When and Why Consumers Respond to Online Privacy Violations

Chi Tran
University of Oregon
ctran4@uoregon.edu

Brandon J. Reich
Portland State University
breich@pdx.edu

Hong Yuan
University of Oregon
hongy@uoregon.edu

Abstract

As consumer privacy concerns become paramount, it is increasingly critical to understand what constitutes a privacy violation and how consumers respond. Using a multi-method approach, this research shows that consumers perceive three privacy violation types with increasing levels of severity, explaining when and why consumers exhibit seemingly paradoxical responses to different violation types. Our theorizing and findings suggest that resource control is the predominant mechanism driving privacy behavior, and that situation controllability (operationalized as level of variability of privacy practices within the industry) may moderate its effects.

Keywords: consumer privacy, privacy violation, resource control, multi-method, text analysis.

1. Introduction

In April 2019, Facebook's CEO Mark Zuckerberg declared to the world that "The future is private" and that the world's largest social media platform would be pivoting to a privacy-focused strategy (Nivea, 2019). This denotes a significant shift from his own declaration in 2010 that privacy was an obsolete social norm of the past (Kirkpatrick, 2010). Other companies have followed suit, changing their regard for privacy from a burdensome cost for risk management (Bamberger & Mulligan, 2010) to a growing recognition of privacy as a potential source of competitive advantage. Apple, for example, recently released a major ad campaign with the simple tagline: "Privacy. That's iPhone." (Wuerthele, 2019).

Clearly privacy is more important to consumers than ever before (Kim, Barasz, & John, 2018; Krishna, 2020), due to growing digitization of consumption (Schmitt, 2019) combined with marketers' own exploitation of consumer data over the past decade (Benes, 2018). More surprisingly, while consumers report unprecedented levels of privacy concerns (PwC 2017) and lack of control over their personal information (Pew, 2019), they continue to patronize the products and services of many of the companies that routinely violate their privacy. For instance, consumers continue to tolerate intrusive privacy practices by companies such as Amazon (Paul 2020) and Google (MacMillan & McMillan, 2018), while Facebook's similarly intrusive privacy violations have produced fierce consumer backlash (Shane, 2018).

For all its importance to modern consumption and marketing, a consensus definition of consumer privacy remains elusive (Acquisti, Taylor, and Wagman 2016; Martin and Murphy 2017), in part because privacy is a constantly evolving and "essentially contested" construct (Mulligan, Koopman, and Doty 2016, p.1). Motivated by this dearth of knowledge, the current research has two core aims. First, we seek to clarify the construct of privacy violation from a consumer perspective. Using an established construct development methodology, we establish that consumers perceive three privacy violation types reflecting a linear increase in severity: *recording* (i.e., merely observing and storing consumers' personal information), *targeting* (i.e., using consumers' information for targeted advertisements), and *sharing* (i.e., sharing consumers' information with a third party). This conceptualization in turn facilitates our second aim, explaining the consumers' paradoxical responses to privacy violations.

In this manuscript, we first present a rigorous review of several literatures to build our conceptual framework. We then use a multi-method approach to test this conceptualization across three studies. Study 1 first clarifies the construct of privacy violation from a consumer perspective, uncovering and confirming three increasingly severe violation types: *recording*, *targeting*, and *sharing* of personal information using survey design and text analysis. We then use this conceptualization to shed light on consumer responses to privacy violations. Using a between-participants experiment (study 2), we test an industry-level

URI: https://hdl.handle.net/10125/103162
978-0-9981331-6-4

manifestation of situation controllability—variability in privacy practices (henceforth "industry variability")—that helps explain opposing consumer responses to otherwise equivalent privacy violations across companies. Using text analysis of scraped Twitter data, study 3 tests these effects in a naturalistic setting and employs a different measure of control-reclaiming behavior.

2. Conceptual Framework

2.1. Consumer Privacy and Privacy Violation

Privacy has traditionally been defined as a right "to be left alone" (Brandeis & Warren, 1890), specifically with reference to one's own physical space. Due to recent shifts in technology, contemporary definitions of consumer privacy often imply information privacy (Goodwin, 1991) especially in an online consumption domain. Indeed, the "information age" has created such widespread consumer demand for privacy over their information that it is sometimes regarded as a commodity that could (and should) be regulated through a market structure (Acquisti et al., 2016; Smith et al., 2011). In theory, this suggests that consumers may willingly exchange privacy and personal information for access to products and services.

Despite this theoretical guidance and the growing attention paid to consumer privacy, there remains a lack of consensus as to what constitutes a violation of privacy (K. D. Martin & Murphy, 2017). How should marketers tread this evolving consumer landscape? The marketing literature has provided some guidance as to conceptualizing privacy violation, loosely defined as unwanted marketing communications, highly targeted advertisement, and secretive online tracking (Nill & Aalberts, 2014). This approach focuses on companies' use of consumer data without consent. While providing a useful foundation, this definition does little to address the increasingly nuanced use of consumer data in the current digital environment. Is seeing an ad about a product one might like "unwanted?" How targeted is "highly targeted?" Would consumers consider an advertisement specifically targeted at them a "privacy violation?" These questions remain open, portending a lack of clarity in the concept of privacy violation from a consumer perspective.

Addressing this gap in the literature, we investigate how consumers conceptualize privacy violations. Past work reveals that consumers might react strongly and negatively only when a large amount of information is collected or when the type of information is highly sensitive (e.g., financial, medical). In this work, we content that the quantity or quality of information collected is indeed important, and add to this literature by focusing on another dimension of privacy violations. This work suggests that consumers emphasize what the company *does* with their information in determining whether and to what extent a privacy violation has occurred. Specifically, using a combination of interview, survey, and experimental methodologies, our results reveal three categories of privacy violation—*recording*, *targeting*, and *sharing*—that clarify variations in perceived privacy violations and represent a linear increase in severity.

