Falsifying and withholding: exploring individuals' contextual privacy-related decision-making

Caroline Lancelot Miltgen, H. Jeff Smith

FALSIFYING AND WITHHOLDING:

EXPLORING INDIVIDUALS' CONTEXTUAL PRIVACY-RELATED DECISION-MAKING

Caroline Lancelot Miltgen *
Audencia Business School
clancelot@audencia.com

H. Jeff Smith
Department of Information Systems and Analytics
Miami University
jeff.smith@MiamiOH.edu

* **Corresponding Author**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Article published in Information and Management (2019), 56(5), 696-717

FALSIFYING AND WITHHOLDING:

EXPLORING INDIVIDUALS' CONTEXTUAL PRIVACY-RELATED DECISION-MAKING

Abstract

As firms rely increasingly on "big data" to segment and target current and potential customers, the challenge of data falsification—individuals providing incorrect personal data in response to requests—is becoming a significant problem. Based on public opinion surveys, within some demographic groups, over three-quarters of individuals confirm that they have given inaccurate information in response to data requests. Obviously, firms that embrace a covert assumption of honesty in online data disclosures are deluding themselves and are likely falling into the trap of "garbage in, garbage out" in their segmenting and targeting.

Despite the frequency and importance of falsification, however, it has received scant attention in the privacy research stream. Most researchers focus on the act of disclosure (and its counter-construct, withholding of data) and overlook that many of the data elements being disclosed may in fact be falsified. To address this weakness in the literature stream, we develop a nomological model that predicts both falsification and withholding behavior, and we test it using a sample collected with the assistance of an online panel provider. We find strong support for the model and show how context could play a significant role in moderating some of the proposed relationships. We then discuss important implications for practice and research.

Keywords: information privacy; falsification; disclosure; withholding; personal data; privacy tradeoff, context

FALSIFYING AND WITHHOLDING:

EXPLORING INDIVIDUALS' CONTEXTUAL PRIVACY-RELATED DECISION-MAKING

The Pizza Shack, a small pizza delivery service, created a social media application that asked visitors to enter (among other things) their date of birth when configuring their profile on the app. The Pizza Shack's owners envisioned the creation of a birthday-related promotion that would send a coupon to each registered customer one week in advance of his or her birthday each year. However, when their IT staff personnel attempted to test the "birthday coupon" notification feature, they were astounded to find that 43% of the coupons were scheduled for distribution on Christmas Day, December 25. Confused, the staff looked carefully at the database and discovered that almost half of the registrants had chosen "1" and "January" as their day and month of birth, respectively. After randomly querying some of these customers who were willing to discuss the matter, the staff concluded that a large number of registrants had not wanted to reveal their actual birthdate to the Pizza Shack so had simply chosen a year (from the "year" pulldown menu) that made them over 18 years old and had selected the first day and month that appeared on the "day" and "month" pulldown menus. Before totally canceling the promotion, the company decided to conduct further interviews to better understand why registrants had reacted this way.

Most of the interviewees said they were unsure of what could be done with this information, suggesting a lack of trust in the company running the app. Following this feedback, Pizza Shack decided to better explain, in the registration screen, why this specific information was needed at that point and how it would be used (i.e., to send people some promotion coupons for their birthday), two elements which had been omitted in the first version of the app. This resulted in a significant reduction of registrants claiming to be born on January 1 in the database and led to a successful promotion campaign.

INTRODUCTION

When faced with a request for personal data, an individual has three basic and mutually exclusive alternatives to the request:

a) Honestly disclose the information;

b) Withhold the information (sometimes called "non-disclosure"); and

c) Disclose false information.[1]

Unfortunately, firms often ignore the possibility of individuals embracing the third option and instead assume that the decision is a binary one (disclose or withhold). This leads them to the inaccurate conclusion that all disclosed data are accurate. Similarly, researchers tend to focus on options a and b, whereas we focus on option c.

