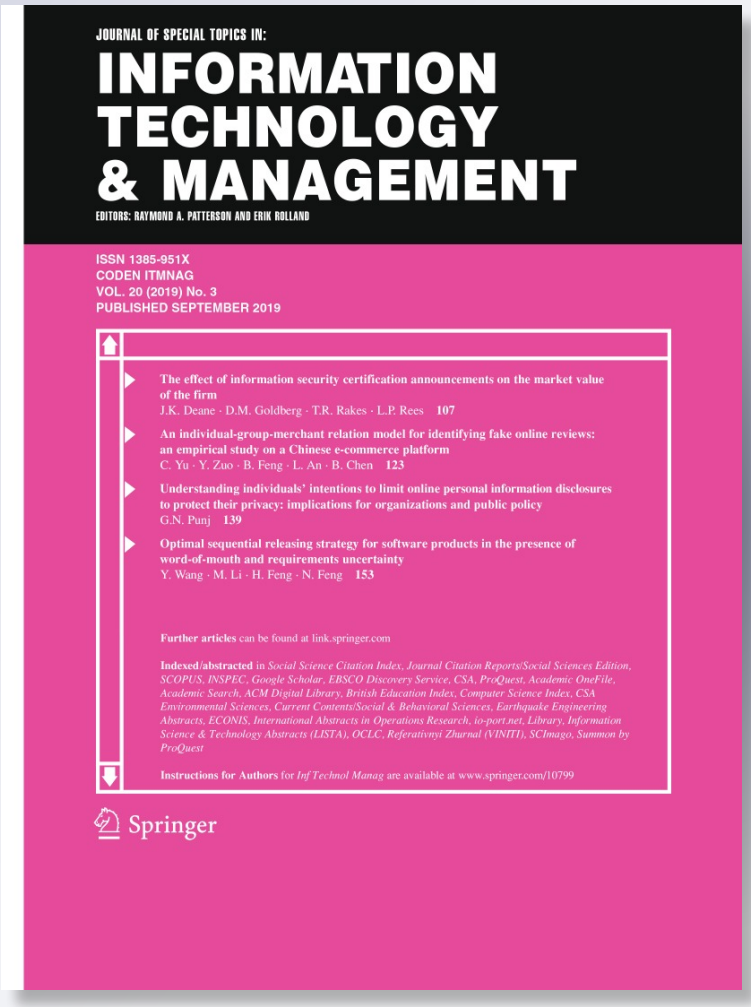
Understanding individuals' intentions to limit online personal information disclosures to protect their privacy: implications for organizations and public policy

Girish N. Punj

Information Technology and Management

ISSN 1385-951X
Volume 20
Number 3

Inf Technol Manag (2019) 20:139-151
DOI 10.1007/s10799-018-0295-2



Springer

Information Technology and Management (2019) 20:139–151
https://doi.org/10.1007/s10799-018-0295-2

Understanding individuals' intentions to limit online personal information disclosures to protect their privacy: implications for organizations and public policy

Girish N. Punj[1]

Published online: 7 December 2018

Abstract
The purpose of the present research is to understand individuals' intentions to limit their personal information online to partially anonymize their digital identity. Key concepts from several privacy theories are used to generate hypotheses that can be used to understand the behavior of interest. Data from a national probability sample of 792 adults is used to test the hypotheses. The results indicate that the size of an individual's digital footprint, their need for control over personal information, and past privacy violations are important determinants of their online information limiting behavior. The findings have important implications for theory and practice. From a theoretical perspective, the findings indicate that individual intentions to limit personal information online seem to be based on a desire to balance their current online exposure with their need to control their personal online information. Past privacy violations also exert an influence on online information disclosures. The research is important for organizations and policy makers in designing privacy policies and proposing regulation that recognizes the dilemma that individuals encounter when they share information online with an organization for mutual benefit.

Keywords Privacy · Information limiting behavior · Anonymity · Digital identity · Public policy · Information protection · Regulation

1 Introduction

A 2013 report by the non-profit Pew Research Center found that 50% of individuals are concerned about the amount of information about them that was available online and the size of their "digital footprint" in cyberspace [47]. The personal data collected by organizations is typically with the permission of individuals who provide their consent by signing-off on the privacy policies of firms. But research shows that an estimated 95% of individuals do not bother reading the fine print in these policies [28]. Individuals disclose their personal information online to organizations in an effort to build or enhance their relationship with the organization for mutual benefit, but at the same time they are concerned about their privacy.

A strategy increasingly being used by people is to limit their online information disclosure to partially anonymize their identity to protect sensitive information, while sharing non-sensitive data for developing affinity with the organization. The Pew Research Center report cited earlier estimates that a startling 86% of people have engaged in this behavior. The trend toward striking a balance between online information disclosure and privacy protection is also prevalent on social media where individuals construct virtual personas that only bear a partial resemblance to their real-life identities [65]. By divulging large amounts of personal information on social media, people unwittingly create "digital skeletons in the closet" that can exist into perpetuity [1] and risk serious privacy violations.

When individuals share their personal information with organizations, they face the risk of a privacy violation due to the possibility of the shared information being misused by the company, or more often by a third party with harmful intent. Recent events reported in the press, such as the Facebook—Cambridge Analytica data-sharing scandal have only heightened people's concerns about disclosing their personal information online [52]. The limited disclosure or partial

✉ Girish N. Punj
Girish.Punj@business.uconn.edu

[1] Department of Marketing, School of Business, University of Connecticut, 2100 Hillside Road, Storrs, CT 06269-9013, USA

anonymization of personal information online provides a mechanism by which an individual can share information with an organization, while also reducing the risk of a privacy violation. It balances two conflicting goals relating to privacy protection and information disclosure and is an unintended consequence of an individual's apprehension about their personal information being misused or stolen and the privacy policies of firms.

2 Research purpose

Despite the research on online information protection that has been reported in the literature [27, 38, 46, 55, 61], surprisingly few studies have examined why and how individuals seek to limit their personal information online. Research by Lwin and her colleagues found that consumers may falsify their personal information online as a defensive reaction to corporate overreach [32] and also when they do not perceive a moral obligation to provide accurate information [31]. But not much is known about what leads to limiting personal information disclosure as opposed to outright falsification.

Hence, the purpose of the present research is to understand individuals' intentions to limit their personal information online disclosures to partially anonymize their digital identity. Is a person's decision to limit their online information primarily based on the desire to balance their current online exposure with their need to control their personal online information? Or is it mainly influenced by past privacy experiences? Or a combination of both factors?