We further propose and show that these violation types also affect consumer response *because* they represent increasing levels of privacy violation. Given that consumer privacy is defined in terms of consumers' perceived control over the collection and use of their personal information (Goodwin, 1991), it follows that the more companies exploit consumers' data, the less control consumers have over their information, and in turn the more their privacy is perceived as being violated. Across these three violation types, *recording* represents the least amount of usage or manipulation over consumer data. Conversely, *targeting* and *sharing* both involve further manipulation and exploitation of such data, suggesting that consumers likely perceive these violation types as more severe.

Interestingly, *targeting* and *sharing* are often referred to interchangeably as "personalization" in the marketing literature (Vesanen & Raulas, 2006). For example, behavioral targeting is defined as utilizing consumer behavioral data from multiple sources to deliver personalized ads to users (Summers, Smith, & Reczek, 2016); or adaptive personalization involves constantly updating user preference using data shared among platforms and systems to deliver targeted ads to consumers (Chung, Wedel, & Rust, 2016; Kazienko & Adamski, 2007). These examples suggest that delivering targeted recommendations to consumers and customizing ads by sharing consumer data with a third-party would both be considered the same "personalization" practice according to current conceptualizations offered by extant literature. Yet, gossip theory (K. D. Martin, Borah, & Palmatier, 2017) suggests that consumers may perceive *sharing* to be more intrusive than *targeting*. In social relationships, people feel reduced control over their personal information when it is shared with another party (Kurland & Pelled, 2000; Wert & Salovey, 2004). When they learn that they are the subject of gossip, for instance, most people react negatively *because* their privacy has been severely violated (Baumeister, Zhang, & Vohs, 2004; Foster, 2004). Consistent with this theoretical perspective, we expect that consumers will perceive *sharing* (vs. *targeting*) as a more severe privacy violation because it reduces their perceived control over their own information, and both will be

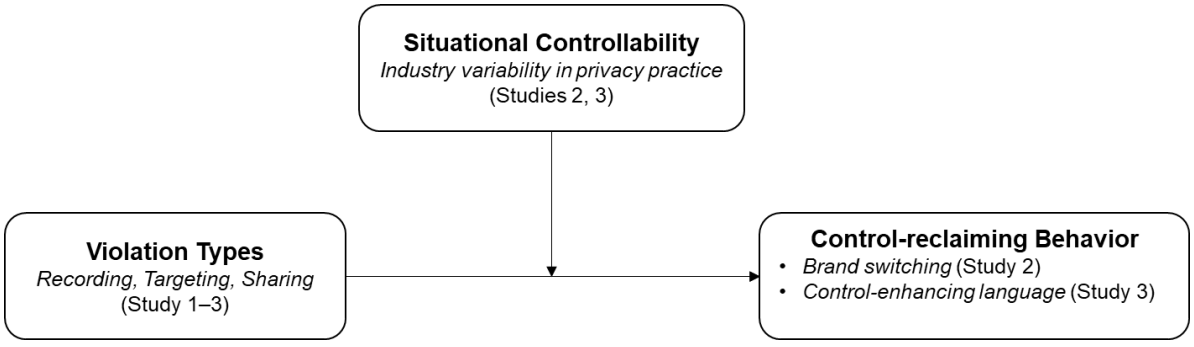
perceived as more severe than *recording* for this same reason. Note that in this work we focus specifically on these intentional company actions (*recording, targeting, sharing*) in online environments, and do not cover physical intrusion or data breaches (in which both the company and consumers are victims of illegal practices).

2.2. Resource Control and Situational Controllability

We have thus far conceptualized privacy violation in terms of consumers' perceived control over their personal information. Lacking control is generally considered aversive and the literature suggests that consumers often try to regain control when it is threatened or lost. (Whitson & Galinsky, 2008). This suggests that, as the severity of privacy violation increases, a decrease in resource control leads to an increased likelihood that consumers will attempt to regain control over their information. Yet, as illustrated by our opening example, a privacy paradox exists (Norberg, Horne, & Horne, 2007; Sheehan & Hoy, 1999) whereby consumers sometimes fail to respond appropriately following a privacy violation. To resolve these conflicting theoretical perspectives, we draw from literatures on self-determination and coping. Self-determination theory corroborates our prediction that individuals cope with threats to control by engaging in control-reclaiming behaviors (Ryan & Deci, 2000) such as domain-specific attempts to solve a problem (e.g., through brand selection; (Inesi, Botti, Dubois, Rucker, & Galinsky, 2011; Schiele & Venkatesh, 2016) or as generalized self-affirmations of one's overall sense of control (e.g., through verbal expressions; Thimm, Rademacher, & Kruse, 1995). However, the coping literature adds nuance and suggests that consumers may respond more passively in situations in which regaining control is objectively impossible or seems impossible (Skinner & Zimmer-Gembeck, 2011). In such cases, a sense of helplessness inhibits control-reclaiming behaviors and no action is taken (Dweck, 2000; Sedek, Kofta, & Tyszka, 1993; Skinner, 1996). That is, in a consumer privacy context, consumers might "resign to giving up their data" when they feel "an undesirable outcome is inevitable and [they are] powerless to stop it" (Turow, Hennessy, & Draper, 2016).

