting alcohol consumption among university students: findings from a formative development study. *JMIR mHealth and uHealth*, 4(4):e119, 2016.

[213] Eran Toch, Justin Cranshaw, Paul Hankes Drielsma, Janice Y Tsai, Patrick Gage Kelley, James Springfield, Lorrie Cranor, Jason Hong, and Norman Sadeh. Empirical models of privacy in location sharing. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, pages 129–138, 2010.

[214] Mina Tsay-Vogel, James Shanahan, and Nancy Signorielli. Social media cultivating perceptions of privacy: A 5-year analysis of privacy attitudes and self-disclosure behaviors among facebook users. *new media & society*, 20(1):141–161, 2018.

[215] Markus Tschersich and Reinhardt Botha. Exploring the impact of restrictive default privacy settings on the privacy calculus on social network sites. *In Proc. ECIS*, 2014.

[216] Hans Van Der Heijden. Ubiquitous computing, user control, and user performance: conceptual model and preliminary experimental design. *A Research Agenda for Emerging Electronic Markets*, page 107, 2003.

[217] Anthony Vance, David Eargle, Jeffrey L Jenkins, C Brock Kirwan, and Bonnie Brinton Anderson. The fog of warnings: how non-essential notifications blur with security warnings. In *Fifteenth Symposium on Usable Privacy and Security ({SOUPS} 2019)*, 2019.

[218] Marco Vanetti, Elisabetta Binaghi, Elena Ferrari, Barbara Carminati, and Moreno Carullo. A system to filter unwanted messages from osn user walls. *IEEE Transactions on Knowledge and data Engineering*, 25(2):285–297, 2011.

[219] Jonathan van't Riet, Robert AC Ruiter, Marieke Q Werrij, Math JJM Candel, and Hein De Vries. Distinct pathways to persuasion: The role of affect in message-framing effects. *European Journal of Social Psychology*, 40(7):1261–1276, 2010.

[220] BS Vidyalakshmi, Raymond K Wong, and Chi-Hung Chi. Privacy scoring of social network users as a service. In *2015 IEEE International Conference on Services Computing*, pages 218–225. IEEE, 2015.

[221] Sami Vihavainen, Airi Lampinen, Antti Oulasvirta, Suvi Silfverberg, and Asko Lehmuskallio. The clash between privacy and automation in social media. *IEEE pervasive computing*, 13(1):56–63, 2013.

[222] Arun Vishwanath, Weiai Xu, and Zed Ngoh. How people protect their privacy on facebook: A cost-benefit view. *Journal of the Association for Information Science and Technology*, 69(5):700–709, 2018.

[223] Sandra Wachter, Brent Mittelstadt, and Luciano Floridi. Why a right to explanation of automated decision-making does not exist in the general data protection regulation. *International Data Privacy Law*, 7(2):76–99, 2017.

[224] Qiaozhi Wang, Jaisneet Bhandal, Shu Huang, and Bo Luo. Content-based classification of sensitive tweets. *International Journal of Semantic Computing*, 11(04):541–562, 2017.

[225] Yang Wang, Pedro Giovanni Leon, Alessandro Acquisti, Lorrie Faith Cranor, Alain Forget, and Norman Sadeh. A field trial of privacy nudges for facebook. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 2367–2376, 2014.

[226] Yang Wang, Gregory Norcie, Saranga Komanduri, Alessandro Acquisti, Pedro Giovanni Leon, and Lorrie Faith Cranor. " i regretted the minute i pressed share" a qualitative study of regrets on facebook. In *Proceedings of the seventh symposium on usable privacy and security*, pages 1–16, 2011.

[227] Logan Warberg, Alessandro Acquisti, and Douglas Sicker. Can Privacy Nudges be Tailored to Individuals' Decision Making and Personality Traits? In *Proceedings of the 18th ACM Workshop on Privacy in the Electronic Society*, WPES'19, pages 175–197, New York, NY, USA, November 2019. Association for Computing Machinery.

[228] Logan Warberg, Alessandro Acquisti, and Douglas Sicker. Can privacy nudges be tailored to individuals' decision making and personality traits? In *Proceedings of the 18th ACM Workshop on Privacy in the Electronic Society*, pages 175–197, 2019.