The research is important for organizations because under the current environment organizations are expected to exercise restraint and self-regulate themselves in safeguarding personal data [3], while individuals are expected to be self-manage their online identity by empowering themselves with information [37, 65]. The research is also important for policy makers as they contemplate potential regulatory approaches to safeguarding the online personal information of individuals in light of recent data breaches that have been reported in the media [33]. Most of the time the collection and analysis of online information is invisible to individuals and ironically that may also be the reason why people are concerned about how their personal information is being used and with whom it is being shared.

3 Literature review

Privacy studies have typically investigated the role "privacy concerns" play in individual decisions to disclose or protect personal information [35, 54]. Several mechanisms have been used to link the antecedents of privacy concerns to its consequents (see literature reviews by Acquisti et al. [1], Pavlou [42] and Smith et al. [55]). Primary among these is the "privacy calculus" where individuals are assumed to make a trade-off between the risk of a privacy loss and benefit of information disclosure [4].

The benefits typically relate to receiving messages (e.g., promotional offers, communications that promote a common cause), while the privacy risk relates to a loss of anonymity and potential privacy violations [7]. Recent research indicates that individuals find it difficult to make the trade-off needed to adequately self-manage privacy because they tend to under-estimate the risks of privacy loss and over-estimate the benefit of revealing personal information online [1].

Consequently, individuals may *limit* their online information disclosures so that the risk of a privacy loss is mitigated, while also attaining a better balance with the benefits of information disclosure as per the privacy calculus [9]. *Online information limiting behavior* will likely be influenced by at least three factors. First, it will depend on the amount of personal online information that already exists in cyberspace. In other words, the individual's current online exposure or *size of digital footprint* is likely to be an important influence. The essential idea behind the concept relates to the digital trace that individuals knowingly leave or unwittingly share with organizations and entities as a result of their online interactions. The presumption is that their future information disclosure behavior is influenced by their current online exposure. In other words, the more information there is out there about them in cyberspace, the less they care about disclosing more. Hence, people who already have a sizable digital footprint are likely to be de-sensitized to the risks of additional online information disclosures [50, 57]. The concept of a digital footprint on online behavior has not previously been examined in a privacy context, although it has been studied in other contexts [15, 26, 39].

Second, the individual's *need for control* over their personal information is also likely to have an effect. In fact, the desire to control information is often viewed as essential to the conceptualization of privacy [46, 55]. Individuals exercising a greater need for control over their personal information are more likely to limit their online information disclosures. Third, the individual's desire to limit their online information disclosures is likely to be based on their *past privacy violations*. People who have experienced a privacy loss or violation in the past can be expected to be sensitized to the possibility of another one occurring in the future. Hence, they are likely to limit their online information disclosures.

Based on the research purpose and a review of the associated literature, the following constructs were selected to examine the behavior of interest. *Online Information Limiting Behavior* was defined as behaviors by individuals that were intended to limit or partially anonymize their personal

online information. *Size of Digital Footprint* was defined as the amount of personal information currently available online for the individual, of which the individual was aware. *Need for Control* was defined as the individual's need for control over their personal information online. *Past Privacy Violations* were defined as privacy violations the individual had previously experienced.

4 Hypothesis development

The present research selectively draws on concepts from the vast literature on "privacy concerns" to obtain an understanding of the influences that lead individuals to limit their personal information to partially anonymize their digital identity. Previous research has found that individuals possess dual tendencies toward information protection and information disclosure [36], which can be exhibit independently of each other [25]. In other words, individuals may engage in both behaviors but to varying degrees. Most individuals seek to strike a balance between these two behavioral tendencies by assessing the risks and benefits associated with them [19]. Yet, there is evidence that seems to indicate that individuals find it difficult to weigh the privacy risk of information disclosure against the benefits of information disclosure. Hence, their decision to limit their online personal information disclosures to partially anonymize their digital identity may be partly based on current online behavior and partly on past privacy experiences [43], which is used as the basis for hypotheses development.

When individuals disclose personal information to organizations, they may be doing so as part of a privacy calculus where they expect to receive benefits in exchange for these disclosures [9]. The benefits typically relate to receiving messages (e.g., promotional offers, communications that promote a common cause), while the privacy risk relates to a loss of anonymity and potential privacy violations [7]. Individuals with a significant amount of online exposure seem to value the benefits of information disclosure to a greater extent than the risk of such disclosure, because they trust the company to safeguard their private information [9]. Research indicates that become less concerned about disclosing personal information when they establish trust with a company [23].

Further, it is probable that individuals who already have significant amount of online information exposure (i.e., maintain a large "digital footprint") have been desensitized to the risks of information disclosure due to the phenomena of psychological habituation [50]. The processes of desensitization and habituation have been examined both as a general-process theory of motivation [34] and as well as an associative conditioning model [12]. The underlying premise is that individuals can become less concerned about defending their privacy [57] through these processes. Harris, Brookshire and Chin [17] found that de-sensitization was positively related to trust and negatively related to risk while examining consumer intentions to install a mobile app. Similarly, Romanosky et al. [51] found that that too many unnecessary notifications of data breaches desensitized individuals to the risk of identity theft.

H1 Individuals with a larger digital footprint are less likely to limit their personal information online.

Based on the precepts of protection motivation theory [36, 64], past privacy violations can be expected to reduce the coping efficacy of individuals to deal with similar threats in the future. Hence, they are likely to limit their personal information online to guard against re-occurrences.

Also, individuals are known to create a virtual information space and seek to contain their personal information within its boundaries [44, 56, 58]. Attempts to cross these boundaries is viewed as an invasive act. Individuals who have experienced a privacy violation may consider that their information space has been violated. They are then likely to guard against future intrusions and resort to tactics such as limiting their online information [8, 62].

Past research has also found that when a digital identity has been previously compromised, individuals begin to anonymize their online personas [65]. Lwin et al. [32] found that individuals are likely to fabricate their personal information online when their concerns are heightened due to weak privacy policies of companies or a lack of adequate public policy regulation. The opportunity to remain anonymous in online settings enhances intentions to limit personal information online [31].

H2 Individuals with past privacy violations are more likely to limit their personal information online.

The need to protect one's personal information is central to the conceptualization of privacy [22, 35, 46]. In fact, most definitions of privacy relate the term to the ability to control personal information [45, 55].

Thus, the need for control may be a key factor in the individual's decision to partially anonymize their digital identity [11]. Sheehan and Hoy [54] identify two aspects of control, with the first relating to the individual's awareness that their personal information is being collected and the second concerning how that information is used. Individuals will seek to limit their online exposure if they have concerns about how their personal information is used [11, 35, 45].