The Pizza Shack scenario—disguised but based on an actual situation—illustrates the danger of firms' relying on the accuracy of users' self-reported inputs. Indeed, one industry observer has lamented that the increasing preponderance of such a phenomenon could destroy the value of "big data" (Lima 2015). A recent survey found that "two in three [consumers] admit…to deliberately giving incorrect information" (Lobel 2015, p. 1), and the frequency of such falsification varies across age groups, with 81% of those between ages 18 and 24 "admitting to provide, at least occasionally, wrong information when asked for personal information" (Lobel 2015, p. 1). As can be seen in Table 1, the frequency of falsification varies for different data types, with about one-third of individuals in one survey reporting that they have falsified their birthday to a "website, service, or a mobile form that asked for personal information" (Hodder et al. 2013, p. 2). Given these survey findings, it is hardly surprising that the firm in our opening vignette found that the data collected in its "birthday promotion" were flawed.

[1] We use the phrase "false information" and the verb "falsify," although some other researchers have embraced the phrase "misrepresentation" and the verb "misrepresent," which seem to be synonymous.

Table 1. FALSIFICATION BY DATA TYPE

Answers to the query "Have you ever made up and submitted incorrect information to a website, service or a mobile form that asked for personal information? (Please check all that apply.)"

I have provided…	Percentage
…an incorrect phone number	34.24%
…a poor or "spam" email address	33.62%
…an incorrect birthday	33.50%
…an incorrect name	31.64%
…incorrect identity information	22.11%
…incorrect preferences	15.85%
…incorrect employment information	13.81%

Source: Hodder et al. 2013. Total respondents: 1,615, geographically distributed U.S. residents, aged 18 and older.

In spite of this problematic phenomenon, it is quite rare for information systems (IS) researchers to focus on its occurrence or factors associated therewith. This is indeed ironic, because digital disclosure of personal information from an individual to an online firm (our focus here) appears to be the most frequently employed dependent variable (DV) in empirical privacy-related IS research (Smith et al. 2011). However, most researchers seem to rely on a covert assumption that individuals' disclosures of personal information are accurate ones. Appendix 1 shows that 42 studies employing online disclosure to an organizational entity as a DV have appeared in the top IS journals,[2] but only one addressed the phenomenon of individuals providing false information during digital disclosure to a firm: Son and Kim (2008).[3] Son and Kim (2008) examined Internet users' "privacy-protective responses," which included two alternatives associated with "information provision" (refusal and misrepresentation), two alternatives associated with "private action" (removal and negative word-of-mouth), and two alternatives associated with "public action" (complaining directly to online companies and complaining indirectly to third-party organizations). However, only one of

[2] See Appendix 1 for details regarding the 10 journals included in this category.

[3] Jiang et al. (2013) also considered misrepresentation as a "privacy-protective behavior" in their study, but their focus was on disclosure in synchronous online social interactions with peers, which differs from our focus here.

their independent variables (perceived justice) was significant in predicting misrepresentation behavior, and that relationship explained only 3% of the variance in misrepresentation.

Looking beyond these top IS journals, we found another IS article that has addressed this issue of false information: Keith et al. (2013) included a DV that addressed whether individuals' data disclosures were in fact honest ones. They discovered that what at first appeared to be a weak relationship between subjects' stated disclosure intentions and their actual disclosures was actually a strong one when they considered *honest* disclosures instead of disclosures. However, their assessment of falsification was conducted in a post hoc survey, and it was not the focus of their model or theoretical derivation.

As it turns out, this lack of attention to falsification in disclosure to online firms is not restricted to IS journals. Our search of the top marketing journals and our overarching search of titles in other disciplines through Google Scholar revealed only five articles that have directly addressed this topic through empirical studies (Malheiros et al. 2013; Metzger 2006; Sheehan et al. 1999; Wirtz et al. 2007; Xie et al. 2006). Although some of the studies did yield quite interesting findings (e.g., Malheiros et al. (2013) found that perceptions of "fairness" were associated with levels of falsification), none of those five viewed falsification as a primary focus of an overarching theoretical model.