The literature discusses how a specific marketing approach (e.g., geo-targeting, personalized email; Sahni, Wheeler, & Chintagunta, 2018) a platform characteristic (e.g. privacy notice—K. E. Martin, 2014), or a particular industry (e.g., social media—Hajli & Lin, 2016) may affect consumers' privacy response. Drawing from this research, we propose an industry-level variable that may determine situational controllability in consumer privacy contexts, and therefore explain the privacy paradox. Industry variability—the amount of variability between companies within an industry in terms of their privacy practices—may indeed explain why severe (vs. moderate) privacy violations prompt consumer action in some market contexts (e.g., social media) and not others (e.g., online retail). This adds to the current literature in privacy, which often focuses on individual differences in explaining privacy behaviors. We contend that in addition to individual differences, certain market conditions could also enable or inhibit consumers' behavioral responses to privacy violation. In sum, we predict that a severe (vs. moderate) privacy violation will prompt greater control-reclaiming behaviors, but only when the violating company operates in an industry with high variability of privacy practices. This is because, in this market context, regaining resource control is still objectively possible. However, in a homogenous industry in which companies share a standardized privacy practice, consumers are unlikely to do so regardless of violation severity because regaining control, objectively, is not possible. Figure 1 summarizes our conceptual framework.

Figure 1: Conceptual Framework



3. Empirical Findings

3.1. Study 1

In Study 1, we asked participants to provide their own definition of privacy and example of privacy violation and subjected these open-ended responses to content analysis. In addition, participants completed a scenario-based survey, and quantitative analyses were employed to identify common themes among actual privacy incidents. Together, findings converged into the conceptualization of three violation types.

3.1.1. Method. Study 1 utilized several actual instances of consumer privacy violations that have occurred in the past ten years. Participants (N = 251 US residents; M_{Age} = 36.32, SD_{Age} = 11.70; 41.01% female) were recruited from Amazon Mechanical Turk (MTurk) and a university's subject pool and were first asked to provide their own privacy definition in a text box. They were then asked to provide an example of a privacy violation. Next, they were presented with twelve scenarios, each presented on its own page and in randomized order (see figure 1 for scenarios). To create these scenarios, we constructed a list of actual privacy incidents reported in the major business press in the past ten years (e.g., The Wall Street Journal, Business Insider, Forbes, etc.) using the business press database Factiva. We excluded incidents of data breaches in which both the company and consumer were victims and filtered out incidents caused by personal errors. We then selected scenarios that garnered substantial media interest at the time and adapted them into short descriptions without mentioning specific companies' names. Each participant's self-reported privacy definition was piped back into a display under each scenario for reference, and participants rated their perceived level of privacy violation for each using a single item: "This action is…" (1 = *Not at all a privacy violation*; 7 = *An extreme privacy violation*).

3.1.2. Principal Component Analysis (PCA). A PCA with Varimax rotation (figure 2) found that the twelve ratings converged into three components with step levels of perceived privacy violation. Component 1 (*recording*; M = 3.38, SD = 1.30) consisted of scenarios in which companies merely collected consumer information (e.g., "CCTV in the supermarket aisle that tracks customers' activity in-store."). Component 2 (*targeting*; M = 5.16, SD = 1.21) included situations in which companies used consumer information for targeting purposes (e.g., "A retailer uses a customer's purchase history to predict that a customer is pregnant and send pregnancy advertisements to her address."). Component 3 (*sharing*; M = 6.12, SD = 1.06) consisted of actions involving sharing information with third parties (e.g., "A social media app allows third-party developers access to users' data."). A follow-up

Figure 2: Study 1. PCA Results (numbers in the composition column represent factor loadings)

S		violation			
Discount Scan	A retailer offers discounts to its members when they scan their mobile barcode from the membership app.	2.55	0.705		
CashierZip	A cashier asks for a customer's zip code at check out	3.19	0.711		
Coupon Phone	A supermarket app provides coupons of produce on a customer's phone when s/he around the produce section in-store.	3.78	0.756		Recording
CCTV	CCTV in the supermarket aisle that tracks customers' activity in-store.	4.01	0.651		
Search History	A company uses search history to provide ads that are specifically tailored to users.	4.89	0.768		
DNAsite	DNA information submitted to an online genealogy website is used to identify a suspect in a criminal case.	4.91	0.606		
SMTarget Ads	A social media app uses users' shared information to provide targeted advertising under sponsored posts.	5.12	0.777		Targeting
Home-Sharing	A home sharing app uses users' search history to predict high-demand dates and set prices accordingly.	5.14	0.743		
Pregnancy Ads	A retailer uses a customer's purchase history to predict that a customer is pregnant and send pregnancy advertisements to her address.	5.49	0.704		
Mobile Service	A mobile network provider offers a service to other companies to use the provider's customer call & browsing info to map locations & apps usage	6.04		0.764	
Developer Access	A social media app allows third-party developers access to users' data.	6.09		0.611	Sharing
Record Convo	A smart home device accidentally records and sends a conversation of a user to another person.	6.24		0.842	

repeated-measures ANOVA comparing the composite means of these components suggested a significant omnibus difference (Huynh-Feldt $F(1.77, 243.67) = 262.28$, $p < .001$, $\eta_p^2 = .65$), and planned polynomial contrasts showed a significant linear trend ($F(1, 138) = 376.36$, $p < .001$). This suggests that consumers perceive progressively increasing levels of privacy violation depending on whether the violation type is *recording*, *targeting* or *sharing*.

