Privacy Lies: Understanding How, When, and Why People Lie to Protect Their Privacy in Multiple Online Contexts

Shruti Sannon, Natalya N. Bazarova, Dan Cosley
Cornell University
Ithaca, NY, USA
{ss3464, nnb8, drc44}@cornell.edu

ABSTRACT

In this paper, we study online *privacy lies*: lies primarily aimed at protecting privacy. Going beyond privacy lenses that focus on privacy concerns or cost/benefit analyses, we explore how contextual factors, motivations, and individual-level characteristics affect lying behavior through a 356-person survey. We find that statistical models to predict privacy lies that include attitudes about lying, use of other privacy-protective behaviors (PPBs), and perceived control over information improve on models based solely on self-expressed privacy concerns. Based on a thematic analysis of open-ended responses, we find that the decision to tell privacy lies stems from a range of concerns, serves multiple privacy goals, and is influenced by the context of the interaction and attitudes about the morality and necessity of lying. Together, our results point to the need for conceptualizations of privacy lies—and PPBs more broadly—that account for multiple goals, perceived control over data, contextual factors, and attitudes about PPBs.

Author Keywords

Privacy; privacy-protective behaviors; deception; computer-mediated communication.

ACM Classification Keywords

K.4.1. Computers and Society: Public Policy Issues: Privacy.

INTRODUCTION

The proliferation of digital technologies has brought about an unprecedented amount of data collection and data sharing in people's everyday lives. People's concerns about sharing their personal information online have grown alongside these technological developments: in 2016, 74% of Americans believed that it is "very important" to be in control of personal information online, though only 9% of those surveyed believed they have such control [55]. In the absence of such control, people engage in a variety of privacy-protective behaviors (PPBs), from choosing to withhold personal information to complaining to regulators [63].

A less-studied PPB is to provide false information when faced with a concerning request: that is, to lie. This paper introduces the concept of *privacy lies* as the deliberate provision of false personal information to protect some facet of an individual's privacy. Such lies are fairly common: almost half of American teenagers have entered false information on their online profiles [37], while 40% of a sample of Internet users report lying to commercial websites [26].

Despite its prevalence, there is only limited understanding of this practice, including what motivates people to do it, how they engage in it, and to what effect. In particular, we see a need to consider factors beyond privacy concerns in order to understand privacy lies, as the well-known privacy paradox shows us that privacy concerns are often a poor indicator of people's actual behavior [6; 20; 47].

One such factor is context, which dynamically influences privacy goals, concerns, and behaviors [46; 51]. Situation-specific factors (such as an individual's current affect) can mediate the effect of pre-existing dispositional factors (such as general privacy concerns) on disclosure [32; 34], while making particular privacy-related needs and goals [1; 69] more or less salient. Another factor is that privacy lies are, after all, lies, and the literature on deception suggests that attitudes about lying are varied [43], and are likely to be important in understanding privacy lie decisions.

To get a better sense of how contexts, goals, and attitudes affect decisions to tell privacy lies, we surveyed 356 participants about their privacy-related concerns, attitudes, and behaviors, and asked them about online interactions in which they had told privacy lies. Following other privacy research that broadly characterizes context as either interpersonal or commercial, we prompted people to share both (a) *person lies*, when individuals interact directly with humans using computer-mediated communication (for example, text messages, online chat rooms, gaming chats, and dating sites) and b) *system lies*, when individuals interact with technologies themselves (for example, web forms, mobile applications, games, virtual agents, and chatbots) without an immediately salient human recipient. Participants who had not told privacy lies were asked for their reasons for not engaging in this PPB.

We found that attitudes about the morality, prevalence, and effectiveness of privacy lies, as well as perceptions of informational control and the use of other PPBs, were

significant predictors of lying frequency; in contrast, self-reported privacy concerns were insignificant when considering these other variables. While the model explained 35% of the variance, analyzing people's examples of privacy lies revealed that their decisions also varied based on their privacy goals, the nature of the request and the requestor, and whether they were interacting with a person or a system. Non-liars' reasons were more categorical: they either saw no need to lie, or saw privacy lies as morally or practically bad. Our findings suggest the need for models to incorporate context-specific factors that indirectly measure privacy concerns, as well as the need to understand how PPBs can have multiple underlying motivations and serve a range of functions beyond protecting personal data.

RELATED WORK

Below we lay the groundwork for these claims, starting with an overview of research on PPBs, focusing on privacy lies. We then discuss individual characteristics beyond privacy concerns that might affect decisions about privacy lies. Finally, we outline work at the intersection of privacy and motivations, and how these may vary by context.

An Overview of Online PPBs, Particularly Privacy Lies

Studies of online PPBs typically focus on one of two broad contexts: commercial interactions with systems or websites, and interpersonal interactions on social media. A taxonomy of PPBs in commercial settings suggests people can respond to privacy concerns in three main ways: a) misrepresenting, obfuscating, or withholding information; b) taking private action against the site by removing their data or negative word-of-mouth; or c) taking public action by complaining to the company or third-party regulators [63]. Multiple PPBs are also possible during interpersonal interactions on social media, including using ambiguous language [11], lying to interaction partners [70], changing privacy settings to manage audiences [5], and deleting content entirely [11].

As with PPBs in general, privacy lies are generally examined with respect to either commercial or interpersonal situations. Research on lying in commercial contexts indicates that as privacy concerns increase, people are more likely to lie [60], and they weigh costs and benefits when deciding whether to truthfully disclose personal information [27]. Based on interviews, Poddar et al. identified three additional factors that influence consumers' decision to falsify personal information in commercial Internet transactions: the necessity, invasiveness, and fairness of the request [54].

Privacy lying research also spans a number of interpersonal contexts, from dyadic exchanges with a chat partner (e.g., [71]) to broadcasting false information on social networking sites [21; 36]. As in commercial contexts, the likelihood of lying rises along with the level of privacy concerns [73]. Individual characteristics like gender can influence PPBs and the privacy concerns that drive them; for example, women are less likely to engage in PPBs despite having higher privacy concerns than men [58], and they engage in deception in online chatrooms mainly for safety reasons [71].

Individual Factors Beyond Privacy Concerns

While privacy concerns can predict participants' likelihood to falsify [60; 73] or withhold information [39], or to use privacy-enhancing tools, the relationship between privacy concerns and PPBs can be relatively weak [45]. More importantly, privacy concerns do not explain the whole story behind people's decision to tell privacy lies. Predictive models have not yet identified the sources of a large part of the variance for telling privacy lies. For instance, Jiang, Heng, and Choi's model to predict misrepresentation using privacy concerns and social rewards had limited predictive power (R^2=.10) [30], which suggests a need to explore other factors that affect privacy lies in interpersonal contexts. Son and Kim also tried to predict misrepresentation using privacy concerns in a commercial context; they found a non-significant relationship, and suggested that factors other than privacy concerns must underlie misrepresentation in commercial settings [63]. Below, we discuss potential factors that prior research has identified as both promising and potentially measurable for predicting privacy lies.

Informational control. A key aspect of privacy is the desire for informational control [19]. In 2015, 90% of American adults said it was important to be able to control what personal information was collected about them, and 93% said it was important to control who had access to it [40]. The ability to use privacy-enhancing technologies that allow people to control the flow of their information is positively associated with perceived control of data, which in turn mitigates privacy concerns [72]. This, however, can then lead to more self-disclosure [6] and risky behaviors [56].