We can conclude that despite the obvious importance of the topic of falsification, surprisingly little attention has been paid to this topic in the IS literature and, for that matter, in academic research as a whole. Generally speaking, studies that attempt to explain individuals' disclosure of personal data covertly assume that the mere disclosure of data is sufficient to enable subsequent use. However, as highlighted by the old adage "garbage in, garbage out," to the extent that the integrity of databases is reduced through falsification, the implications are stark, especially when juxtaposed against greater analytical use of "big data" to both categorize and target specific consumers. As was noted by a Chief Technology Officer, "… more and more people are lying to their

service providers. A big part of the…ecosystem is big data, analytics and the power of information. That starts falling to pieces if we find a lot of people are lying" (Lima 2015, p. 1).

To understand better what is contributing to such behavior, we attempt to answer the following research questions:

1) To what extent does a privacy tradeoff – based on perceived benefits, perceived risks, and trust – explain a) individuals' data disclosure decisions, such as withholding, and b) the phenomenon of individuals providing falsified data in privacy-related decisions?
2) To what extent can individuals' perceptions of risks, benefits, and trust in privacy-related decisions be explained by the contextual factor of perceived relevance?

Further, we consider an exploratory question that, while not having been addressed with much theoretical rigor in the past, deserves consideration in the future research stream:

3) How does the context of a data request influence individuals' decisions regarding data disclosure (including withholding) and falsification?

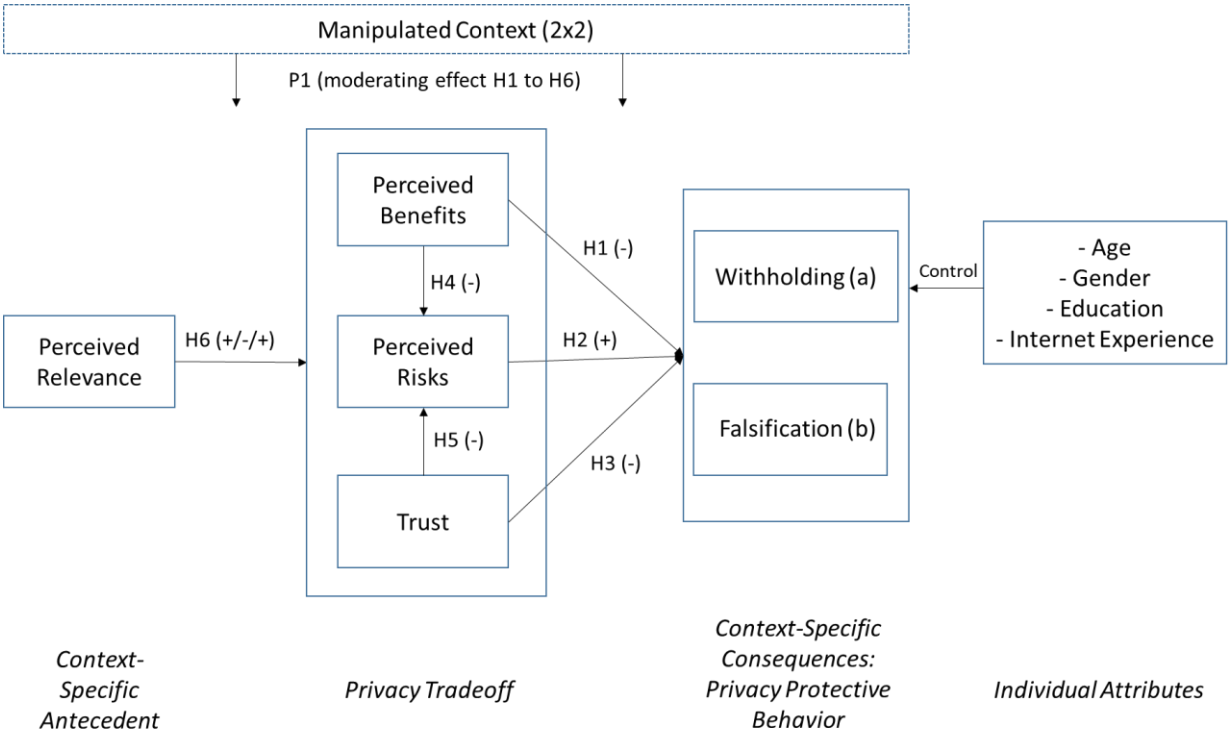
We derive a model that explains not only disclosure (more precisely, withholding) of personal information but also falsification thereof. With the cooperation of a French online access panel provider, we conducted an online experiment in which participants confronted a real-world situation that would prompt an online request for personal data from a commercial website. Our results confirm that large percentages of the variance in both withholding and falsification can be explained by a tradeoff among perceptions of risks, benefits, and trust, each of which, in turn, is associated strongly with the contextual factor of perceived relevance. These findings are particularly unique for the DV of falsification, as this is the first study to have demonstrated this phenomenon.