According to the theory of planned behavior (TPB), it is likely that beliefs that are directly related to the anticipated behavior [48, 63] also have an influence on the current online exposure. Thus, an individual's need for control

over their personal information will lead them to reduce the size of their digital footprint. When individuals believe that they lack such control they will limit their personal information online. They may also believe that they have a "right to be forgotten" and hence will take the steps necessary in the pursuit of that right [40]. Zwick and Dholakia [65] identify three tactics individuals use to control their personal information online. These include anonymity, secrecy and confidentiality based on the amount of personal information that has already been externalized and the accuracy of it.

H3 Individuals with greater need for control over their personal information online will have a smaller digital footprint.

H4 Individuals with a greater need for control are more likely to limit their online personal information.

Individual privacy concerns have been found to be related to individual differences such as income, age, education and gender [4, 8, 21, 29, 38]. Hence, demographic characteristics are likely to be also related to the individual decision to limit personal information online. For instance, higher-income individuals may exhibit a tendency toward information protection, because even though they might value the benefits of information disclosure, the monetary and opportunity costs of a privacy violation (e.g., identity theft) are greater for them. Similarly, individuals with more education may also have a tendency toward information protection, because they are more likely to have the cyber fluency (i.e., web expertise) to calibrate and understand privacy risk.

With regards to gender, women are more concerned about their online privacy [53] but are known to value the social aspect of information [60] and participate in virtual communities to a greater degree [14]. But, they have also been observed to be more protective of their personal information online than men [21]. There are also likely to be important generational differences in online privacy behavior [21]. Younger individuals (e.g., Millennials) are less likely to limit or anonymize personal information online because they have been de-sensitized to the risks of a privacy loss due to their ubiquitous use of social media.

H5 The relationship between demographic characteristics and the limiting of online personal information will be positive for (a) income, (b) education, (c) age and (d) gender (female).

Taken together, the hypotheses can be used to propose a framework for understanding how individuals self-manage their digital identity. The framework proposes that the decision to limit personal information online is primarily influenced by the three constructs discussed earlier, namely, size of digital footprint, need for control, past privacy violations, and demographic characteristics as shown in Fig. 1.

5 Research method

5.1 Data collection

Data from a national probability sample of 792 adult internet and smartphone users, 18 years and older, living in the continental United States were used to test the hypotheses. The data were gathered through a telephone survey conducted in July 2013 by Princeton Associates on behalf of the Pew Research Center's Internet and American Life Project [47]. The survey data were collected using a dual-frame sample design. Both landline and cellular random digit dial (RDD) samples were used.

The landline sample was a list-assisted random digit sample of telephone numbers selected from landline telephone exchanges in the continental United States. The cell phone sample (including those without a landline phone) was drawn from dedicated cellular exchanges based on the most recently available Terminating Point Master (TPM) data file for the continental United States. The combined sample generalizes to the American population of adult internet and smartphone users, with a margin of sampling error of ± 3.8 percentage points, as computed by statisticians of the sponsoring organization [47]. The non-profit sponsoring organization is an authoritative source of information on how Americans use the internet and the data provided by it is often used by federal agencies in formulating government policy.

The questionnaire was administered using professionally trained and experienced personal interviewers. Information on the constructs in the study was gathered through both pre-coded and open-ended responses. The likelihood that respondents made recall errors was minimized by asking respondents to report online privacy behavior and activities in which they had recently engaged. Based on comparisons between the current sample and other samples on identical survey questions that are part of the annual longitudinal privacy surveys conducted by Pew Internet, no evidence of any systematic bias could be detected. But that does not mean there are no non-sampling errors in the data file. Yet, because of the use of a highly-regarded survey organization and professional interviewers trained to probe respondent behavior, it was felt that the data were of sufficiently high quality to merit their use in testing the hypotheses. The questions in the survey instrument that provided information on the four primary constructs of interest in this study are listed in the Appendix, and are a part of the Pew Research Center's 2013 report titled "Anonymity, Privacy, and Security Online."