Privacy literacy and experiences. Privacy knowledge and experience might also be useful for predicting privacy lies. Privacy literacy scales that rely on the number of actual PPBs people report performing in daily life, such as Turow and Hennessy's [65], serve both as proxies for knowledge and evidence of ability. Further, peoples' past experiences with privacy violations may impact their privacy concerns [14] and predict PPBs [8]. For example, Facebook users who have personally experienced an invasion of privacy are more likely to engage in PPBs by changing their privacy settings [15]. This suggests the importance of examining privacy-related experiences such as hearing privacy-related news or being a victim of a privacy violation [62].

General trust and willingness to disclose information. Personal privacy experiences might affect general trust of the Internet, i.e., the belief that institutional actors will protect and not disclose consumer data [65]. Similarly, an individual's general tendency to disclose personal information can predict PPBs [23] and helps understand disclosure decisions [38], including truthfulness.

Attitudes about deception and lying online. Finally, because privacy lies are, well, lies, decisions about them are likely influenced by people's general attitudes about lying, such as its effectiveness and appropriateness for protecting privacy. Some people may see misrepresentation for the sake of

privacy as deceitful and be less likely to engage in it [9]. Similarly, protection motivation theory [41] suggests people are more likely to engage in protective behaviors they view as being effective, and perceptions of social norms can influence the frequency of many behaviors [4] including self-disclosure [10]. Finally, research on location-sharing disclosures on social media suggests that the propensity to lie is a stable characteristic that can have a positive direct effect on online privacy concerns [49].

Taken together, this work suggests many individual characteristics beyond privacy concerns that are likely to affect decisions around privacy lies. Our first research question, then, is about whether these factors improve our ability to predict privacy lying behavior:

RQ1: To what extent can individual-level characteristics beyond privacy concerns, including privacy-related views, actual privacy-related experiences, and attitudes about lying as a PPB, help predict privacy lies?

Privacy Goals and Functions, in Context
Most research on PPBs has studied privacy-protective behaviors through the lens of the privacy calculus, based on a weighing of the costs and benefits [16; 34]. For example, people weigh the perceived enjoyment they could derive from revealing information against the likelihood of a privacy violation when making disclosure decisions [35].

However, the considerations and weightings individuals bring to the privacy calculus are only partially determined by persistent, underlying traits such as general privacy attitudes [50]. The "privacy in context" perspective argues that privacy decisions and PPBs reflect contextual interpretations and perceptions that interact with traits such as general privacy attitudes [46; 50; 52]. For example, people make in-the-moment, context-driven decisions about whether to lie about their location on social media based on factors such as their current physical distance from another user [22].

We argue that, in order to understand the nature of privacy lies, we have to look into the motivations that drive them in a more nuanced way than simply the general goal of protecting privacy or data. According to Westin's taxonomy of privacy functions [69], there are at least four main privacy-related motivations: autonomy, the absence of manipulation by others; self-evaluation, a safe space to reflect on feelings and identity free from the threat of social judgment; emotional relief, the ability to deviate from social norms and expectations; and limited and protected communication, the ability to dictate intimacy and boundaries with interaction partners. These different conceptualizations and functions of privacy are lost if online PPBs are considered through the lens of protecting informational privacy alone.

As with perceived costs and benefits, motivations also likely vary by context. Recent research on context and motivations in disclosure suggests that contextual factors such as audience representations in social media platforms amplify different types of disclosure motivations, which, in turn, help

explain online disclosure decisions [3; 12]. Although this line of research has focused on disclosure attributes other than truthfulness, we see its insights as likely to apply to decisions around privacy lies as well.

Further complicating matters, relative to the common decision in privacy work to focus on *either* systems or people, is that many online contexts raise interpersonal and system-oriented considerations simultaneously. Decisions to provide truthful information in a Facebook profile, for instance, affect both what Facebook knows about you and how others perceive you. This doesn't mean that both are always salient. For example, Young et al. found that on balance, college students were concerned about protecting their social privacy on Facebook but not their institutional privacy, i.e., data collected by the platform [74]; Felt, Egelman and Wagner found similar sentiments for privacy concerns in mobile applications [17]. Specific design decisions in a system's interface can have major impacts in how people perceive context. For instance, design features such as perceived anonymity and availability of social cues can affect people's risk/benefit calculus, and consequently, their privacy concerns [30].

In addition to the macro context, research suggests that the specific interaction partner matters [48]. In general, people are more likely to provide information to partners they trust [32] and are familiar with [74]. Characteristics of the requestor also matter in commercial contexts, where decisions to disclose information to websites are based on factors such as regard for the company and trust in the website [44].

The nature of the specific request is also important. People are more likely to disclose information to websites when they see the request as relevant to the situation [75], fair, and not invasive [54]. Similarly, people draw on a range of privacy rules when making decisions about disclosure in interpersonal exchanges, where the nature of the request, including the type of question asked, its depth, and its appropriateness, matter [52].

Overall, this literature positions context itself as a malleable rather than a fixed property, to which different people may assign different meanings, amplifying certain goals, concerns, and elements while disregarding others. This complicates the relationship between motivations, disclosure truthfulness, and context, making it important to look at how motivations manifest in and depend on specific contexts. Together, these findings give rise to the following two research questions:

RQ2: What are people's motivations for telling privacy lies to people and systems? In what situations do they arise and what specific privacy functions do they serve?

RQ3: How do contextual factors, such as people's perceptions of both information requesters and requests, influence people's decisions to tell privacy lies?

How Can Non-Liars Teach Us About Privacy Lies?
The prior research questions focus on decisions about lying and how context may affect these decisions, but some people choose not to tell privacy lies. We believe studying these *non-users* [2] of privacy lies as a PPB could also be revealing. Based on prior work, we foresee three individual-level characteristics that may drive people to not tell privacy lies. First, some people may simply be less worried about their online privacy than others [59] and thus may not engage in PPBs, including privacy lies. Second, people may develop a sense of learned helplessness when faced with repeated privacy violations online [61], making them less likely to attempt to protect their privacy by lying. Third, people may have moral reasons for not engaging in privacy lies. Some users who tested a differential privacy solution felt that obscuring or fuzzing their information was unethical [9], and moral beliefs are inversely related to misrepresenting oneself in interpersonal interactions online [30]. However, although lying is often considered immoral, identity concealment to support privacy is viewed as a relatively acceptable form of deception [67]. Thus, privacy lies may also be seen as acceptable. These observations lead to our final research question:

RQ4: What are the reasons people do not tell privacy lies?

METHOD
To address these questions, we collected both quantitative and qualitative data about people's privacy lying behavior and attitudes, their concerns around and willingness to share personal data, and their views and experiences around privacy in different online contexts.

Procedure
We collected the data through a survey of Amazon Mechanical Turk (MTurk) workers that invited them to participate in a study on online behaviors and attitudes. After giving consent, participants were first asked for demographic information in case later exposure to information on privacy lies would influence them to lie about such data. Participants were then asked about their general willingness to disclose various kinds of data online and their privacy-related experiences and behaviors.