In the following section, we propose a research model with several specific hypotheses and one proposition. We then provide a literature review that justifies those hypotheses and explains that proposition. We follow this with a discussion of methods of this study and an analysis of our data. We conclude with a discussion of implications for research and practice.

BACKGROUND AND MODEL DEVELOPMENT

As is obvious from the discussion above, no theory regarding falsification of data in response to disclosure requests from organizational entities has emerged yet in any literature stream. Therefore, we appeal to the literature on disclosure itself and derive an exploratory model that incorporates both withholding and falsification as DVs in an umbrella construct known as "privacy protective behavior" that we borrow from Jiang et al. (2013).[4] At its core, consistent with a thorough literature review in the domain of data disclosure (Appendix 1), our model considers the privacy tradeoff that is most often associated with the privacy calculus paradigm and some extensions thereto—in particular, perceived relevance. Also, in an attempt to consider our third, exploratory research question regarding context, we test for moderating effects of some contextual factors across the model. Figure 1 and Table 2 summarize our model and hypotheses/proposition, which we explain below.

FIGURE 1. RESEARCH MODEL



[4] Although Jiang et al. (2013) focused on disclosure and falsification of data in a social networking environment rather than organizational entities, we nevertheless rely on their nomenclature.

TABLE 2. SUMMARY OF HYPOTHESES/PROPOSITION

H1: Higher levels of perceived benefits will be associated with lower levels of a) withholding and b) falsification.
H2: Higher levels of perceived risks will be associated with higher levels of a) withholding and b) falsification.
H3: Higher levels of trust will be associated with lower levels of a) withholding and b) falsification.
H4: Higher levels of perceived benefits will be associated with lower levels of perceived risk.
H5: Higher levels of trust will be associated with lower levels of perceived risk.
H6: Higher levels of perceived relevance will be associated with a) lower levels of perceived risk, b) higher levels of perceived benefits, and c) higher levels of trust.
P1: Relationships in H1-H6 (above) will be moderated by contextual factors associated with data disclosure requests.

The Privacy Tradeoff

To the extent that one theoretical perspective has been more broadly embraced than others in attempts to explain data disclosures, that perspective would undoubtedly be the privacy calculus, which assumes that individuals perform a "risk-benefit analysis to assess the outcomes they would face in return for the [data] and respond accordingly" (Smith et al. 2011, p. 1001). In the past, the concept of "respond accordingly" within the privacy calculus subdomain has usually been construed as disclosing the data or refusing to do so (i.e., withholding). As discussed earlier, in this study, we extend these previous findings by postulating that another option is conceivable: an individual might provide falsified data.