[229] David Watson, Lee Anna Clark, and Auke Tellegen. Development and validation of brief measures of positive and negative affect: the panas scales. *Journal of personality and social psychology*, 54(6):1063, 1988.

[230] Jason Watson, Heather Richter Lipford, and Andrew Besmer. Mapping user preference to privacy default settings. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 22(6):1–20, 2015.

[231] Adam Waytz, Joy Heafner, and Nicholas Epley. The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, 52:113–117, 2014.

[232] Alan F Westin. Privacy and freedom. *Washington and Lee Law Review*, 25(1):166, 1968.

[233] Anita Whiting and David Williams. Why people use social media: a uses and gratifications approach. *Qualitative Market Research: An International Journal*, 2013.

[234] Primal Wijesekera, Arjun Baokar, Lynn Tsai, Joel Reardon, Serge Egelman, David Wagner, and Konstantin Beznosov. The feasibility of dynamically granted permissions: Aligning mobile privacy with user preferences. In *2017 IEEE Symposium on Security and Privacy (SP)*, pages 1077–1093. IEEE, 2017.

[235] Daricia Wilkinson, Öznur Alkan, Q Vera Liao, Massimiliano Mattetti, Inge Vejsbjerg, Bart P Knijnenburg, and Elizabeth Daly. Why or why not? the effect of justification styles on chatbot recommendations. *ACM Transactions on Information Systems (TOIS)*, 39(4):1–21, 2021.

[236] Daricia Wilkinson, Saadhika Sivakumar, David Cherry, Bart P Knijnenburg, Elaine M Raybourn, Pamela Wisniewski, and Henry Sloan. User-tailored privacy by design. In *Proceedings of the Usable Security Mini Conference*, 2017.

[237] Robert E Windom, James O Mason, and C Everett Koop. The surgeon general's 1989 report on reducing the health consequences of smoking: 25 years of progress: Executive summary. *Morbidity and Mortality Weekly Report*, 38(S2):i–32, 1989.

[238] Pamela Wisniewski. The privacy paradox of adolescent online safety: A matter of risk prevention or risk resilience? *IEEE Security & Privacy*, 16(2):86–90, 2018.

[239] Pamela Wisniewski, AKM Islam, Bart P Knijnenburg, and Sameer Patil. Give social network users the privacy they want. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pages 1427–1441. ACM, 2015.

[240] Pamela Wisniewski, AKM Islam, Heather Richter Lipford, and David C Wilson. Framing and measuring multi-dimensional interpersonal privacy preferences of social networking site users. *Communications of the Association for information systems*, 38(1):10, 2016.

[241] Pamela Wisniewski, Bart P Knijnenburg, and H Richter Lipford. Profiling facebook users privacy behaviors. In *SOUPS2014 Workshop on Privacy Personas and Segmentation*, 2014.

[242] Pamela J Wisniewski, Bart P Knijnenburg, and Heather Richter Lipford. Making privacy personal: Profiling social network users to inform privacy education and nudging. *International Journal of human-computer studies*, 98:95–108, 2017.

[243] Priscilla NY Wong, Duncan P Brumby, Harsha Vardhan Ramesh Babu, and Kota Kobayashi. Voices in self-driving cars should be assertive to more quickly grab a distracted driver's attention. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pages 165–176, 2019.

[244] Mu Yang, Yijun Yu, Arosha K Bandara, and Bashar Nuseibeh. Adaptive sharing for online social networks: A trade-off between privacy risk and social benefit. In *2014 IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications*, pages 45–52. IEEE, 2014.

[245] Haoti Zhong, Anna Squicciarini, and David Miller. Toward automated multiparty privacy conflict detection. In *Proceedings of the 27th ACM international conference on information and knowledge management*, pages 1811–1814, 2018.

[246] Mark Zuckerberg. Facebook, social media privacy, and the use and abuse of data. In *Joint hearing before the COMMITTEE ON COMMERCE, SCIENCE, AND TRANSPORTATION UNITED STATES SENATE and the COMMITTEE ON THE JUDICIARY UNITED STATES SENATE, One Hundred Fifteenth Congress, Second Session*, 2018.