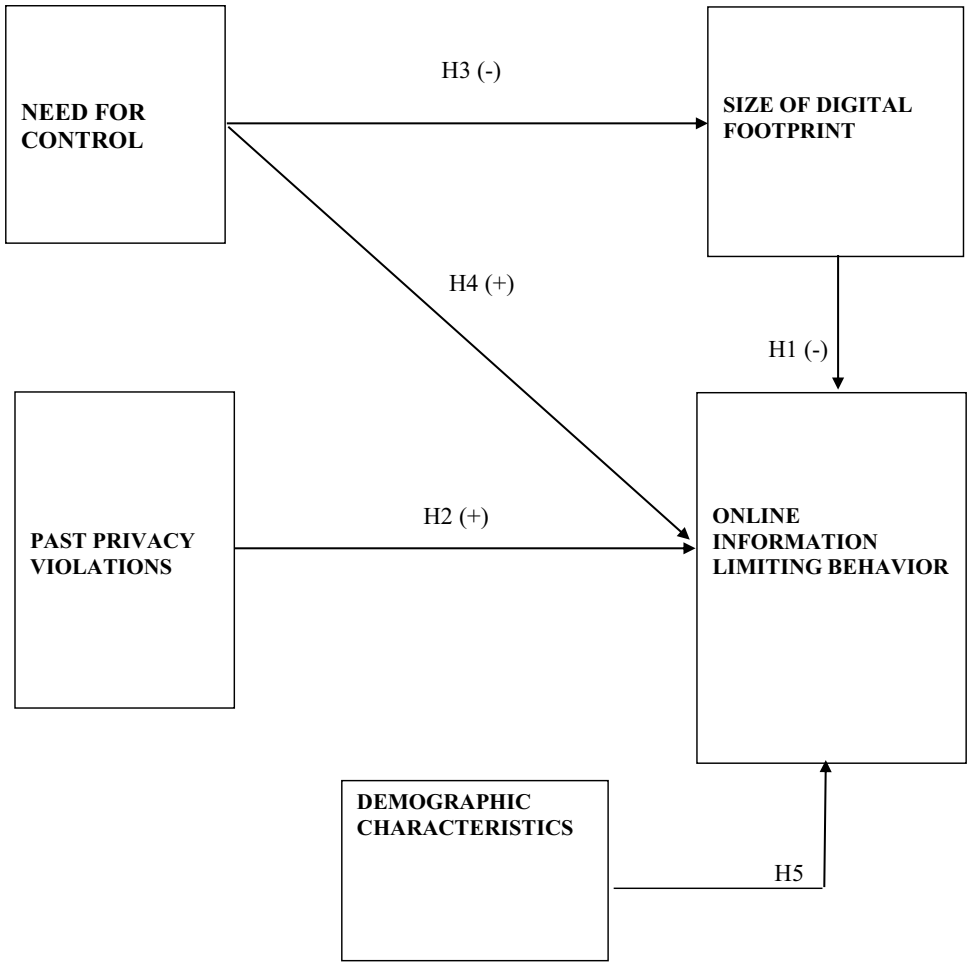


Fig. 1 Proposed model of online information limiting behavior

5.2 Measures

Partial Least Squares—Structural Equation Modeling (PLS—SEM) was used to identify *formative* measures corresponding to the four latent factors in the proposed model (see Fig. 1) from the multiple items available in the dataset. The logic for using formative as (opposed to reflective) measures for measuring the unobserved constructs was that each measure only captured a particular aspect of the latent factor's domain, and that a combination of the selected measures when considered together best summarized the meaning of the construct [10]. In order to select the *formative* measures from those available in the dataset through the survey questions, the procedure(s) recommended by Ringle et al. [49] and Lowry and Gaskin [30] were adopted. These included selecting choosing only items from those available based on statistical significance ($p < .05$), collinearity statistics (VIF < 3.0), and researcher judgment of the *content validity* of the selected items [10]. The final set of items used to measure the study constructs are described below, and depicted in Table 1.

5.2.1 Online information limiting behavior

The construct was operationalized using answers provided by the respondent to different dichotomous scale (1 = yes; 0 = no) items that measured the steps taken by the respondent to protect their personal information. The five items selected were "used a temporary username," "gave inaccurate information," "cleared cookies and browser history," "deleted previously posted information," and "sought removal of posted information." Taken together, the items sought to encapsulate the extent to which respondents had engaged in behavior designed to limit their personal information online, as depicted in Table 1.

5.2.2 Size of digital footprint

The construct was operationalized using answers provided by the respondent to different dichotomous scale (1 = yes; 0 = no) items. The six items represented personal information on respondent that he/she believed was already available about them online, such as their "home phone number,"

Table 1 Descriptive information for online privacy variables

	Frequency	Percent
Online information limiting behavior		
Cleared cookies and browser history	496	(63)
Deleted previously posted information	287	(36)
Used a temporary username	190	(24)
Sought removal of posted information	131	(17)
Gave inaccurate information	94	(12)
Size of digital footprint (items available online)		
Video	574	(73)
Group affiliations	446	(56)
Home phone number	436	(55)
Personally identifying information	385	(49)
Birthdate	295	(37)
Photo	256	(32)
Need for control (over listed items)		
Content of email	534	(67)
Files downloaded	434	(55)
Content of online chat	389	(49)
Applications and programs used	327	(41)
Past privacy violations		
Email account hacked	147	(19)
Personal information stolen	80	(10)
Personal information misused	77	(10)
Stalked or harassed online	66	(8)

Entries are number (percentage) of "yes" responses to row items

"personally identifying information," "photo," "video," "group affiliations," and "birthdate." Collectively, the items attempted to depict the amount of personal information that the respondent believed that was available about them online, with the caveat that personal information online of which the individual was unaware was not captured by the construct.

5.2.3 Need for control

The construct was operationalized using a dichotomous scale that assessed the importance (1 = important; 0 = not important) of the respondent's need for control over different items relating to their personal information online. The four items chosen were "files downloaded," "applications and programs used," "content of email," and "content of online chat." When considered together, the items denoted the degree of control the individual sought over their personal information online.

5.2.4 Past privacy violations

The construct was operationalized using answers provided by the respondent to different dichotomous scale (1 = yes; 0 = no) items that sought to measure the degree to which respondents had had their privacy compromised. The four items selected asked respondents whether they had their "personal information stolen, "email account hacked," "personal information misused," and been "stalked or harassed online." Taken together, the items sought to depict the extent to the individual had experienced privacy violations in the past.

5.2.5 Demographic characteristics

Information on the demographic characteristics of individuals such as income, education, age and gender was also directly measured. *Income* was measured as the total household income from all sources before taxes in 2012. A six-point ordinal scale derived from cut-offs used by the US Census Bureau (1 = less than $20,000; 2 = $20,000–$39,999; 3 = $40,000–$74,499; 4 = $75,000–$99,999; 5 = $100,000–$149,999; 6 = more than $150,000) was used. *Age* was measured using a six-point ordinal scale that used break-points in chronological age that are normally used by demographers (1 = 18–24 years; 2 = 25–34 years; 3 = 35–44 years; 4 = 45–54 years; 5 = 55–64 years; 6 = 65+ years). *Education* was measured using a five-point ordinal scale (1 = less than high school; 2 = high school graduate; 3 = some college or two year associate degree; 4 = four year college graduate; 5 = postgraduate or professional degree). *Gender* was recorded on a dichotomous scale (1 = male; 2 = female) scale by the phone interviewer.

The modal categories for *Income* and *Education* were $40,000–$74,499 of annual household income, and some college or vocational school, respectively. The modal category for *Age* was 55–64 years. For *Gender*, the sample was almost evenly split between males (51%) and females (49%). Overall, the sample distributions on the study variables closely matched the demographic profile of the American population of internet users, as expected, due to the use of a national sample frame and probability sampling. Descriptive statistics on the demographic variables are reported in Table 2.

6 Results

6.1 Measurement model

SmartPLS3 structural equation modeling software was used to estimate weights for the four latent factors in the proposed model, namely, *Size of Digital Footprint*, *Need for Control*, *Past Privacy Violations* and *Online Information Limiting Behavior*. In order to estimate the significance of the estimated weights the bootstrapping algorithm (n = 2000) in the software was used. With only one exception, all estimated

Table 2 Descriptive information for demographic characteristics

	Frequency	Percent
Income		
Less than $20,000	117	(14.8)
$20,000–$39,999	142	(17.9)
$40,000–$74,999	196	(24.7)
$75,000–$99,999	90	(11.4)
$100,000–$149,999	105	(13.3)
$150,000 or more	57	(7.2)
Missing	85	(10.7)[a]
Education		
High school incomplete	30	(3.8)
High school graduate	182	(23.0)
Some college, no degree	247	(31.2)
Four year college degree	181	(22.9)
Postgraduate or professional degree	150	(18.9)
Missing	2	(0.3)
Age		
18–24 years	97	(12.2)
25–34 years	109	(13.8)
35–44 years	100	(12.6)
45–54 years	134	(16.9)
55–64 years	184	(23.2)
65+ years	146	(18.4)
Missing	22	(2.8)
Gender		
Male	400	(50.5)
Female	392	(49.5)

[a]A classification and regression tree (CART) algorithm was used to predict and insert missing values for Income