As can be seen in Appendix 1, of the 36 (out of 42) studies that claimed a reliance on one or more theoretical bases, privacy calculus (or a closely related variant such as utility maximization) was embraced more often than any other theory (by 14 of the studies claiming a theoretical basis, or 39%). Although privacy calculus's tradeoff assumption has been challenged as relying on a sometimes unrealistic assumption of high-effort cognitive processing on individuals' parts (Dinev et al. 2015), its widespread use in disclosure-focused studies suggests that it should be embraced as the starting point for our model.

The fundamental theoretical argument undergirding privacy calculus was stated succinctly by Culnan and Bies (2003, p. 327) as follows:

People disclose personal information to gain the benefits of a relationship; the benefits of

> disclosure are balanced with an assessment of the risks of disclosure…In other words, individuals will exchange personal information as long as they perceive adequate benefits will be received in return—that is, benefits which exceed the perceived risks of information disclosure…

This perspective, which has since been widely embraced across the privacy literature stream, has been adopted in investigations of various forms of data disclosure, and it has proven robust across almost all of them. The few articles that have empirically examined falsification behavior have not relied on this theory, per se, although Xie, Teo, and Wan (2006) refer to utility theory and did consider "rewards" (a correlate to perceived benefits) as an antecedent to the provision of accurate personal information. Because theoretical development for falsification has not yet emerged, we postulate a simple converse effect to that for disclosure in these core privacy tradeoff hypotheses:

H1: Higher levels of perceived benefits will be associated with lower levels of a) withholding and b) falsification.

H2: Higher levels of perceived risks will be associated with higher levels of a) withholding and b) falsification.

Although the privacy tradeoff forms the core of our model, we have also included some additional constructs in our investigation of the deeper complexities of privacy decision-making.

Impact of trust on dependent variables

The trust construct has played a role in a large number of studies in the overall IS privacy research stream and, especially, in those studies that have embraced disclosure as their DV. As can be seen in Appendix 1, of the 42 studies employing online disclosure to an organizational entity as a DV, 18 (i.e., 42.8%) included a trust-related construct in their model. None of the studies that measured falsification as a DV included trust in their model, although Xie et al. (2006) did consider a requesting firm's "reputation," and Metzger (2006) showed that reputation did impact trust. Although trust is often viewed as an interpersonal dyadic variable in psychological research, the perspective in privacy studies (and, for that matter, most IS studies in general) has been trust in an

organizational entity, and that is the perspective we embrace here.

A consistent finding with the privacy literature stream—which is supported by the entries in Appendix 1—is that individuals are less likely to withhold personal data when they trust the entity in the context of the request. From a theoretical perspective, authors' arguments are varied. Recently, for example, Bansal et al. (2016) relied on prospect theory and argued that trust increases individuals' perceived utility in disclosing information, and Ozdemir et al. (2017) relied on previous research into trustor-trustee relationships to confirm a similar hypothesis regarding disclosure. We expect the same result to hold here, and it stands to reason that the converse effect will hold for withholding and falsification:

> **H3:** Higher levels of trust will be associated with lower levels of a) withholding and b) falsification.

Perceived benefits and risks

Although both perceived benefits and perceived risks are important components in the privacy calculus tradeoff, it remains unclear to what extent the two constructs impact one another. Based on previous precedent (e.g., Dinev et al. 2013), we hypothesize that higher levels of perceived benefits will be associated with lower levels of perceived risks. Dinev et al. (2013, p. 302), who also relied on privacy calculus as the core of their model, explained as follows:

> The notion of privacy calculus assumes that there is a consequential tradeoff of costs and benefits salient in an individual's privacy decision-making. Overall, the calculus perspective of privacy suggests that when asked to provide personal information to service providers or companies, consumers…behave in ways that they believe will result in the most favorable net level of outcomes. Consequently, we argue that consumers are more likely to accept the potential risks that accompany the disclosure of personal information as long as they perceive that they can achieve a positive net outcome. Hence, when a positive outcome of information disclosure is anticipated, risk beliefs are hypothesized to decrease.