We then presented participants with a short tutorial on privacy lies, adapted from another study on deception [24], to both walk participants through the definition of privacy lies and to reduce the stigma of admitting to socially unacceptable behavior. We reassured participants that we did not view lies as good or bad, but were simply interested in examining an important part of human communication. We kept the definition of privacy lies broad ("providing false information about yourself online to protect your privacy") so as to not limit notions of privacy, privacy goals, and contexts in which participants had privacy concerns. To progress to the remainder of the survey, participants had to correctly answer a 2-item quiz to distinguish privacy lies from other forms of deception; they received feedback on incorrect answers, which they could then correct.

After passing the quiz, participants were asked whether they had ever told a privacy lie online. Those who had told privacy lies were then asked whether they had told these lies to people in interpersonal contexts or directly to websites or other devices where another human was not directly involved, then to describe a detailed example of when and why they had told a privacy lie to a person and/or a system. Participants who reported not having told privacy lies were asked to explain their reasoning in an open-ended response. All participants then answered questions about their privacy concerns and attitudes, as well as their perceptions of the prevalence, morality, and effectiveness of privacy lies based on hypothetical scenarios describing privacy lies to people and systems. The scenarios and questions are available as supplemental material in the ACM digital library.

We first piloted the survey with a student sample (*N*=10) to refine questions. After refinement, the final survey was posted on MTurk at various times during the morning, afternoon, and night on multiple days to gain variation in the sample. We restricted the survey to MTurkers located in the United States. Since the survey took the first 20 MTurk participants an average of 12 minutes to complete, we compensated participants $1.50 to comply with U.S. federal minimum wage standards.

Participants
Of 409 respondents, we removed 34 who failed the attention check at the end of the survey and 19 whose responses did not fit the definition of privacy lies. Of the remaining 356 participants, 54% were male and 82% identified as Caucasian. The average age was 34 years. On average, participants estimated that they used the Internet for 7.6 hours a day. Most reported having some college education or higher (87%) and being employed full-time (68%), as well as having worked on MTurk for over 1 year (66%). Only 23% of participants reported never having told a privacy lie, and 18% told them rarely, whereas 41% told them sometimes, and 18% told them often or all the time.

Measures
Frequency of Privacy Lies
Frequency of telling privacy lies was computed from two items: first, participants were asked if they ever told privacy lies (yes/no); participants who selected "yes" were asked to rate the frequency with which they tell privacy lies. These two items were combined into a 5-point scale (never, rarely, sometimes, often, all the time).

Privacy-Related Views
To measure privacy-related views, we used several scales: a 2-item *Trust of the Internet* scale, *α*=.77 [65]; a 5-item *Perceived Control of Data* scale, *α*=.86 [72]; and a 10-item *Users' Information Privacy Concerns* scale, *α*=.79 ([42], changing "online company" to "online party"). We also asked about participants' *Willingness to Disclose Online* for each of 9 personal data points (name, age, gender, sexual orientation, birth date, email address, location, browsing history, and content of online messages), adapting this

		Individual	Reduced Model		Full Model	
	Variables	**Adj. R^2**	**Coeff.**	**S.E.**	**Coeff.**	**S.E.**
Control variables	Age	.002	-.002	.005	-.002	.005
	Gender (Female as reference category)	<.001	.063	.097	.042	.098
Privacy-related views	Trust of the Internet	.036**	-	-	-.049	.036
	Perceived Control of Data	.040**	-.162*	.061	-.12	.067
	Users' Informational Privacy Concerns	.142**	-	-	.004	.053
	Willingness to Disclose Online	.075**	-.219*	.069	-.188*	.070
Actual privacy-related experiences and behaviors	Number of PPBs	.096**	.139**	.035	.126**	.037
	Freq. of being a Privacy Victim	.003	-	-	.043	.042
	Freq. of Privacy-related News	.049**	-	-	.019	.044
	Freq. of Privacy-concerning Situations	.039**	-	-	.031	.049
Attitudes	Attitudes about Privacy Lies	.275**	.463**	.046	.460**	.053

Table 1. Values for the full model (R^2= .37; adjusted R^2 = .35), the reduced model (R^2 = .36; adjusted R^2 = .35), and individual predictors (controlling for age and gender) predicting frequency of telling privacy lies (N=356).
Significance is indicated by * for $p < .05$ and ** for $p < .001$.

measure from previous privacy studies [23; 53]. We summed the number of data points they were willing to disclose, ranging from 0 to 9.

Actual Privacy-Related Behaviors and Experiences
We adapted binary items used in prior work to gauge participants' use of PPBs [65], including whether they used "do not track" plugins, virtual private networks (VPNs), invisible browsing, ad blockers, and secure passwords, as well as whether they regularly cleared browser cookies, controlled privacy settings on social network sites, and read app and website privacy policies. We summed the *number of PPBs* participants engaged in, ranging from 0 to 8.

We measured privacy experiences with three 7-point bipolar items: how often they had personally been the victim of a privacy violation in the past (*Freq. of Privacy Victim)*, how frequently they had seen or heard about online privacy violations in the past year (*Freq. of Privacy-related News)* used in Smith et al. [62], and an original item measuring their perceived frequency of coming across privacy-concerning situations online, with anchors of 1="Never" and 7="All the time" (*Freq. of Privacy-concerning situations)*.

Attitudes about Privacy Lies
Based on prior work that uses hypothetical scenarios to measure attitudes about lying [43], we presented participants with two short scenarios of lies to people and to systems online. They rated each scenario on three original 7-point bipolar items about the degree to which they perceived privacy lies to be common, ethical, and effective in protecting privacy. Since principal components analysis extracted one factor for the six items (measuring lies to both people and systems), we computed a variable that represents people's overall *attitudes about privacy lies*, α=.82.

RESULTS
Our results are divided into three sections. First, we put forth a model that accounts for the individual characteristics that predict lying as a PPB, with the goal of understanding factors beyond privacy concerns that account for this behavior (RQ1). Then, we explain the results of our thematic analysis of privacy lie examples that shed light on various privacy goals, functions, and contextual factors underlying the decision to tell privacy lies in person and system contexts (RQ2 and RQ3). Finally, we present the results of our thematic analysis of the reasons why people do not tell privacy lies to understand these non-use behaviors and to reveal hidden factors that might be at play (RQ4).

RQ1: Predicting Privacy Lies
Correlations between all variables ranged from <.001 to .35, except for between privacy concerns and attitudes about privacy lies (r=0.54). We ran a test for multicollinearity using the mctest package in R, and the highest variation inflation factor (VIF) was 1.79 (for privacy concerns). Based on the relatively low VIF values, we retained all variables in the analyses without concern for multicollinearity.

We then ran a general linear model that included 11 variables (9 predictor and 2 control variables) with frequency of lying as a continuous dependent variable. Using a stepwise regression procedure, variables were deleted one at a time based on largest p-value. Likelihood ratio tests were then run to compare the reduced model with the full model until removing a variable led to a significant difference between models. The beta coefficients, standard errors, and p-values of the reduced and full models are reported in Table 1. We also computed models for each individual predictor, controlling for age and gender; adjusted r-squared values for those models are also shown in Table 1.

Controlling for age and gender, frequency of privacy lies in the reduced model was predicted by a) perceived control of data, b) general willingness to disclose online, c) the number of other PPBs people engage in online, and d) attitudes about privacy lies. Perceived control of data and general willingness to disclose online were negatively associated with privacy lies; the number of PPBs and attitudes about privacy lies were positively related to the frequency of privacy lies. Together, these four factors explained 35% of variance, whereas a model with privacy concerns only explained 14% of the variance. Privacy concerns was not a significant predictor in the full or reduced models that included the above-mentioned factors.