3.1.3. Content Analysis. To further validate our categorization of privacy violation, we conducted follow-up content analyses using both automated topic modelling technique (Berger et al., 2020; Humphreys & Wang, 2018) and conventional thematic coding (Kassarjian, 1977) on the examples of privacy violation that participants provided at the beginning of the study. To prepare the data for automated analysis, we first converted all words to lower case and cleaned the data for typos, emojis, URLs, stop words, and punctuations (Berger et al., 2020). We then applied Latent Dirichlet Allocation (LDA; Blei, Ng, and Jordan 2003) technique to extract potential topics and themes underlying participants' provided example of privacy violations. This approach, similar to PCA, allows groupings of words that frequently occur together into themes or topics, and responses might overlap with the same set of topics in varying probabilities. However, different from the PCA procedure, the LDA algorithm is unconstrained by pre-defined categories and pre-selected scenarios and therefore provides a more naturalistic consumer-generated categorization of privacy violation.

We conducted our analysis with different numbers of topics to generate the set of topics that best balances parsimony with discrimination from one another based on the most frequent keywords. Multiple iterations of the procedure resulted in a final set of five topics, labeled as "recording," "targeting," "sharing," "data breach," and "illegal/ physical intrusion" based on the individual words and responses that represent each topic (Table 1). Although the latter two topics are clearly relevant to privacy in general, they represent aberrant violations and are therefore less germane to the current research context of privacy violations routinely levied by marketers in ordinary consumption situations. Consequently, data breach and illegal/physical intrusion were discarded from our conceptualization of privacy violation. The remaining three topics therefore added support for our three-dimension conceptualization of privacy violation faced by consumers.

Table 1: Study 1: Topic Modelling (LDA) Results On Privacy Violation Examples

Topics	Top keywords*	Representative response text**
Recording	monitoring, track, violation, record, permission	"The electronics that we use as a consumer (laptops, phones, cameras). Companies monitoring what you're doing on those machines. Even websites as well" "Google keeping record of what I do on the internet to personalize my page"
Targeting	Targeting, online, search, ads, history	"The facebook fiasco. Targeting political ads based on search history." "The fact that any information I put out in the internet (search and purchase history) is used to generate hundreds of ads on my phone"
Sharing	personal, selling, consent, sharing, (third) party(ies)	"Someone sharing my personal information without my consent. Ex: different websites knowing what I am viewing." "Facebook selling your private information to the highest bidder."
Data breach	data, breach, stolen, break, leaked	"My username and password get leaked in a data breach." "My information was stolen through a data breach once."
Illegal/ physical intrusion	people, hack(-ers,-ing,-ed), card, information, windows	"A customer's information from applying for a credit card is given or misused by the store employee they applied with." "People peeking into my windows."

3.1.4. Discussion. Study 1 explored and confirmed the nature of privacy violation from a consumer perspective using a robust combination of qualitative and quantitative methodologies. Specifically, a PCA using pre-defined scenarios and content analysis using consumer-generated examples of privacy violation converged to show that consumers may perceive distinct levels of privacy violation depending on what companies do with their information, increasing progressively in a linear trend across recording, targeting and sharing violation types. Our findings also highlight the distinction between sharing and targeting, concepts which extant literature has traditionally confounded. In addition, a follow-up analysis of the text showed that 73.46% of privacy definitions provided by participants were related to control over personal information (i.e., resource control), our theorized mechanism underlying privacy violation's effects. In study 2, we corroborate this conceptualization of violation types with an experimental design and test our

control-based explanation for the privacy paradox directly.

3.2. Study 2

Study 2 aimed to replicate and extend the findings of study 1 using a controlled experimental design among fictitious brands and with a different measure of control-reclaiming behavior (brand switching). We again expected increased control-reclaiming behavior in response to a sharing (vs. targeting) privacy violation when the industry is highly variable in its privacy practice. However, this effect should be mitigated when the industry is standardized. In this and the subsequent study, we focus more narrowly on the distinction between *targeting* and *sharing* to provide a more conservative and streamlined test of the underlying control-based mechanism. This also more accurately reflects most market situations, in which *recording* is inherently involved in both *targeting* and *sharing*, but the latter two need not include each other. In other words, *recording* is ubiquitous (Acquisti, Brandimarte, & Loewenstein, 2020) but companies differ in terms of whether consumer data is used for *targeting* or *sharing* with third parties.

3.2.1. Method. We used a 2 (action: targeting, sharing) × 2 (industry variability: low, high) full-factorial design with 321 MTurk participants (N = 214 after attention check exclusions; M_{Age} = 37.38, SD_{Age} = 11.43; 53.27% female). Participants were given information about "Industry X," containing four companies (A, B, C, and D) and were asked to imagine themselves as current customers of "Company A," the market leader. In the low (high) variability condition, all companies in Industry X used consumer data in the same way (different ways). As in the pilot study, participants in the targeting (sharing) condition then saw a pop-up detailing Company A's cookies policy of using consumers' personal data to deliver targeted ads (share with third-parties). As a manipulation check, we measured perceived privacy violation with the same four-item scale (α = .85). As the core dependent variable, participants were asked to choose whether they would continue onto Company A's website (0) or switch to another brand (1). Lastly, participants were presented with two attention checks: One identical to that used in the pilot study and another asking which action was taken by Company A (delivering targeted ads or sharing with third parties).