We embrace this same logic and hypothesize:

> **H4:** Higher levels of perceived benefits will be associated with lower levels of perceived risk.

Trust and perceived risk

While the nomological role of trust has not been fully clarified (Smith et al. 2011), it is widely viewed as being an important input into disclosure decisions. What has never been established—not only in the domain of privacy research but also in the broader domain of IS research in general—is the specific relationship between perceived trust, perceived risk, and intended behaviors. Looking across a sample of models that have included these constructs (e.g., Bansal et al. 2010; Buttner et al. 2008; Nicolaou et al. 2006; Robert et al. 2009; Zimmer et al. 2010b), one can find numerous combinations of mediation and moderation among these constructs.

In our model, we follow the path of Zimmer et al. (2010b), who demonstrated in a privacy-related study that perceived trust impacts perceived risk (which, in our model, impacts privacy protective behaviors). Zimmer et al. (2010b, p. 117) argued that "[c]onsumers that trust a website believe that there is more predictability regarding usage of information by the exchange party, reducing transaction risk for the consumer." They extended those concepts to include disclosure, and we follow their path with our hypothesis:

H5: Higher levels of trust will be associated with lower levels of perceived risk.

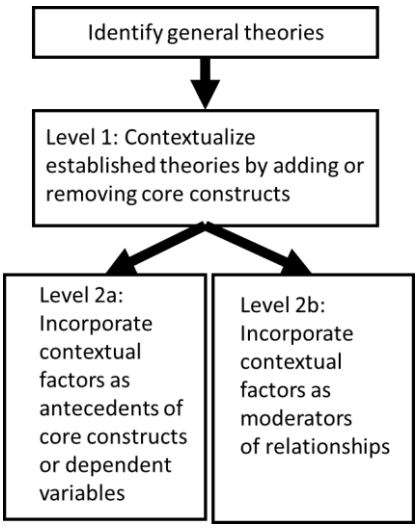
Context

Context has been defined as "situational opportunities and constraints that affect the occurrence and meaning of organizational behavior as well as functional relationships between variables" (Hong et al. 2014, p. 112). Nissenbaum (2014) noted that there are four interpretations of context that may prove salient: context as technology system or platform, context as business model or business practice, context as sector or industry, and context as social domain. In terms of a definitional boundary, "context encompasses stimuli and phenomen[a] that surround and thus exist in the environment *external to the individual*" (Bansal et al. 2016, p. 2, italics added). In that light, one would consider contextual variables to include those that cannot be controlled by an individual

decision-maker but which may well impact the outcome of his/her decision-making process. Factors that are internal to the decision-maker (e.g., personality traits such as introversion/extroversion) are *not* examples of context but are, rather, components that can be considered for inclusion in the basic decision-making model itself.

As was clearly detailed by Hong et al. (2014), contextual theory can be developed in one of two ways, although only one is salient for our immediate purposes. The approach that is meaningful in our study is what is termed "single-context theory contextualization," which assumes at least a moderately well-developed theory exists in a particular domain of interest. By adding or removing core constructs, the theory is then contextualized to account for new factors. As can be seen in Figure 2, this can occur through either "Level 2a," in which additional contextual factors are added as antecedents or through "Level 2b," in which contextual factors are added as moderators. As shown below, our model will be contextualized in both ways.[5]

Figure 2
Approaches to incorporating context into theorizing[6]



[5] We also note that although it would not contribute to our immediate endeavor, one might turn to replication of an existing theory across numerous contexts and then consolidate the findings through theory-grounded meta-analyses. See Hong et al. (2014, Section 3.2) for details.
[6] Adapted from Hong et al. (2014, p. 115). We included only the domain of "single-context theory contextualization" and omitted the parallel domain of cross-context theory replication, which is usually associated with grounded theory, as it is less relevant to our immediate efforts. Also, we omitted a third Level 2

In our model, we incorporate the perceived relevance factor under Level 2a and the moderating impacts of important contextual factors across our model under Level 2b.