Frequency of being a victim of a privacy violation did not predict the likelihood of telling privacy lies. The frequencies of reading about privacy violations in the news and coming across privacy-concerning situations online were significantly related to frequency of privacy lies taken individually, but were no longer significant in either the full or reduced models.

Finally, attitudes about privacy lies (in terms of their prevalence, morality, and effectiveness) were the strongest predictor of lie frequency. Overall, the reduced model provides a parsimonious explanation of individual-level characteristics—perceived control of data, general willingness to disclose online, use of other PPBs, and attitudes about privacy lies—that predict frequency of engaging in privacy lies.

RQ2 and RQ3: Context and Motivations for Lying

The model explains about one third of the variance in the frequency of telling privacy lies, suggesting that there are other factors to uncover. Thus, we turn to an investigation of contextual factors that may influence privacy lies: the reasons people have for telling privacy lies, and how these are influenced by their perceptions of the information request and the requestor in both person and system contexts.

To do this, we looked at the open-ended responses from 275 participants who reported telling privacy lies. Overall, 13% of these participants had told such lies only to people, 38% had told only told them to systems, and 51% had told both system and people lies. We report on the themes that emerged from analyzing 175 person lies and 236 system lies.

Two members of the research team open-coded the first 75 examples of each set, at which point we reached code saturation and then iterated. We found that participants often provided multiple reasons for telling any given privacy lie, so we allowed for multiple overlapping codes. Ultimately, we developed a framework with four main themes comprised of 11 subcodes, shown with subcode frequencies in Table 2. Those themes and subcodes were used to code the remainder of the dataset of lies to people (*Cohen's kappa*=.70) and lies to systems (*Cohen's kappa*=.68). The first two themes involve people's perceptions of the information request and the requesting party, while the second two themes represent

Themes	#	Subcodes	Person	System
Judgments about the request for information	1	Unneeded	16 (8.7%)	52 (21.1%)
	2	Inappropriate/ Discomfiting	50 (27.3%)	36 (14.6%)
Perceptions of the Requestor	3	Unfamiliarity	45 (24.6%)	8 (3.2%)
	4	Distrust	15 (8.2%)	28 (11.3%)
Lies for Gain	5	Gaining a benefit	2 (1.1%)	152 (61.5%)
	6	Exploring a persona	28 (15.3%)	26 (10.5%)
	7	Privacy for privacy's sake	15 (8.2%)	20 (8.1%)
Lies of Avoidance	8	Avoiding harm in general	63 (34.4%)	54 (21.9%)
	9	Avoiding offline repercussion	65 (35.5%)	22 (8.9%)
	10	Avoiding communication	30 (16.4%)	65 (26.3%)
	11	Avoiding misuse of data	0	54 (21.9%)

Table 2. Breakdown of main themes in examples of privacy lies in interactions with people (*n*=175) and systems (*n*=236). Note that lies could exhibit multiple themes; thus, percentages total over 100%.

people's motivations to gain or avoid a particular outcome. Below, we discuss each theme in turn.

1. *Judgments about the request for information*: Many participants expressed strong opinions about the request for information. A common reason to tell a privacy lie was the perception that the data requested was unneeded (*Subcode 1: Unneeded*), more so in interactions with systems (21.1%) than with people (8.7%). Participants voiced frustration that requests from systems were often "unwarranted and not necessary to know" (P23) for the purposes of the interaction, such as a birthdate for "a simple registration" (P51). Research indicates that consumers are more likely to disclose their information to companies if they feel they are being treated fairly [63]; our findings suggest that a perception of unfairness can also motivate people to tell privacy lies. Such perceptions of unfairness also sometimes elicited emotional reactions in comments such as "they don't deserve to know" (P34), and "it's none of their business" (P73).

Participants often reported that the requests for information were not just unneeded, but also uncomfortable (*Subcode 2: Inappropriate/Discomfiting*), here more so when lying to people (27.3%) than systems (14.6%). Comments from these participants suggest they wanted control over who got to see their personal information; privacy lies were a way that they could avoid "very, very personal questions" (P51) they saw as inappropriate, particularly during interactions with relative strangers who made them feel uncomfortable. People also told privacy lies in response to "invasive" (P30) requests from systems, such as P170: "I use a false social media persona I've set up, to avoid giving systems access to my

genuine accounts. I've dedicated Twitter and Facebook accounts to deal with this nonsense." They also wanted to provide information of their own volition rather than be inundated by requests, as voiced by P82, "I enjoy giving the info I want myself without being interrogated."

2. *Perceptions of the Requestor*: 24.6% of participants who told privacy lies to another person did so because they were interacting with someone they were relatively unfamiliar with on platforms such as Reddit, Facebook, dating websites, and online games (*Subcode 3: Unfamiliarity*). In some cases, they did not know or want to know the person at all; in others, did want to interact, but their partner overstepped personal boundaries:

"I was recently on a forum and had a man start talking to me. He seemed nice and was interested in the same thing I was. We were just having friendly conversations, then he started asking more personal information. I did not feel comfortable giving him that so I lied about my name but I did tell him my age" (P70).

These findings are in line with deception research that people are more likely to tell self-serving lies to people who are not well-known to them, since it would be harder to get away with telling such lies to people with whom they are familiar [70].

In contrast, unfamiliarity was rarely called out in privacy lies to systems (3.2%); instead, participants mentioned telling privacy lies to websites they used frequently. For instance, some Facebook users were concerned about both the other people who could view their data (P69) and the company itself: "Most of my Facebook information is filled with privacy lies, I think Zuckerberg demands/resells too much information" (P60). In both person (8.2%) and system (11.3%) contexts, perceived untrustworthiness could be the immediate trigger for telling a privacy lie (*Subcode 4: Distrust*). Given that trust in the requestor can be a powerful mediator between privacy concerns and eventual disclosure [31], we would have expected trust to occur more frequently in participants' responses. This is, however, consistent with the statistical model, in which general trust of the Internet was found not to be significant.

3. *Lies for Gain*: A factor that was much more strongly tied to privacy lies in interactions with systems (61.5%) than people (1.1%) was the motive to gain some form of benefit (*Subcode 5: Gaining a Benefit*). These privacy lies mainly involved gaining access to services or information that required personal data during the signup process. Related to the "Unneeded" theme, participants would sometimes assess which information they thought was truly needed to accomplish their gains, being truthful about those elements while lying about parts that they deemed unnecessary:

"Recently a website who doesn't need my birth date, requested it. I give the correct year, but lied about the month and day in order to protect my privacy. This information shouldn't be asked for a simple registration. I agree if they want to see if I'm over 18 (21) they could ask for my birth year, but not month and day" (P51).

Some participants told privacy lies to anonymously try out an alternate identity (*Subcode 6: Exploring a Persona*). This was relatively common in interactions with both people (15.3%) and systems (10.5%). For example, P20 lied about his age and gender, saying "I wanted to change people's perspective of me and to make sure I couldn't be identified." Others told privacy lies to be able to engage in stigmatized activities; for example, P64 misrepresented himself on a marijuana forum and P134 lied on a sex chat site to avoid being identified and embarrassed, while P78 provided false information to a pornographic website, stating "Given that there are but a couple of people who I can talk freely about topics such as this, and that it is a taboo subject in our culture, I'd rather not disclose [my] information." People told similar lies during direct interactions with people, for example, when expressing an unpopular political opinion (P78). In Westin's taxonomy of privacy functions [69], such lies would serve the purpose of emotional release by allowing people a private way to deviate from social norms and rules.