3.2.2. Results/ Discussion. To check the action manipulation, we first conducted a 2 (action) × 2 (industry variability) ANOVA (*df* for *F*-tests = 1, 210; see table 2 for means and SDs) on perceived privacy violation. Results confirmed greater perceived privacy violation in the *sharing* (vs. *targeting*) condition (F = 13.22, p < .001), but no other effects (ps > .15). To test our core prediction, we conducted a separate ANOVA on switching behavior. We observed main effects of action (F = 9.12, p = .003) and industry variability (F = 6.81, p = .01), such that switching was more likely in response to *sharing* (vs. *targeting*) and when variability was high (vs. low). More importantly, we observed the expected interaction (F = 4.39, p = .04). Planned contrasts revealed that, in the high industry variability condition, participants were significantly more likely to switch brands following a *sharing* (vs. *targeting*) privacy violation (F = 11.26, p = .001). However, in the low industry variability condition, company action had no effect on participants' switching behavior (p = .50). These results replicate those observed in study 1, further suggesting that consumers only attempt to reclaim control after a severe privacy violation when market constraints permit such a possibility.

Table 2: Study 2 Mean (SD) for each condition

Study 2	High Industry Variability		Low Industry Variability	
	Targeting	*Sharing*	*Targeting*	*Sharing*
Privacy violation	4.23 (1.33)	4.97 (1.22)	3.94 (1.59)	4.67 (1.53)
Switching likelihood	0.24 (0.43)	0.56 (0.50)	0.21 (0.41)	0.26 (0.45)

3.3 Study 3

Our first two studies have shown consistent support for our theorizing in constructed scenarios involving forced choices among fictitious brands. To build generalizability, study 3 aims to replicate our findings in a more naturalistic setting involving actual brands in distinct industries and employing a directly observable measure of control-reclaiming behavior from a secondary source.

Specifically, study 3 used text analysis of Twitter data to compare linguistic response to targeting or sharing privacy violations from Facebook and Amazon, creating a naturalistic 2 (violation type: targeting, sharing) × 2 (industry variability: high, low) quasi-experiment. Facebook and Amazon represent two industries (social media and online retail, respectively) that naturally differ in terms of variability of privacy practices. We expect increased control-reclaiming behavior in response to a sharing (vs. targeting) privacy violation from Facebook because it operates in an industry with high variability in privacy practice (Ahmad 2018; Norton 2020). However, because Amazon operates in a more standardized industry (Paul

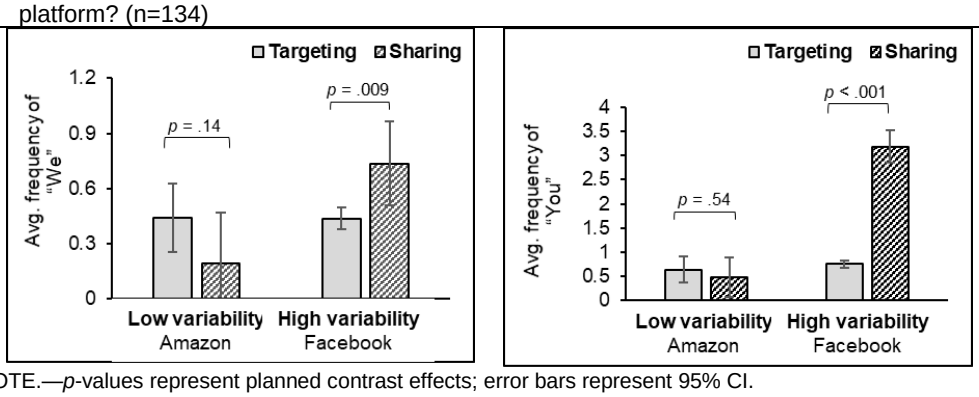
2020), a sense of helplessness should inhibit consumers' control-reclaiming behavior regardless of the company's privacy-violating action.

To operationalize control-reclaiming behavior, we draw from research in psycholinguistics on the use of personal pronouns. Personal pronouns have been shown to demonstrate a speaker's mental state or traits and influence cognitive processes (Chung and Pennebaker 2007; Kacewicz et al. 2014). In marketing and related fields, research into personal pronouns has grown in prominence in recent years thanks to the increasing availability of text data (Humphreys and Wang 2018). Research in this domain has shown that use of personal pronouns can affect relationships between firms and customers (Packard, Moore, and McFerran 2018) or influence cultural trends (Packard and Berger 2020). More relevant to the present study, Kacewicz et al. (2014) established that use of first-person plural ("we") and second-person singular ("you") pronouns both serve as verbal expressions of control because they imply a collectivist- and other-orientation, respectively (Cassell et al. 2006). Use of such control-enhancing language can serve as a way for individuals to reclaim control once it has been threatened, even if the individual is not consciously aware of it (Thimm et al. 1995). We therefore expect greater frequency of these control-enhancing pronouns among consumers tweeting about Facebook's sharing (vs. targeting) privacy violations because situation controllability is relatively high. Conversely, in response to Amazon (where situation controllability is low), pronoun frequency should not differ regardless of whether the tweets refer to sharing or targeting.

3.3.1. Method. We used package Rtweet to collect 4,165 tweets (N = 3,805 after removing duplicates) from the Twitter Developer API and matched keywords (Hewett et al. 2016) reflecting violation type (targeting or sharing) with either Amazon or Facebook. Keywords regarding targeting (sharing) were "targeting," "targeted," "targets" ("sharing data," "shares data," "shared data," "third party data," "3rd party data"). Following Berger et al. (2020), we excluded retweets and replies to ensure that the content was not biased towards the most popular tweets, and removed URLs, punctuation, digits, username tags, and emojis (see figure 2 for example tweets). We ran the cleaned data through the LIWC Dictionary (Pennebaker et al. 2015) to analyze frequencies (per tweet) of two linguistic expressions of control enhancement: Pronouns "we" and "you."