Table 3 Measurement model: outer weights

	Outer weights	Significance
Online information limiting behavior		
Cleared cookies and browser history	0.31	$p < .01$
Deleted previously posted information	0.47	$p < .01$
Used a temporary username	0.25	$p < .01$
Sought removal of posted information	0.30	$p < .01$
Gave inaccurate information	0.24	$p < .01$
Size of digital footprint		
Video	0.31	$p < .01$
Group affiliations	0.28	$p < .01$
Home phone number	0.11	$p < .05$
Personally identifying information	0.28	$p < .01$
Birthdate	0.24	$p < .01$
Photo	0.37	$p < .01$
Need for control		
Content of email	0.43	$p < .01$
Files downloaded	0.37	$p < .01$
Content of online chat	0.54	$p < .01$
Applications and programs used	−0.21	$p < .05$
Past privacy violations		
Email account hacked	0.45	$p < .01$
Personal information stolen	0.23	$p < .01$
Personal information misused	0.48	$p < .01$
Stalked or harassed online	0.42	$p < .01$

weights were positive and significant at the $p < .01$ level, indicating a very strong fit to the measurement model, as reported in Table 3. One of the estimated weights had an associated negative sign. Yet, it was retained as a formative measure for its corresponding construct because it was significant ($p < 05$), and the negative sign was most likely an artifact of the estimation algorithm. An examination of the collinearity statistics showed that the variance inflation factors (VIF's) for all measures were below the recommended thresholds (< 3) [16, 18].

6.2 Structural model

The proposed model of *Online Information Limiting Behavior* was also estimated using SmartPLS3 structural equation modeling software. The overall model $\chi^2 = 590.51$ was significant ($p < .01$). Model fit indices indicate the structural model also provided a strong fit to the data (SRMR = .07; NFI = 0.8) with both fit indices meeting the recommended thresholds, namely, (SRMR < .08) and (NFI > 0.8) [16, 18].

The estimated correlations among the unobserved constructs are depicted in Table 4.

The squared multiple correlations (adjusted R^2's) for the structural equations for *Size of Digital Footprint* and *Online Information Limiting Behavior* were 0.28 and .41 respectively. A bootstrapping algorithm (n = 2000) indicated that both these values were significant at the $p < .01$ level.

6.3 Hypothesis tests

The relationship between *Size of Digital Footprint* and *Online Information Limiting Behavior* was significant and in the predicted direction ($\beta = -0.28$; $p < .01$). Thus H1 was supported. Both *Past Privacy Violations* and *Need for*

Table 4 Correlations among online privacy constructs

Construct	SDF	PPC	NC	OILB
Size of digital footprint (SDF)	–			
Past privacy violations (PPC)	−0.31	–		
Need for control (NC)	−0.28	0.10	–	
Online information limiting behavior (OILB)	−0.50	0.47	0.24	–

Control were found to be positively related to *Online Information Limiting Behavior* as expected (β = 0.31; $p < 0.01$) and (β = 0.10; $p < 0.01$), respectively. Thus, H2 and H4 were also supported. *Need for Control* was found to be negatively related to *Size of Digital Footprint* as predicted (β = −0.29; $p < 0.01$). Thus H3 was also supported The results of all the hypothesis tests along with the standardized β estimates and significance levels are depicted in Table 5.

The magnitude of the standardized regression weights (β's) suggest that *Past Privacy Violations* are more important than both *Need for Control* and *Size of Digital Footprint* in influencing the individual's decision to limit their personal information online. Interestingly, it seems that *Past Privacy Violations* and *Size of Digital Footprint* have opposing but similar effects on online information limiting behavior. Thus, it appears that the two influences work in tandem in influencing the individual's decision to limit their personal information online. Further, it seems *Size of Digital Footprint* mediates the relationship between *Need for Control* and *Online Information Limiting Behavior.* At issue is whether *Size of Digital Footprint* fully mediates the relationship between *Need for Control* and *Online Information Limiting Behavior.* Such a possibility was tested using mediational analysis using the procedure recommended for PLS-SEM models [16, p. 239] which involved estimating both the direct effect of *Need for Control* on *Online Information Limiting* as well as the indirect effect through *Size of Digital Footprint*. Full mediation would be identified if the indirect effect was significant, but not the direct effect, while partial mediation would be revealed if both the direct and indirect effects were significant. The results of the mediational analyses showed that both the direct effect of *Need for Control* on *Online Information Limiting Behavior* (β = 0.10; $p < .01$) and the indirect effect through *Size of Digital Footprint* (β = 0.08; $p < .01$) were significant. Thus, partial mediation was confirmed as implied by the proposed model.

For the hypothesized relationships between demographic characteristics and *Online Information Limiting Behavior*, only *Age* and *Education* were found to be positively related to *Online Information Limiting Behavior* as predicted (β = 0.24; $p < .01$) and (β = 0.07; $p < .05$), respectively. Disappointingly, no relationships between *Income and Gender* with *Online Information Limiting Behavior* were found. Thus, H5(b) and H5(c) were supported, but not H5(a) and H5(d). To further examine the possible effects of the demographic characteristics on *Online Information Limiting Behavior,* multi-group moderation analysis was used for *Income, Education, Age and Gender,* both as individual constructs and in combination. Once again, only the effects of *Age* and *Education* were found to be significant. The demographic finding that emerges is that older adults with more education are more likely to limit their online information, while younger individuals with less education are less likely to do so.

Taken together, the results of the research indicate that individuals with a large digital footprint are less likely to limit or partially anonymize their personal information and those with past privacy violations are more likely to do so. The need for control over personal information influences online information behavior directly as well as indirectly, which indicates that it has dual influence on online information limiting behavior. The demographic effects confirm that the younger generations (e.g., millennials) are less likely to limit their personal information online in comparison to older generations (e.g., baby boomers), possibly because of their greater use of social media (e.g., Instagram).

Table 5 Structural model: path coefficients

	β	Significance
Constructs		
H1: Size of digital footprint → online limiting behavior	−0.28	$p < .01$
H2: Past privacy violations → online limiting behavior	0.31	$p < .01$
H3: Need for control → size of digital footprint	−0.29	$p < .01$
H4: Need for control → online limiting behavior	0.10	$p < .01$
Demographic factors		
H5 (a) Income → online limiting behavior	0.03	n.s.
H5 (b) Education → online limiting behavior	0.07	$p < .05$
H5 (c) Age → online limiting behavior	0.24	$p < .01$
H5 (d) Gender (female[a]) → online limiting behavior	0.03	n.s

[a]Gender (male) was used as the control category

Model fit

Adjusted R^2 (size of digital footprint) = 0.28

Adjusted R^2 (online limiting behavior) = 0.41

$\chi^2 = 590.51$; SRMR = .07; NFI = 0.80

6.4 Study limitations

The study is high in external validity because it is based on a nationally representative sample of American individuals with internet access in 2013. To achieve the high degree of external validity certain compromises had to be made during the data collection process. Some of the variables were measured using ordinal or dichotomous scales. While multiple indicators would have been preferred, the extent to which repeated measurements of the same underlying behavior might cause respondent fatigue and lead respondents to prematurely terminate the phone interview was a critical consideration in the study design.