The final subcategory in the "lie for gain" category was what we conceptualized as "privacy for privacy's sake", which was cited just over 8% of the time for both person and system lies (*Subcode 7: For Privacy's Sake*). Comments from these participants included statements about their identities or emotions, such as "I'm a bit secretive person" (P177) and "I was feeling very private" (P16). These participants seemed to value their privacy regardless of whether they perceive an additional benefit or negative consequence from revealing their information. For example, P84 stated: "I was uncomfortable using my real information, even though I felt reasonably confident that the person online would not personally do any damage with the information."

4. *Lies of Avoidance*: Many participants were concerned with avoiding undesired consequences in all interactions, which emerged as another major category for engaging in privacy lies with both people (34.4%) and systems (21.9%) (*Subcode 8: Avoiding General Harm*). Such concerns might entail avoiding "unwanted attention" (P66), fraud or identity theft (P30), or harassment (P163). During interactions with people, participants were especially worried about offline repercussions such as being stalked, and consequently often lied about their location (35.5%), while offline repercussions of systems collecting data were much less salient (8.9%) (*Subcode 9: Avoiding Offline Repercussions*).

People also lied to avoid further communication with people (16.4%) and systems (26.3%) (*Subcode 10: Avoiding Communication*). With people, privacy lies served to stop unwanted communication altogether, or (as with "Inappropriate" requests) to avoid a line of questioning that was perceived to be invasive while remaining in contact (P162). In the latter case, lies were a useful way to assert personal boundaries. With systems, avoiding communication mainly involved telling privacy lies to avoid spam, often by

giving fake email addresses, "so I can still investigate the offers but am not bothered by solicitations that typically follow" (P273). Finally, 21.9% of lies to systems were triggered by concerns about data management and reuse (*Subcode 11: Avoiding Misuse of Data*). These participants were often less concerned about the initial request than what may happen to their data later: "If for some reason the information is compromised I do not want it to be extremely easy for someone to find me and figure out who I am" (P196). In contrast, people did not mention concerns about third party sharing in their responses about privacy lies to people at all.

RQ4: Reasons to Not Lie

We also sought to understand the reasons for not engaging in privacy lies through a thematic content analysis of responses from 91 participants who reported never telling privacy lies. We developed a framework of reasons based on open coding the first 25 reasons to not tell privacy lies, at which point we reached code saturation. We found that most participants provided exactly one reason for not telling privacy lies, so we used a mutually exclusive coding scheme for these responses. Two coders coded the responses according to our framework (*Cohen's kappa*=.69); all disagreements were resolved through discussion. Our findings can be organized into four main categories (Table 3), two of which cast lying as unnecessary, and two of which cast lying as bad.

The most common reason people gave for privacy lies being unnecessary was that they did not feel a need to protect their privacy online through *any* PPB. These participants had "very little to hide" (P61, P145) or were "like an open book" (P229). One participant drew an analogy to people who do not have the option to remain private in public spheres, stating "Famous people have to let all their information hang out there in the public, essentially that is, most of their information is known, so if they can do it I can do it" (P160). Some of them believed that no damage could come from making their personal information available online, stating "I have nothing of value in my name that someone could benefit from if they stole my identity" (P184), or "I don't feel a threat that is tangible from exposing this information" (P45). Other participants rarely came across what they believed to be unwarranted requests for personal information: "Most of the times I needed to use my personal information it has been for very necessary reasons" (P80). It's possible that some of these participants would be more willing to lie for privacy reasons if they encountered services that asked for unneeded information, though their threshold for determining whether an information request is warranted may be different compared to those who do engage in privacy lies.

Another reason privacy lies could be unneeded is reliance on other PPBs, as reported by about one-fifth of non-liars. Instead, omission was a common strategy in the face of privacy threats, where participants would "simply refrain from giving the information as opposed to giving false information" (P126). In some cases, these participants may

Reasons for Not Telling Privacy Lies	Frequency
Unneeded	41 (45%)
Preference for another PPB	19 (21%)
Immoral/Unethical	20 (22%)
Could have negative repercussions	11 (12%)

Table 3. Reasons given by non-liars (*n*=91) for not telling privacy lies (percentages total to 100%)

compromise, providing some information but omitting the rest: "I sometimes omit things such as my exact birthday, such as only putting month and year or just month or just year" (P357). However, sometimes certain data points cannot be omitted because they are required to access online services. In such cases, participants either choose not to sign up for the service (P357), to avoid the website altogether (P176), or only use websites that they trust (P86). Similarly, when communicating with an individual, participants like P342 "would probably just stop talking to them or not answer, or find a way to circumvent the question." Other PPBs reported by participants included secure browsing (P132) or privacy settings to ensure their information was restricted to only a few people (P153).

The other main general theme reflected the negative view on lying in general. For about one-fifths of our participants, decisions about privacy lies were first and foremost a moral issue. These participants saw it as "taking on a fake identity" (P4) and viewed themselves as truthful people, stating that they "prefer to be honest" (P79). They also disliked lying in general regardless of motivation, viewing it as "morally wrong" (P312) or citing religious reasons, such as P344: "The Bible says not to lie so I try not to. Whether we justify them with one thing or another, it's still wrong." P270 was specifically concerned about the personal impact of privacy lies: "If I lied about myself online I would feel a sense of shame that I don't think I would be able to shake." In fact, some participants who *did* lie often offered partial truths—for example, saying they lived in a neighboring town—perhaps to balance privacy concerns with moral reservations.

Finally, a smaller number of participants saw lying not as intrinsically bad, but as a bad idea because privacy lies could have negative consequences: "Telling lies online can lead to a cascade of problems if someone decides to try and dig deeper into my personal history" (P154), "especially if you end up becoming friends with the person over time" (P289). This reasoning has some parallels to deception in online dating, where people will only tell small lies to help manage social constraints so as not to adversely affect potential future interactions [64]. Participants also worried about "getting in trouble" or "banned" (P48) for providing false information in various online contexts, particularly if the "falsities didn't line up properly" (P330). Some people saw maintaining consistency with and remembering privacy lies as too much effort: "If I provide lies online, I would have to provide a different answer to each situation. I would then have to remember those details if I want to use these websites in the future. That requires a lot of effort. I decided that the

benefit of doing this is not worth the cost" (P115), although several people who did tell privacy lies managed this by using "fake identities" where they provided false information consistently across websites.

DISCUSSION

In the following sections, we discuss the implications of our results for privacy research and design.

Implications for Privacy-Related Measures

Our findings support the need to go beyond self-reported privacy concerns and to measure privacy without asking about it directly [7]. The predictive model captured just over one-third of the variance in telling privacy lies; however, self-reported privacy concerns were not significant in this model. These findings align with research that people who report high privacy concerns may not actually act in a privacy-protective manner [8; 47]. Non-liars' responses were interesting in this regard: in the thematic content analysis, almost half of their responses were about lying, or PPBs in general, being unnecessary—but their quantitative responses about privacy concerns were still above 5 on a 7-point scale. This suggests that explicit questions about privacy concerns might pose a social desirability, experiment demand characteristic, or priming bias, and a corresponding need for less direct measurements of privacy concerns. For instance, the actual use of non-lying PPBs was a significant predictor for telling lies, probably because this identifies people who were more willing and able to translate concerns into concrete PPBs. Questions about PPBs might also have a smaller effect on priming people to be concerned about their privacy versus explicit questions about privacy concerns.