Against that background, we now consider the "Level 2a" antecedent in our model: perceived relevance.

Contextual antecedent: perceived relevance

As can be seen from the underlined constructs in the "Antecedents to DVs" column of Appendix 1, a large number of contextual factors have been considered in studies associated with data disclosure, but the domain has not converged on a single and theoretically justified grouping. Because it would obviously be impossible to incorporate all of them in a single model, we include only one important contextual antecedent in our model: perceived relevance. We are persuaded by the findings of Zimmer et al. (2010b) who demonstrated the important role of perceived relevance in explaining perceived risk—our core mediating variable—and we extend their own model by hypothesizing relationships not only with perceived risk but with the full mediating block in our model including also perceived benefits and trust.

Perceived relevance and perceived risks

Frequently, an individual who receives a data request assesses it in light of the context of the request: that is, does the individual view the data as salient for the intended transaction(s)? In a 2003 study (Hodder et al., 2003), 1553 out of 1704 (91%) respondents acknowledged that they withheld or submitted incorrect information at a website or mobile app. A reason cited by 67.5% of those 1533 was as follows: "I was afraid of what the site or app would do with the information." Only limited research attention has been focused on this relationship: Li et al. (2010) showed that individuals usually consider the relevance of a particular data element to an intended transaction in the rubric of

action item ("Decompose core constructs into contextual factors") because our model derivation was intentionally parsimonious based on prior research.

"exchange fairness," and this is associated with their "privacy risk belief" in their privacy calculus. Also, Knijnenburg, Kobsa, and Jin (2013) hypothesized that perceived risk is influenced by a subject's perception of "purpose specificity" in a data request, such that, for a given website, "perceived risk is lower for the type of information that clearly matches the purpose of the website," and for a given type of information, "perceived risk is lower for the website that has a clear purpose for requesting the information" (Knijnenburg et al. 2013, p. 4). They found significant support for this hypothesis. Additionally, Zimmer et al. (2010b) found that perceived relevance strongly influences perceived risk.

While one might argue that such an inference is unwarranted rationally (after all, the risk of providing the data is rationally a function of protections for the data against unintended uses, intrusion, etc., rather than the intended use itself), individuals' perceptions are not always formed through purely rational assessments (e.g., Acquisti et al. 2005; Goes 2013). On the basis of limited previous findings and the argument above, we hypothesize

H6a: Higher levels of perceived relevance will be associated with lower levels of perceived risks.

Perceived relevance and perceived benefits

Although previous research has included little focus on the relationship between perceived relevance and perceived benefits, we expect that it will be the converse as for perceived risks: that is, individuals will (perhaps irrationally) infer that a transaction itself will provide greater benefits when they perceive the data request as relevant to that transaction. This is largely due to an affective contextual effect: an individual will infer that the situation is a positive one (Petty et al. 1986) and will therefore infer that greater benefits will accrue. Acknowledging the exploratory inference, we offer the following hypothesis:

H6b: Higher levels of perceived relevance will be associated with higher levels of perceived benefits.

Perceived relevance and trust

Two studies shed tangential light on the linkage between perceived relevance and trust. Liu, Marchewka, Lu, and Yu (2004) found strong support for a relationship between an individual's understanding of how data would be used after its collection and trust, suggesting that being convinced that the data were relevant for a particular transaction was important in building trust. Suh and Han (2003) tested the relationships between several subconstructs of "control"—which is related to relevance due to the individual's ability to dictate that data may be used for only what (s)he deems as relevant purposes—and trust and found mixed support across the various subconstructs. Looking across these two studies, we conclude that, while it appears that there is some evidence of a relationship between perceived relevance and trust, the precise parameters of that relationship have not previously been clarified.