But this limitation may be offset by the advantage of examining stated online information limiting behaviors as opposed to attitudes or behaviors that might have to be primed (i.e., simulated) in an experimental setting. Another limitation relates to the study population. The empirical results would likely be different if the study were conducted in a different geographic region (e.g., the E.U.) where the "right to be forgotten" [40] is an important influence on both public policy and the online data capture practices of companies.

7 Discussion and conclusion

The present study seeks to be among the first to formulate and test a conceptual model of individual intentions to *limit* online personal information disclosures in an effort to protect privacy. The research findings show that past privacy violations and the need for control are important influences on individual intentions to limit their personal information online, with the former effect dominating the latter. But, these influences that work together in limiting online information disclosures are offset by the magnitude of the individual's current online exposure, which has an opposite effect. The pattern of effects suggests that individuals may be assessing their behavioral experience (i.e., past privacy violations and size of digital footprint) separately from attitude (i.e., need for control).

An examination of these three effects on online information disclosure potentially also contributes to the research on privacy concerns [5, 55] because it identifies a behavior by which individuals may address those concerns. Although outside the scope of the present study, the research findings may also shed light on reasons for the "privacy paradox" [2, 24, 41] where a mismatch between stated privacy concerns and actual behavior has been observed [43]. Consequently, a viable partially strategy for individuals to address their privacy concerns may be to limit online information disclosures to partially anonymize their online identity.

7.1 Future research

The present study examines an important but little researched area on the online behavior of individuals that has important implications for both organizations and public policy. Several avenues for future research emerge from the current study findings.

First, the online information limiting behavior of individuals represents a serious threat to e-commerce firms as their business models are frequently built on their ability to monetize the information provided to them by individuals. When their people limit or partially anonymize their online personal information, the ability of companies to reach potential customers with revenue-generating offers is diminished. The advertising revenues generated by online media firms (e.g., Google, Facebook) is a big business which relies on targeting prospects with personalized messages that are specifically intended for them. Hence, there is a compelling need to replicate or substantiate the present findings in an e-commerce context. A possible outcome of such an effort could be that individuals have become de-sensitized to potential privacy losses in e-commerce contexts due to the high frequency with which they receive marketing messages and communications.

A second area of future research relates to studying the behavior of interest from a normative or motivational perspective as it has an ethical dimension to it [6]. Under such a perspective, individuals may be motivated to partially anonymize their online identity if they regard their relationships with firms as being unfair or imbalanced [13]. Carried to the extreme, they may also falsify or purposely provide misleading information to organizations in an attempt to get even with them. Thus, perceived unfairness, power imbalance and other similar motivational constructs could be investigated as potential influences on the individual's decision to partially anonymize online personal information [20].

A third area of future research is to determine *which* individuals are demographically more likely to limit or partially anonymize their personal information. So doing can help organizations develop strategies intended to pre-empt them from so doing. The results of the present research indicate that there are minor demographic differences in individual intentions to limit personal information online, but more detailed studies are needed.

7.2 Implications

The study findings have several implications for organizations, policy makers and individuals, some with unintended consequences. From an organizational perspective, the study findings indicate that individuals may be limiting their online information disclosures and partially anonymizing their digital identities as a *pre-emptive* behavior conducted

in anticipation of a privacy loss. Organizations should be concerned about such a development because it is a direct threat to the trust-based relationships they seek to build with people for mutual benefit. As online information disclosures are highly situational and context-dependent, organizations need to better understand *why* and *how* individuals limit or partially anonymize their personal information as it specifically applies to them. Is it because there is a general erosion in trust on the part of individuals in organizational entities? Or is it due to the need for self-preservation in a fast changing online media landscape? The challenge for organizations is to educate consumers on the advantages of information disclosure when it is for mutual benefit, while simultaneously establishing greater safeguards to protect personal data. Individuals being fearful of sharing personal information online also speaks to a larger problem. They do not believe that existing privacy policies of organizations provide adequate safeguards.

From a public policy perspective, it is important for regulatory agencies to take pro-active steps when they believe that organizations are not doing an adequate job in protecting the online personal information of individuals and/or people are having difficulty self-managing their privacy. The General Data Protection Regulation (GDPR) adopted by the EU is symbolic of the intervention of governmental agencies when organizations fail to establish adequate safeguards to protect the personal information of individuals. The recently launched probe of the privacy protection practices of Facebook by the FTC (Federal Trade Commission) in the USA also illustrative of the need for possible policy-based intervention.

From a societal perspective, it is important to recognize that *some* individuals are highly protective of their personal information online and believe that their privacy is a fundamental right that may not be violated under any circumstances [59]. Their view of privacy is akin to being left alone. Others are less protective of their personal information and are willing to reveal it in exchange of benefits. For them, privacy is more like a commodity that can be traded when incentives for so doing are provided. There are still others are unconcerned about the protection of their personal information [59]. Whatever may be the case, organizations should play a more active role in managing individual expectations regarding information disclosure and highlight the mutual benefit to both parties. Under the right circumstances, people will share personal information with organizations provided they believe the organization has a legitimate need for the information and the information exchange is for mutual benefit [57].

Finally, from the perspective of individuals, it is important to understand that the information exchange that typically occurs between an individual and an organization is in the pursuit of a shared benefit. Thus, they also share joint responsibilities in protecting their own online personal information. For individuals, this means exercising due diligence while revealing their personal information online, but not to the degree that it makes the relationship they seek untrustworthy. It is time-consuming and difficult for individuals to build and maintain trust, but relatively easy to lose it. When individuals limit online information disclosures to partially anonymize their digital identity they could be doing so because they believe the information demands of organizations are unduly intrusive (e.g., seeking access to "contacts" lists). Hence, it is incumbent on organizations to communicate that they have a legitimate need for the personal information they seek from individuals.

7.3 Conclusion

The present study examines an important but little researched area on the online behavior of individuals that has important implications for both organizations and public policy. Individual intentions to mask their personal information online and anonymize their digital identity provides a mechanism by which an individual can share information with an organization, while also protecting their privacy. It balances two conflicting goals relating to privacy protection and information disclosure and is an unintended consequence of an individual's apprehension about their personal information being misused or stolen and the privacy policies of firms.