Our study also indicates a need to understand people's attitudes about specific PPBs when studying how their privacy concerns map onto their protective responses. Although only about 20% of non-liars explicitly cited moral concerns as their main reason for not lying, the predictive model shows that people's attitudes about privacy lies explained more variance than any other variable alone. Based on our qualitative analysis, we suspect that attitudes about (particular forms of) deception are important to consider in studying privacy lies and PPBs, as with Bullek's finding that some people regard differential privacy solutions that obscure data as immoral [9]. Designs could target these attitudes about deception, framing PPBs as more or less deceptive or socially normative in an effort to encourage or discourage particular PPBs depending on the design goals at hand.

The predictive model also showed that higher perceived control of data was associated with lower frequency of privacy lies. Research on the privacy control paradox has shown that perceptions of informational control can make people more likely to disclose sensitive information [6] and more likely to both engage in risky behaviors and ignore risk-related information [56]. Our results suggest that the privacy control paradox extends to privacy lies as well; there may be value in designs that help people reflect on how much control

they have over their data. Consider, for instance, a privacy nudge [68] that shows not just friends but also unknown people who might see a Facebook post based on its privacy setting.

From Data Control to Privacy Goals

The prior finding highlights a danger for individuals focused on control over their data. We suggest that common research framings of privacy as control over data and preventing information loss may also be harmful, because those framings can lose sight of the range of goals served by privacy (c.f., [69]).

For example, participants reported telling privacy lies to be able to explore an identity within a safe space (such as sexual orientation), in line with Westin's concept of self-evaluation, while others told lies to vent or to explore socially unacceptable opinions without social censure, in line with Westin's conceptualization of privacy as an emotional release. The goals of these privacy lies fit well with arguments that individuals need a backstage where they can safely explore ideas [52]. People also told privacy lies to limit their communication to specific audiences, in line with Westin's privacy function of protected communication, and to assert their freedom from external manipulation, in line with Westin's privacy function of autonomy. Although all of these motivations can be expressed in terms of control over information, we see such control more as a *means*, rather than an *end*—and considering how people think about these ends is likely to be fruitful for better understandings of privacy and privacy lies.

Consider the case of unneeded/inappropriate requests for information, which featured in about one-third of participants' examples of privacy lies. Generally, participants were less worried about losing information than they were that the request was *appropriate* in the context of the interaction. When they judged that it was not, as when a person requested information that was too intimate or a system requested information that was not needed, participants used privacy lies to not just protect privacy but also to assert autonomy. This aligns with Fuster's argument that providing inaccurate information online allows for the "preservation of the informational autonomy of individual[s]" in the face of unprecedented levels of data collection that offers them few other forms of recourse [18]; here, too, the protection of information is a means toward the function of autonomy.

Going a step further down this path, it might be more fruitful to move from thinking about "privacy functions" to thinking about "functions" more generally, and how PPBs support them. Consider Hecht et al.'s paper on how people use the location field in Twitter profiles, which highlighted a number of reactions to providing location information ranging from commentary on the need for such information ("NON YA BISNESS!!") to self-expression and playfulness [25]. Analyzing these data purely from a privacy and information

control/accuracy lens might miss important factors that help drive what might generically look like misrepresentation.

Contextual Differences

Comparing privacy lies to people and to systems helped illuminate how people make decisions about this practice based on the interaction context. People do tell privacy lies to both people and systems online, and some high-level motivations cut across contexts (e.g., lying to avoid general harms, to explore a persona, or for the sake of being private). Other motivations, however, tended to be more salient in one context or the other. For example, people were much more likely to tell privacy lies to gain benefits from systems than from people, while potential offline repercussions were a more salient driver of privacy lies to people than to systems. Key considerations of how people evaluated the request and requestor also appeared to be context-dependent. Requests from systems that led to privacy lies tended to be "unneeded" while those from people tended to be "uncomfortable". Meanwhile, people were more likely to describe a person rather than a system they lied to as "unfamiliar", and though "trust" was a concern raised in both contexts, compared to familiarity it was much less of a concern during interactions with people than with systems.

Note that we are discussing these findings not because we think the main takeaway is "people raise more concerns about X in context Y than context Z" (though such information is likely useful in guiding privacy inquiries and design). Rather, we worry that too often—including in our own statistical model—questions about privacy concerns, willingness to disclose information, and so on, are treated in a context-independent way. This is almost certainly wrong: the answer to "how willing am I to disclose my income" or "how worried am I about my data being shared with third parties" must be "it depends"—and research on privacy needs to be careful to both set context and bound conclusions.

Limitations and Future Work

In our own case, the definition of context as systems versus people is coarse, and though it is in line with much of the other privacy literature, we see it as a limitation. Further, as discussed before, the distinction between systems and people is often blurry online. A privacy lie on Facebook could be undertaken either to protect against the company, from other people on Facebook, third parties, or some combination of all of them. As with surveys, system designs can (and should) provide information and guidance in ways that make certain goals, concerns, and behaviors more salient, as with the privacy nudges work from Wang et al. [68].

Using an MTurk sample is common in recent CHI privacy research [9; 28; 66]. MTurk samples have been found to track reasonably well with national benchmark samples [13], though one risk to using them in privacy research is that MTurkers can have higher privacy concerns than average [33]. As an MTurk sample, it may under-represent older adults [29] or over-represent people who spend a lot of time online; however, in our sample, neither age (p=.61) nor time spent online (p=.16) predicted lying frequency. Finally, since we restricted our sample to MTurkers in the U.S., our findings speak to American contexts, and future work could explore the variation in privacy lies across cultures.

Our study also focused on self-reported recall-based lies, and though prior deception research argues that recall-based lies are acceptable [57], studies of lies in situ might lead to additional insights. More generally, we see value in bringing the literature on deception to bear on studying privacy lies and other PPBs perceived to be deceptive. This paper starts down that road by considering attitudes about privacy lies, and our measure of it is based on factors known to affect truthful disclosure. However, we couldn't say much about the individual factors of perceived prevalence, morality, and effectiveness as predictors of privacy lies, as in our data they collapse into a single reliable measure. Exploratory analyses using these as individual predictors do show that each is individually significant, suggesting that a deeper exploration of the constructs and mechanisms called out by the deception literature has potential to move privacy research and models of deceptive PPBs forward.

Our results suggest there is room for such motion toward a more comprehensive and predictive theoretical model of privacy lies. Given that our predictive model explained 35% of the variance in frequency of privacy lies, our findings suggest that there is a need to explore how other contextual and motivational factors may be captured, both for understanding privacy lies as well as other PPBs in general.

CONCLUSION

In this study, we put forth an improved model to predict privacy lies that incorporates actual privacy-protective behaviors, perceived control over data, and attitudes about lying in addition to commonly used scales of privacy-related concerns. We then identified how contextual factors influence privacy lies during interactions with people and with systems, such as perceptions of the request and requestor. Finally, we characterized the different types of privacy lies, the functions they serve, and the reasons people have for telling or abstaining from privacy lies in multiple contexts. Overall, our findings highlight the need to examine a range of contextual factors and motivations, beyond general privacy concerns, in understanding privacy lies as well as privacy-protective behaviors more generally.