Figure 3: Study 3. Example Tweets and Effects of Violation Type × Industry Variability Interaction on Usage Frequency of "We" and "You"

	Amazon	Facebook
Targeting	"Targeted marketing tech allows the food industry to tailor promotions, transforming a one-time temptation into an all-the-time ad. 2/3 of products we purchased from @amazon were healthy, but the majority of ads after were for unhealthy food" (n= 299)	"Facebook refuses to restrict untruthful political ads and micro-targeting. You know what to do." (n=3171)
Sharing	Amazon has fired a number of employees after they shared customer email address and phone numbers with a third-party. Are you building your brand on Amazon? How do you control the security of your customers' private #data on that platform? (n=134)	"Facebook is normalizing sharing your healthcare data #privacy" (n=201)



NOTE.—*p*-values represent planned contrast effects; error bars represent 95% CI.

3.3.2. Results. For robustness and clarity, we examine effects on frequency of "we" and "you" separately. A 2 (violation type) × 2 (industry variability) ANOVA (df = 1, 3801 for F-tests) treating "we" frequency as the

dependent variable revealed no main effects (*p*s > .17), but a significant interaction (F = 6.85, p = .009, η_p^2 = .002). Planned contrasts supported our theorizing. For tweets referencing Facebook (figure 2, left panel), *sharing* (vs. *targeting*) was associated with significantly greater usage of "we" (F = 6.85, p = .009, η_p^2 = .002, $M_{Sharing}$ = 0.78, $SD_{Sharing}$ = 2.53, $M_{Targeting}$ = 0.41, $SD_{Targeting}$ = 1.58), whereas no effect of violation type was observed among tweets referencing Amazon (p = .14). A parallel ANOVA on "you" frequency displayed a similar pattern. Although we observed main effects of violation type (F = 96.15, p < .001, η_p^2 = .02) and variability (F = 29.38, p < .001, η_p^2 = .01), this was qualified by a significant interaction (F = 70.49, p < .001, η_p^2 = .02; figure 2). Planned contrasts again showed, with reference to Facebook, usage of "you" was significantly greater in response to *sharing* (vs. *targeting*; F = 189.60, p < .001, η_p^2 = .05; $M_{Sharing}$ = 3.06, $SD_{Sharing}$ = 4.60, $M_{Targeting}$ = 0.74, $SD_{Targeting}$ = 2.25), whereas no difference emerged with reference to Amazon (p = .54, $M_{Sharing}$ = 0.74, $SD_{Sharing}$ = 2.18, $M_{Targeting}$ = 0.65, $SD_{Targeting}$ = 2.14; see figure 3).

3.3.3. Discussion. Consistent with our theorizing, study 3 findings suggest that consumers use control-signaling pronouns in a similar manner to brand switching. Specifically, "we" and "you" were used much more frequently when discussing Facebook's *sharing* (vs. *targeting*) actions. The same effect, however, is not observed when discussing Amazon, even though its actions are no less intrusive than Facebook's. Our theorizing and previous findings suggest that this is because Facebook and Amazon, while equally intrusive with consumers' data, operate in industries with dramatically different levels of variability in terms of privacy practice.

One potential alternative explanation for this pattern is that the pronoun "you" was more popular with Facebook due to the highly social aspect of the platform. Indeed, Packard and Berger (2020) found that songs with "you" are much more popular because "you" directly signals attentional focus and inspires other-activation. Yet the authors did not find the same effect on usage of "we." In contrast, this quasi-experiment found consistent findings for both "we" and "you" usage and therefore, in conjunction with our prior studies, provides robust support for our theorizing around consumers' usage of these pronouns as a means to reclaim control.

4. General Discussion

Privacy is more prescient to marketers and consumers than ever before (K. D. Martin et al., 2017). Yet, the literature provides little clarity around privacy violation as a construct (K. D. Martin & Murphy, 2017) and consumers' paradoxical responses to privacy violations (Krasnova, Spiekermann, Koroleva, & Hildebrand, 2010). The current research aimed to shed light on these overlapping areas. Study 1, using a multi-method approach, suggested three dimensions of privacy violation (*recording*, *targeting*, and *sharing*) that progressively increase in severity. Results of two follow-up studies converged to show that consumers are more likely to engage in control-reclaiming behaviors following a *sharing* (vs. *targeting*) privacy violation when the company operates in an industry with high (but not low) variability in privacy practices. In Study 2, this pattern was shown in a controlled between-participants experiment using fictitious companies. Study 3 demonstrated the effect using text analysis in a naturalistic setting.

Our findings contribute conceptual clarity to the privacy literature, suggesting that privacy violation is primarily a matter of what companies do with consumer data rather than the type or quantity of data collected. In particular, we disentangle the nuanced difference between *targeting* and *sharing*, two actions traditionally muddled as "personalization" in the literature (Vesanen & Raulas, 2006). Our research establishes that these company actions have direct consequences on perceived privacy violation and control-reclaiming behaviors.

Furthermore, this research draws from theories of self-determination (Ryan & Deci, 2000) and coping (Hong, Chiu, Dweck, Lin, & Wan, 1999) to help explain consumers' conflicting responses to privacy violations. Our findings show an industry-level characteristic (i.e., variability in privacy practice) that moderates the relationship between privacy-violating company action and consumer response, providing a theory-driven explanation for this market phenomenon.

Pragmatically, our research shows marketers that not all privacy violations are created equal, and so asking consumers for ubiquitous privacy consent might instigate backlash (e.g., through brand switching) as they later discover intrusive company actions. Of course, certain information (e.g., crash reports) is needed to improve consumer experience. Based on our findings, companies might take a more balanced approach whereby consent is requested separately for each of three types of data usage (*recording*, *targeting*, and *sharing*). Ideally, this would enhance transparency of privacy practices and consumers' perceived control.