Appendix: Survey questions in the pew research center research report used as measures of study constructs

Size of digital footprint

We'd like to know if any of the following information about you is available on the internet for others to see. It doesn't matter if you put it there yourself or someone else did so. As I read each item, you can just tell me yes or no—if you're not sure if something is on the internet, just say so.

How about (insert items in order):

a. Your email address
b. Your home address
c. **Your home phone number**
d. Your cell phone number
e. Your employer or a company you work for
f. Your political party or political affiliation
g. **Things you've written that have your name on it**
h. **A photo of you**
i. **Video of you**
j. **Which groups or organizations you belong to**

k. **Your birth date**

Response Categories

1. Yes
2. No
3. Does not apply to me
8. Do not know or not sure
9. Refused

Note Only the **highlighted items** were used as formative measures of the construct.

Need for control

Now, here is a list of some things that you might do online. For each activity, how much do you care that only you and those you authorize should have access to the following kinds of information? First, is it very important to you, somewhat important, or not too important to you that only you and those you authorize have access to?

How about (insert items in order):

a. The websites you browse
b. The place where you are located when you use the internet
c. **The content and files that you download**
d. The times of day you are online
e. **The applications or programs you use**
f. The searches you perform
g. **The content of your email**
h. The people you exchange email with
i. **The content of your online chats or hangouts with others**

Response Categories

1. Very important
2. Somewhat important
3. Not too important
4. Not at all important
5. Does not apply to me
8. Do not know
9. Refused

Note Only the **highlighted items** were used as formative measures of the construct.

Online information limiting behavior

While using the internet, have you ever done any of the following things? First, have you ever while you used the internet?

How about (insert items in order):

a. **Used a temporary username or email address**
b. Used a fake name or untraceable username
c. **Given inaccurate or misleading information about yourself**
d. Set your browser to disable or turn off cookies
e. **Cleared cookies and browser history**
f. Used a service that allows you to browse the web anonymously, such as a proxy server, or software, or a virtual personal network
g. Encrypted your communications
h. Decided not to use a website because they asked for your real name
i. **Deleted or edited something you posted in the past**
j. **Asked someone to remove something that was posted about you online**
k. Used a public computer to browse anonymously

Response Categories

1. Yes
2. No
3. Does not apply to me
8. Do not know
9. Refused

Note Only the **highlighted items** were used as formative measures of the construct.

Past privacy violations

As far as you know, have you ever had any of these experiences as a result of your online activities? Have you ever had this experience as a result of your online activities?

How about (insert items in order):

a. **Had important personal information stolen such as your social security number, your credit card, or bank account information**
b. **Had an email or social networking account of yours compromised or taken over without your permission by someone else**
c. Been the victim of an online scam and lost money
d. **Been stalked or harassed online**
e. Lost a job opportunity or educational opportunity because of something you posted online or someone posted about you online
f. **Experienced trouble in a relationship between you and a family member or a friend because of something you posted online**
g. Had your reputation damaged because of something that happened online

h. Something happened online that led you into physical danger

Response Categories

1. Yes
2. No
8. Do not know
9. Refused

Note Only the **highlighted items** were used as formative measures of the construct.

References

1. Acquisti A, Brandimarte L, Loewenstein G (2015) Privacy and human behavior in the age of information. Science 347(6221):509–514
2. Acquisti A, Grossklags J (2005) Privacy and rationality in individual decision making. IEEE Secur Priv 1:26–33
3. Ashworth L, Free C (2006) Marketing dataveillance and digital privacy: using theories of justice to understand consumers' online privacy concerns. J Bus Ethics 67(2):107–123
4. Awad NF, Krishnan MS (2006) The personalization privacy paradox: an empirical evaluation of information transparency and the willingness to be profiled online for personalization. MIS Q 30:13–28
5. Bélanger F, Crossler RE (2011) Privacy in the digital age: a review of information privacy research in information systems. MIS Q 35(4):1017–1042
6. Boatright M (2000) Privacy, ethics and the conduct of business, 3rd edn. Prentice-Hall, Saddle River, pp 159–183
7. Chellappa RK, Sin RG (2005) Personalization versus privacy: an empirical examination of the online consumer's dilemma. Inf Technol Manag 6(2):181–202
8. Chen K, Rea AI Jr (2004) Protecting personal information online: a survey of user privacy concerns and control techniques. J Comput Inf Syst 44(4):85
9. Culnan MJ, Armstrong PK (1999) Information privacy concerns, procedural fairness, and impersonal trust: an empirical investigation. Organ Sci 10(1):104–115
10. Diamantopoulos A, Winklhofer HM (2001) Index construction with formative indicators: an alternative to scale development. J Mark Res 38(2):269–277
11. Dinev T, Xu H, Smith JH, Hart P (2013) Information privacy and correlates: an empirical attempt to bridge and distinguish privacy-related concepts. Eur J Inf Syst 22(3):295–316
12. Epstein LH, Temple JL, Roemmich JN, Bouton ME (2009) Habituation as a determinant of human food intake. Psychol Rev 116(2):384
13. Fukukawa K, Ennew C (2010) What we believe is not always what we do: an empirical investigation into ethically questionable behavior in consumption. J Bus Ethics 91(1):49–60
14. Gefen D, Ridings CM (2005) If you spoke as she does, sir, instead of the way you do: a sociolinguistics perspective of gender differences in virtual communities. ACM SIGMIS Database 36(2):78–92
15. Golder SA, Macy MW (2014) Digital footprints: opportunities and challenges for online social research. Ann Rev Sociol 40:129–152
16. Hair JF Jr, Hult GTM, Ringle C, Sarstedt M (2017) A primer on partial least squares structural equation modeling (PLS-SEM). Sage Publications, London
17. Harris MA, Brookshire R, Chin AG (2016) Identifying factors influencing consumers' intent to install mobile applications. Int J Inf Manag 36(3):441–450
18. Henseler J, Sarstedt M (2013) Goodness-of-fit indices for partial least squares path modeling. Comput Stat 28(2):565–580
19. Hoffman DL, Novak TP, Peralta MA (1999) Information privacy in the marketspace: implications for the commercial uses of anonymity on the Web. Inf Soc 15(2):129–139
20. Horne DR, Norberg PA, Cemal Ekin A (2007) Exploring consumer lying in information-based exchanges. J Consum Mark 24(2):90–99
21. Jai TMC, King NJ (2015) Privacy versus reward: do loyalty programs increase consumers' willingness to share personal information with third-party advertisers and data brokers? J Retail Consum Serv 28:296–303
22. James TL, Nottingham Q, Collignon SE, Warkentin M, Ziegelmayer JL (2016) The interpersonal privacy identity (IPI): development of a privacy as control model. Inf Technol Manag 17(4):341–360
23. Jarvenpaa SL, Tractinsky N, Vitale M (2000) Consumer trust in an Internet store. Inf Technol Manag 1(1–2):45
24. Jensen C, Potts C, Jensen C (2005) Privacy practices of Internet users: self-reports versus observed behavior. Int J Hum Comput Stud 63(1):203–227
25. Jiang Z, Heng CS, Choi BC (2013) Research note—privacy concerns and privacy-protective behavior in synchronous online social interactions. Inf Syst Res 24(3):579–595
26. Lambiotte R, Kosinski M (2014) Tracking the digital footprints of personality. Proc IEEE 102(12):1934–1939
27. Lanier CD, Saini A (2008) Understanding consumer privacy: a review and future directions. Acad Mark Sci Rev 12(2):1–45
28. LaRose R, Rifon N (2006) Your privacy is assured-of being disturbed: websites with and without privacy seals. New Media Soc 8(6):1009–1029
29. Li Y (2012) Theories in online information privacy research: a critical review and an integrated framework. Decis Support Syst 54(1):471–481
30. Lowry PB, Gaskin J (2014) Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: when to choose it and how to use it. IEEE Trans Prof Commun 57(2):123–146
31. Lwin MO, Williams JD (2003) A model integrating the multidimensional developmental theory of privacy and theory of planned behavior to examine fabrication of information online. Mark Lett 14(4):257–272
32. Lwin M, Wirtz J, Williams JD (2007) Consumer online privacy concerns and responses: a power–responsibility equilibrium perspective. J Acad Mark Sci 35(4):572–585
33. McKinnon J, Vartabedian M (August 6, 2018) Tech firms, embattled over privacy, warm to federal regulation. Wall Street Journal
34. McSweeney FK, Swindell S (1999) General-process theories of motivation revisited: the role of habituation. Psychol Bull 125(4):437
35. Milne GR, Boza ME (1999) Trust and concern in consumers' perceptions of marketing information management practices. J Interact Mark 13(1):5–24
36. Milne GR, Labrecque LI, Cromer C (2009) Toward an understanding of the online consumer's risky behavior and protection practices. J Consum Aff 43(3):449–473
37. Milne GR, Rohm AJ (2000) Consumer privacy and name removal across direct marketing channels: exploring opt-in and opt-out alternatives. J Public Policy Mark 19(2):238–249