REFERENCES

1. Anita Allen. 2011. *Privacy Law and Society* (2nd. ed.). West Academic Publishing.

2. Eric P. S. Baumer, Jenna Burrell, Morgan G. Ames, Jed R. Brubaker, and Paul Dourish. 2015. On the importance and implications of studying technology non-use. *Interactions* 22, 2: 52–56.

3. Natalya N. Bazarova and Yoon Hyung Choi. 2014. Self-disclosure in social media: Extending the functional approach to disclosure motivations and characteristics on social network sites. *J Commun* 64, 4: 635–657.

4. Cristina Bicchieri and Erte Xiao. 2009. Do the right thing: but only if others do so. *J Behav Decision Making* 22, 2: 191–208.

5. danah boyd and Eszter Hargittai. 2010. Facebook privacy settings: Who cares? *First Monday* 15, 8.

6. Laura Brandimarte, Alessandro Acquisti, and George Loewenstein. 2013. Misplaced confidences: Privacy and the control paradox. *Soc Psychol Pers Sci* 4.3: 340–347.

7. Alex Braunstein, Laura Granka, and Jessica Staddon. 2011. Indirect content privacy surveys: measuring privacy without asking about it. In *Proceedings of the Seventh Symposium on Usable Privacy and Security* (SOUPS '11). Article 15.

8. Moritz Büchi, Natascha Just, and Michael Latzer. 2017. Caring is not enough: the importance of Internet skills for online privacy protection. *Inform Commun Soc* 20, 8: 1261–1278.

9. Brooke Bullek, Stephanie Garboski, Darakhshan J. Mir, and Evan M. Peck. 2017. Towards understanding differential privacy: when do people trust randomized response technique? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '17), 3833–3837.

10. Daphne Chang, Erin L. Krupka, Eytan Adar and Alessandro Acquisti. 2016. Engineering Information Disclosure: Norm Shaping Designs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '16), 587–597.

11. Jeffrey T. Child, Sandra Petronio, Esther A. Agyeman-Budu, and David A. Westermann. 2011. Blog scrubbing: Exploring triggers that change privacy rules. *Comput Human Behav* 27, 5: 2017–2027.

12. Yoon Hyung Choi and Natalya N. Bazarova. 2015. Self-disclosure characteristics and motivations in social media: Extending the functional model to multiple social network sites. *Hum Commun Res* 41, 4: 480–500.

13. Scott Clifford, Ryan M. Jewell, and Philip D. Waggoner. 2015. Are samples drawn from Mechanical Turk valid for research on political ideology? *Res Politics* 2, 4: 1–9.

14. Mary J. Culnan. 1993. "How did they get my name?": An exploratory investigation of consumer attitudes toward secondary information use. *MIS Quart*: 341–363.

15. Bernhard Debatin, Jennette P. Lovejoy, Ann-Kathrin Horn, and Brittany N. Hughes. 2009. Facebook and online privacy: Attitudes, behaviors, and unintended consequences. *J Comput Mediat Commun* 15, 1: 83–108.

16. Tamara Dinev and Paul Hart. 2006. An extended privacy calculus model for e-commerce transactions. *Inform Sys Res* 17, 1: 61–80.

17. Adrienne Porter Felt, Serge Egelman, and David Wagner. 2012. I've got 99 problems, but vibration ain't one: a survey of smartphone users' concerns. In *Proceedings of the Second ACM Workshop on Security and Privacy in Smartphones and Mobile devices* (SPSM '12), 33–44.

18. Gloria González Fuster. 2010. Inaccuracy as a privacy-enhancing tool. *Ethics Inform Tech* 12, 1: 87–95.

19. Cathy Goodwin. 1991. Privacy: Recognition of a consumer right. *J Pubic Policy Mark* 10, 1: 149–166.

20. Ralph Gross and Alessandro Acquisti. 2005. Information revelation and privacy in online social networks. In *Proceedings of the ACM Workshop on Privacy in the Electronic Society*, Alexandria, Virginia, USA, 71–80.

21. Mariea Grubbs Hoy and George Milne. 2010. Gender differences in privacy-related measures for young adult Facebook users. *J Interact Advert* 10, 2: 28–45.

22. Shion Guha and Stephen B. Wicker. 2015. Spatial subterfuge: an experience sampling study to predict deceptive location disclosures. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (UbiComp '15), 1131–1135. https://doi.org/10.1145/2750858.2804281

23. Babita Gupta, Lakshmi S. Iyer, and Robert S. Weisskirch. 2010. Facilitating global e-commerce: A comparison of consumers' willingness to disclose personal information online in the US and India. *J Electron Commer Res* 11, 1: 28–45.

24. Jeffrey T. Hancock, Jeremy Birnholtz, Natalya Bazarova, Jamie Guillory, Josh Perlin, and Barrett Amos. 2009. Butler lies: awareness, deception and design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '09), 517–526.

25. Brent Hecht, Lichan Hong, Bongwon Suh, and Ed H. Chi. 2011. Tweets from Justin Bieber's heart: the dynamics of the location field in user profiles. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '11), 237–246.

26. Donna L. Hoffman, Thomas P. Novak, and Marcos Peralta. 1999. Building consumer trust online. *Commun ACM* 42, 4: 80–85.

27. Daniel R. Horne, Patricia A. Norberg, and A. Cemal Ekin. 2007. Exploring consumer lying in information-based exchanges. *J Consum Mark* 24, 2: 90–99.

28. Roberto Hoyle, Srijita Das, Apu Kapadia, Adam J. Lee, and Kami Vaniea. 2017. Was my message read?: Privacy and signaling on Facebook messenger. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '17), 3838–3842.

29. Connor Huff and Dustin Tingley. 2015. "Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents. *Res Politic* 2, 3: 1–12.

30. Zhenhui Jiang, Cheng Suang Heng, and Ben C. F. Choi. 2013. Research note—privacy concerns and privacy-protective behavior in synchronous online social interactions. *Inform Sys Res* 24, 3: 579–595.

31. Adam N. Joinson, Carina B. Paine, Tom Buchanan, and Ulf-Dietrich Reips. 2007. Privacy, trust and self-disclosure to web-based surveys. In *Proceedings of the 2007 Association for Survey Computing Conference (ASC '07),* 133–144.

32. Adam N. Joinson, Ulf-Dietrich Reips, Tom Buchanan, and Carina B. Paine Schofield. 2010. Privacy, trust, and self-disclosure online. *Hum Comput Interact* 25, 1: 1–24.

33. Ruogu Kang, Stephanie Brown, Laura Dabbish, and Sara Kiesler. 2014. Privacy attitudes of Mechanical Turk workers and the US public. In *Symposium on Usable Privacy and Security* (SOUPS '14), 37–49.

34. Flavius Kehr, Tobias Kowatsch, Daniel Wentzel, and Elgar Fleisch. 2015. Blissfully ignorant: the effects of general privacy concerns, general institutional trust, and affect in the privacy calculus. *Informat Sys J* 25, 6: 607–635.

35. Hanna Krasnova, Elena Kolesnikova, and Oliver Guenther. 2009. "It won't happen to me!": Self-disclosure in online social networks. In *Proceedings of the 15th Americas Conference on Information Systems* (AMCIS 2009), 1–10.

36. Cliff Lampe, Nicole B. Ellison, and Charles Steinfield. 2008. Changes in use and perception of Facebook. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work* (CSCW '08), 721–730.