In addition, past research has shown that while privacy is often seen by marketers as a burden or cost that could hardly be used strategically (K. D. Martin & Murphy, 2017). However, our findings challenge this assumption. Especially within industries with standardized data privacy practice (e.g., banking, online retail, search engines), companies could use a

differentiated privacy practice as a unique selling point. Indeed, DuckDuckGo—a privacy-protecting search engine platform—has recently gained traction in the otherwise homogenous search engine industry (Lomas, 2019).

5. Limitations and Future Research

Although we have attempted to maximize the rigor of our research, several limitations exist that provide opportunities for future research. First, we hypothesized that consumers' responses to privacy violations are a function of consumers' need to reclaim lost control over their information. However, we have not been able to test this mediation process in this empirical package, which we hope to address in advancing this research. Second, the recent introductions of new privacy laws (e.g., GDPR or CCPA) provide an opportunity for consumers to feel in control of their personal data and identity in the digital environment, which might produce different behavioral responses to privacy violations. Since current work in this legal domain often focuses on institutional and/or economics' consequences (Amjad & Murillo, 2020; Layton & Elaluf-Calderwood, 2019), future research should investigate this critical topic in the domain of user/ consumer privacy. Third, this research primarily focused on the US context, while privacy norms could vary significantly across cultures. The cross-cultural privacy behaviors are a topic worthy of future work. Finally, in this work we identify three violation types and while keeping the information quantity and quality constant in assessing a dimension of privacy violations. Future work could attempt to disentangle the interactions among these different dimensions (e.g., sharing of non-sensitive information vs. recording of sensitive data) in further clarifying the construct of privacy violation.

6. References

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2020). 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.

Acquisti, A., Taylor, C., & Wagman, L. (2016). The Economics of Privacy. *Journal of Economic Literature*, *54*(2), 442–492.

Amjad, O., & Murillo, D. (2020). Setting the Expectations Right: Reassessing the Power of GDPR in Protecting Online Users' Privacy. In *Academy of Management Proceedings* (Vol. 2020, p. 21204). Academy of Management Briarcliff Manor, NY 10510.

Bamberger, K. A., & Mulligan, D. K. (2010). Privacy on the Books and on the Ground. *Stanford Law Review*, *63*, 247–316.

Baumeister, R. F., Zhang, L., & Vohs, K. D. (2004). Gossip as Cultural Learning. *Review of General Psychology*, *8*(2), 111–121.

Benes, R. (2018, April 10). People Believe Ads Are Becoming More Intrusive. *EMarketer*.

Berger, J., Humphreys, A., Ludwig, S., Moe, W. W., Netzer, O., & Schweidel, D. A. (2020). Uniting the Tribes: Using Text for Marketing Insight. *Journal of Marketing*, *84*(1), 1–25.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Brandeis, S. D., & Warren, L. D. (1890). The Right to Privacy. *Harvard Law Review*, *4*(5), 193–220.

Chung, T. S., Wedel, M., & Rust, R. T. (2016). Adaptive Personalization Using Social Networks. *Journal of the Academy of Marketing Science*, *44*(1), 66–87.

Dweck, C. S. (2000). *Self-theories: Their Role in Motivation, Personality, and Development*. United Kingdom: Psychology Press.

Foster, E. K. (2004). Research on Gossip: Taxonomy, Methods, and Future Directions. *Review of General Psychology*, *8*(2), 78–99.

Goodwin, C. (1991). Privacy: Recognition of a Consumer Right. *Journal of Public Policy & Marketing*, *10*(1), 149–166.

Hajli, N., & Lin, X. (2016). Exploring the Security of Information Sharing on Social Networking Sites: The Role of Perceived Control of Information. *Journal of Business Ethics*, *133*(1), 111–123.

Hong, Y., Chiu, C., Dweck, C. S., Lin, D. M.-S., & Wan, W. (1999). Implicit theories, attributions, and coping: a meaning system approach. *Journal of Personality and Social Psychology*, *77*(3), 588.

Humphreys, A., & Wang, R. J.-H. (2018). Automated Text Analysis for Consumer Research. *Journal of Consumer Research*, *44*(6), 1274–1306.

Inesi, M. E., Botti, S., Dubois, D., Rucker, D. D., & Galinsky, A. D. (2011). Power and choice: Their dynamic interplay in quenching the thirst for personal control. *Psychological Science*, *22*(8), 1042–1048.

Kassarjian, H. H. (1977). Content analysis in consumer research. *Journal of Consumer Research*, *4*(1), 8–18.

Kazienko, P., & Adamski, M. (2007). AdROSA—Adaptive Personalization of Web Advertising. *Information Sciences*, *177*(11), 2269–2295.

Kim, T., Barasz, K., & John, L. K. (2018). Why Am I Seeing This Ad? The Effect of Ad Transparency on Ad Effectiveness. *Journal of Consumer Research*, *45*(5), 906–932.

Kirkpatrick, M. (2010, January 10). Facebook's Zuckerberg Says The Age of Privacy Is Over. *The New York Times*.

Krasnova, H., Spiekermann, S., Koroleva, K., & Hildebrand, T. (2010). Online Social Networks: Why We Disclose. *Journal of Information Technology*, *25*(2), 109–125.

Krishna, A. (2020). Privacy is a Concern: An Introduction to the Dialogue on Privacy. *Journal of Consumer Psychology*, *30*(4), 733–735.

Kurland, N. B., & Pelled, L. H. (2000). Passing the Word: Toward a Model of Gossip and Power in the Workplace. *Academy of Management Review*, *25*(2), 428–438.