38. Milne GR, Rohm AJ, Bahl S (2004) Consumers' protection of online privacy and identity. J Consum Aff 38(2):217–232
39. Muhammad SS, Dey BL, Weerakkody V (2018) Analysis of factors that influence customers' willingness to leave big data digital footprints on social media: a systematic review of literature. Inf Syst Front 20(3):559–576
40. Newman AL (2015) What the "right to be forgotten" means for privacy in a digital age. Science 347(6221):507–508
41. Norberg PA, Horne DR, Horne DA (2007) The privacy paradox: personal information disclosure intentions versus behaviors. J Consum Aff 41(1):100–126
42. Pavlou PA (2011) State of the information privacy literature: where are we now and where should we go? MIS Q 35(4):977–988
43. Peltier JW, Milne GR, Phelps JE (2009) Information privacy research: framework for integrating multiple publics, information channels, and responses. J Interact Mark 23(2):191–205
44. Petronio S (1991) Communication boundary management: a theoretical model of managing disclosure of private information between marital couples. Commun Theory 1(4):311–335
45. Phelps JE, D'Souza G, Nowak GJ (2001) Antecedents and consequences of consumer privacy concerns: an empirical investigation. J Interact Mark 15(4):2–17
46. Phelps J, Nowak G, Ferrell E (2000) Privacy concerns and consumer willingness to provide personal information. J Public Policy Mark 19(1):27–41
47. Rainie L, Kiesler S, Kang R, Madden M (September 5, 2013) Anonymity, privacy, and security online. Pew Research Center Report
48. Rifon NJ, LaRose R, Choi S (2005) Your privacy is sealed: effects of web privacy seals on trust and personal disclosures. J Consum Aff 39(2):339–362
49. Ringle CM, Sarstedt M, Straub DW (2012) Editor's comments: a critical look at the use of PLS-SEM in MIS quarterly. MIS Q 36:iii–xiv
50. Romanosky S, Acquisti A (2009) Privacy costs and personal data protection: economic and legal perspectives. Berkeley Technol Law J 24:1061
51. Romanosky S, Telang R, Acquisti A (2011) Do data breach disclosure laws reduce identity theft? J Policy Anal Manag 30(2):256–286
52. Seetharaman D (March 28, 2018) Facebook to streamline privacy settings. Wall Street Journal
53. Sheehan KB, Hoy MG (1999) Flaming, complaining, abstaining: how online users respond to privacy concerns. J Advert 28(3):37–51
54. Sheehan KB, Hoy MG (2000) Dimensions of privacy concern among online consumers. J Public Policy Mark 19(1):62–73
55. Smith HJ, Dinev T, Xu H (2011) Information privacy research: an interdisciplinary review. MIS Q 35(4):989–1016
56. Stanton JM, Stam KR (2002) Information technology, privacy, and power within organizations: a view from boundary theory and social exchange perspectives. Surveill Soc 1(2):152–190
57. Stewart DW (2017) A comment on privacy. J Acad Mark Sci 45(2):156–159
58. Sutanto J, Palme E, Tan CH, Phang CW (2013) Addressing the personalization-privacy paradox: an empirical assessment from a field experiment on smartphone users. MIS Q 37(4):1141–1164
59. Taylor H (March 19, 2003) Most people are privacy pragmatists: who, while concerned about privacy, will sometimes trade it off for other benefits. Harris Interactive
60. Van Slyke C, Comunale CL, Belanger F (2002) Gender differences in perceptions of web-based shopping. Commun ACM 45(8):82–86
61. White TB, Novak TP, Hoffman DL (2014) No strings attached: when giving it away versus making them pay reduces consumer information disclosure. J Interact Mark 28(3):184–195
62. Wirtz J, Lwin MO, Williams JD (2007) Causes and consequences of consumer online privacy concern. Int J Serv Ind Manag 18(4):326–348
63. Xu H, Luo XR, Carroll JM, Rosson MB (2011) The personalization privacy paradox: an exploratory study of decision making process for location-aware marketing. Decis Support Syst 51(1):42–52
64. Youn S (2009) Determinants of online privacy concern and its influence on privacy protection behaviors among young adolescents. J Consum Aff 43(3):389–418
65. Zwick D, Dholakia N (2004) Whose identity is it anyway? Consumer representation in the age of database marketing. J Macromark 24(1):31–43