37. Amanda Lenhart and Mary Madden. 2007. Teens, privacy & online social networks: How teens manage their online identities and personal information in the age of MySpace (p. ii). Pew Internet & American Life Project: Washington, DC, 18 April 2007.

38. Laura H. Lind, Michael F. Schober, Frederick G. Conrad, and Heidi Reichert. 2013. Why do survey respondents disclose more when computers ask the questions? *Pub Opinion Quart* 77, 4: 888–935.

39. May O. Lwin, Jochen Wirtz, and Jerome D. Williams. 2007. Consumer online privacy concerns and responses: a power-responsibility equilibrium perspective. *J Acad Market Sci* 35, 4: 572–585.

40. Mary Madden and Lee Rainie. 2015. Americans' attitudes about privacy, security and surveillance. Pew Research Center. Retrieved from http://www.pewinternet.org/

41. James E. Maddux and Ronald W. Rogers. 1983. Protection motivation and self-efficacy: A revised theory of fear appeals and attitude change. *J Exper Soc Psychol* 19, 5: 469–479.

42. Naresh K. Malhotra, Sung S. Kim, and James Agarwal. 2004. Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Inform Sys Res* 15, 4: 336–355.

43. Steven A. McCornack. 1992. Information manipulation theory. *Commun Monographs* 59, 1: 1–16.

44. Miriam J. Metzger. 2004. Privacy, trust, and disclosure: Exploring barriers to electronic commerce. *J Comput-Mediat Commun* 9, 4: 00–00.

45. George R. Milne, Andrew J. Rohm, and Shalini Bahl. 2004. Consumers' protection of online privacy and identity. *J Consum Aff* 38, 2: 217–232.

46. Helen Nissenbaum. 2004. Privacy as contextual integrity. *Wash Law Rev* 79: 119.

47. Patricia A. Norberg, Daniel R. Horne, and David A. Horne. 2007. The privacy paradox: Personal information disclosure intentions versus behaviors. *J Consum Aff* 41, 1: 100–126.

48. Judith S. Olson, Jonathan Grudin, and Eric Horvitz. 2005. A study of preferences for sharing and privacy. In *SIGCHI Extended Abstracts on Human factors in computing systems* (CHI EA '05), 1985–1988.

49. Xinru Page, Bart P. Knijnenburg, and Alfred Kobsa. 2013. What a tangled web we weave: lying backfires in location-sharing social media. In *Proceedings of the 2013 Conference on Computer-Supported Cooperative Work* (CSCW '13), 273–284.

50. Carina Paine, Adam N. Joinson, Tom Buchanan, and Ulf-Dietrich Reips. 2006. Privacy and self-disclosure online. In CHI '06 Extended Abstracts on Human Factors in Computing Systems (CHI EA '06). ACM, New York, NY, USA, 1187-1192.

51. Leysia Palen and Paul Dourish. 2003. Unpacking privacy for a networked world. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '03), 129–136.

52. Sandra Petronio. 2008. *Communication privacy management.* John Wiley & Sons, Inc.

53. Joseph Phelps, Glen Nowak, and Elizabeth Ferrell. 2000. Privacy concerns and consumer willingness to provide personal information. *J Public Policy Mark* 19, 1: 27–41.

54. Amit Poddar, Jill Mosteller, and Pam Scholder Ellen. 2009. Consumers' rules of engagement in online information exchanges. *J Consum Aff* 43, 3: 419–448.

55. Lee Rainie, 2016. The state of privacy in post-Snowden America. Pew Research Center. Retrieved from http://www.pewresearch.org/

56. Nora J. Rifon, Robert LaRose, and Sejung Choi. 2005. Your privacy is sealed: Effects of web privacy seals on trust and personal disclosures. *J Consum Aff* 39, 2: 339–362.

57. Kim B. Serota, Timothy R. Levine, and Franklin J. Boster. 2010. The prevalence of lying in America: Three studies of self-reported lies. *Human Commun Res* 36, 1: 2–25.

58. Kim Bartel Sheehan. 1999. An investigation of gender differences in on-line privacy concerns and resultant behaviors. *J Interact Mark* 13, 4: 24–38.

59. Kim Bartel Sheehan. 2002. Toward a typology of Internet users and online privacy concerns. *Inform Society* 18, 1: 21–32.

60. Kim Bartel Sheehan and Marlea Grubbs Hoy. 1999. Flaming, complaining, abstaining: How online users respond to privacy concerns. *J Advertising* 28, 3: 37–51.

61. Irina Shklovski, Scott D. Mainwaring, Halla Hrund Skúladóttir, and Höskuldur Borgthorsson. 2014. Leakiness and creepiness in app space: perceptions of privacy and mobile app use. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14), 2347–2356.

62. Jeff H. Smith, Sandra J. Milberg, and Sandra J. Burke. 1996. Information privacy: measuring individuals' concerns about organizational practices. *MIS Quart*: 167–196.

63. Jai-Yeol Son and Sung S. Kim. 2008. Internet users' information privacy-protective responses: A taxonomy and a nomological model. *MIS Quart:* 503–529.

64. Catalina L. Toma, Jeffrey T. Hancock, and Nicole B. Ellison. 2008. Separating fact from fiction: An examination of deceptive self-presentation in online dating profiles. *Pers Soc Psychol Bull* 34, 8: 1023–1036.

65. Joseph Turow and Michael Hennessy. 2007. Internet privacy and institutional trust: insights from a national survey. *New Media Soc* 9, 2: 300–318.

66. Wali Ahmed Usmani, Diogo Marques, Ivan Beschastnikh, Konstantin Beznosov, Tiago Guerreiro, and Luís Carriço. 2017. Characterizing social insider attacks on Facebook. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '17), 3810–3820.

67. Sonja Utz. 2005. Types of deception and underlying motivation: What people think. *Soc Sci Comput Rev* 23, 1: 49–56.

68. Yang Wang, Pedro Giovanni Leon, Alessandro Acquisti, Lorrie Faith Cranor, Alain Forget, and Norman Sadeh. 2014. A field trial of privacy nudges for Facebook. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '17), 2367–2376.

69. Alan Westin. 1967. *Privacy and freedom*. Atheneum, New York.

70. Monica T. Whitty and Siobhan E. Carville. 2008. Would I lie to you? Self-serving lies and other-oriented lies told across different media. *Comput Human Behav* 24, 3: 1021–1031.

71. Monica T. Whitty and Jeff Gavin. 2001. Age/sex/location: Uncovering the social cues in the development of online relationships. *CyberPsychol Behav* 4, 5: 623–630.

72. Heng Xu. 2007. The effects of self-construal and perceived control on privacy concerns. In *Proceedings of the International Conference on Information Systems* (ICIS '07), 1–14.

73. Seounmi Youn and Kimberly Hall. 2008. Gender and online privacy among teens: Risk perception, privacy concerns, and protection behaviors. *Cyberpsychol Behav* 11, 6: 763–765.

74. Alyson L. Young and Anabel Quan-Haase. 2013. Privacy protection strategies on Facebook: The Internet privacy paradox revisited. *Inform Commun Soc* 16, 4: 479–500.

75. J. Christopher Zimmer, Riza Ergun Arsal, Mohammad Al-Marzouq, and Varun Grover. 2010. Investigating online information disclosure: Effects of information relevance, trust and risk. *Inform Manag* 47, 2: 115–123.