Layton, R., & Elaluf-Calderwood, S. (2019). A social economic analysis of the impact of GDPR on security and privacy practices. In *2019 12th CMI Conference on Cybersecurity and Privacy (CMI)* (pp. 1–6). IEEE.

Lomas, N. (2019, March 13). Google has Quietly Added DuckDuckGo as a Search Engine Option for Chrome Users in ~60 Markets. *Tech Crunch*.

Lynskey, D. (2019, October 9). "Alexa, Are You Invading My Privacy?" – The Dark Side of Our Voice Assistants. *The Guardian*.

MacMillan, D., & McMillan, R. (2018, October 8). Google Exposed User Data, Feared Repercussions of Disclosing to Public. *Wall Street Journal*.

Martin, K. D., Borah, A., & Palmatier, R. W. (2017). Data Privacy: Effects on Customer and Firm Performance. *Journal of Marketing*, *81*(1), 36–58. h

Martin, K. D., & Murphy, P. E. (2017). The Role of Data Privacy in Marketing. *Journal of the Academy of Marketing Science*, *45*(2), 135–155. https://doi.org/10.1007/s11747-016-0495-4

Martin, K. E. (2014). Privacy Notices as Tabula Rasa: An Empirical Investigation into How Complying with a Privacy Notice is Related to Meeting Privacy Expectations Online. *Ssrn*, *34*(2), 210–227.

Mulligan, D. K., Koopman, C., & Doty, N. (2016). Privacy is an Essentially Contested Concept: A Multi-dimensional Analytic for Mapping Privacy. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2083).

Nill, A., & Aalberts, R. J. (2014). Legal and Ethical Challenges of Online Behavioral Targeting in Advertising. *Journal of Current Issues and Research in Advertising*, *35*(2), 126–146.

Nivea, R. (2019, April 30). At F8, Zuckerberg Unveils Facebook's New Mantra: "The Future Is Private." *CNET*.

Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The Privacy Paradox: Personal Information Disclosure Intentions Versus Behaviors. *Journal of Consumer Affairs*, *41*(1), 100–126.

Packard, G., & Berger, J. (2020). Thinking of You: How Second-Person Pronouns Shape Cultural Success. *Psychological Science*, *31*(4), 397–407.

Paul, K. (2020, February 3). They Know Us Better Than We Know Ourselves': How Amazon Tracked My Last Two Years of Reading. *The Guardian*.

Pew. (2019). *Americans and Privacy: Concerned, Confused and Feeling Lack Of Control over Their Personal Information*. *Pew Research Center* (Vol. November). Washington, DC.

PWC. (2017). Consumer Intelligence Series: Protect Me, (September), 4.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68.

Sahni, N. S., Wheeler, S. C., & Chintagunta, P. (2018). Personalization in email marketing: The role of noninformative advertising content. *Marketing Science*, *37*(2), 236–258.

Schiele, K., & Venkatesh, A. (2016). Regaining Control through Reclamation: How Consumption Subcultures Preserve Meaning and Group Identity after Commodification. *Consumption Markets & Culture*, *19*(5), 427–450.

Schmitt, B. (2019). From atoms to bits and back: A research curation on digital technology and agenda for future research. *Journal of Consumer Research*, *46*(4), 825–832.

Sedek, G., Kofta, M., & Tyszka, T. (1993). Effects of Uncontrollability on Subsequent Decision Making: Testing the Cognitive Exhaustion Hypothesis. *Journal of Personality and Social Psychology*, *65*(6), 1270–1281.

Shane, D. (2018, September 28). Research Shows Users Are Leaving Facebook in Droves. Here's What It Means For You. *INC.Com*.

Sheehan, K. B., & Hoy, M. G. (1999). Using E-mail to Survey Internet Users in the United States: Methodology and Assessment. *Journal of Computer-Mediated Communication*, *4*(3), JCMC435.

Skinner, E. A. (1996). A Guide to Constructs of Control. *Journal of Personality and Social Psychology*, *71*(3), 549–570.

Skinner, E. A., & Zimmer-Gembeck, M. J. (2011). Perceived Control and the Development of Coping. In S. Folkman (Ed.), *The Oxford Handbook of Stress, Health, and Coping* (pp. 35–59). Oxford University Press.

Smith, H. J., Dinev, T., & Xu, H. (2011). Information Privacy Research: An Interdisciplinary Review. *MIS Quarterly*, *35*(4), 989–1015.

Summers, C. A., Smith, R. W., & Reczek, R. W. (2016). An Audience of One: Behaviorally Targeted Ads as Implied Social Labels. *Journal of Consumer Research*, *43*(1), 156–178.

Thimm, C., Rademacher, U., & Kruse, L. (1995). " Power-Related Talk" Control in Verbal Interaction. *Journal of Language and Social Psychology*, *14*(4), 382–407.

Turow, J., Hennessy, M., & Draper, N. (2016). *The Tradeoff Fallacy*. University of Pennsylivania-Annenberg School for Communication. https://doi.org/10.2139/ssrn.2820060

Vesanen, J., & Raulas, M. (2006). Building Bridges for Personalization: A Process Model for Marketing. *Journal of Interactive Marketing*, *20*(1), 5–20.

Wert, S. R., & Salovey, P. (2004). A Social Comparison Account of Gossip. *Review of General Psychology*, *8*(2), 122–137.

Whitson, J. A., & Galinsky, A. D. (2008). Lacking Control Increases Illusory Pattern Perception. *Science*, *322*(5898), 115–117.

Wuerthele, M. (2019, March 14). "Privacy. That's iPhone" ad Campaign Launches, Highlights Apple's Stance on User Protection. *AppleInsider*.