Extending these earlier findings, we postulate that, if an individual perceives that an entity is requesting personal data that are germane neither to an immediate transaction nor to the individual's ongoing relationship with the entity, (s)he will often perceive that the entity is attempting to garner personal data inappropriately, and this will reduce his or her level of trust in the entity making the request:

H6c: Higher levels of perceived relevance will be associated with higher levels of trust.

Having established these hypotheses for the "Level 2a" antecedent of perceived relevance, we turn now to our "Level 2b" proposition regarding moderation.

Moderation

In the spirit of the approach embraced by Bansal et al. (2016), we manipulate the context in which personal data are requested from individuals and consider the moderating effects of this "manipulated" context on the overall model (i.e., moderating effects on hypotheses H1 through H6). Unlike Bansal et al. (2016), we do not consider industry as a representation of context but rather

consider two contextual cues that could influence disclosure decisions: the amount of data requested (low or high) (e.g., Kanuk and Berenson 1975) and the incentive behind the data request (pecuniary vs. nonpecuniary) (e.g., Acquisti et al. 2013).

As can be seen in the "Moderators" column of Appendix 1, only a few contextual items have been considered as moderators in previous research, and no theoretical framework has emerged. In fact, it is even difficult, a priori, to make directional predictions regarding moderating effects. One can, at best, postulate that some contextual effect may be likely, but whether it is likely to strengthen or weaken the numerous relationships within the model is largely dependent on 1) the contextual cue studied (and its manipulation thereof) and 2) the independent and dependent variables being considered in the model. Because of the exploratory nature of our study and because we are interested in testing whether (or not) the context can influence the privacy-related decision-making of individuals (especially to explain falsification) rather than decomposing each moderating effect, we state the as a proposition (rather than a hypothesis) the following:

> **P1:** Relationships in H1-H6 (above) will be moderated by contextual factors associated with data disclosure requests.

Control variables

In addition to the tested hypotheses above, we also include some control variables. Many prior studies have gathered some demographic data about the involved subjects, but this has almost always been done without any a priori hypothesizing regarding demographic relationships to other constructs.[7] In keeping with this tradition, we include three demographic facets (age, gender, and education) and one experiential element (Internet experience) as control factors.

Having now established the research model and hypotheses for our study, we turn to a discussion of the research method.

[7] One exception is Krasnova et al. (2012), which included a hypothesis regarding gender, that failed to find support.

METHOD

Research Design

We conducted an online experiment to test the proposed model. All participants confronted a real-world situation that would prompt an online request for personal data from a commercial website. Participants were randomly assigned to one of four different treatments. Those treatments correspond to two contextual cues that are interesting to test both from a theoretical and a managerial perspective: 1) the size of the request (i.e., amount of data requested) and 2) the inferred incentive to accrue from fulfilling the request (i.e., why the individual might respond to the data request).

In both experimental conditions, subjects were told to think of their current main mobile phone provider. Half of the participants were told they could create–by filling in an online form–a personal profile on the homepage of the website of their mobile phone provider; by doing so, the subject would be provided with exclusive and personalized information and offers. The other half of the participants were offered the opportunity to participate in a lottery sponsored by their mobile phone provider. In both cases, the basic form was identical, although the detailed data that were requested on the form were also manipulated, so that half of the participants received a short form containing only five fields (name, postal address, zip code, city, and email), whereas the other half received a longer form (same fields as the short form and 15 additional fields, which included some items such as mobile preferences and behavior). These forms and details were designed to match the forms that web surfers typically fill in when considering offers within these contexts (see Appendices 2 and 3).

One hundred and sixty eight web users from a French online access panel participated in the study. The access panel company was told to build a sample of 180[8] French web users who represented the population of interest (mobile phone users) on the basis of two criteria: age and

[8] We requested a sample of this size because we wanted to have more than 40 participants per treatment, which should ensure normal distributions for the variables tested, while taking into account incomplete questionnaires or respondents who did not match our criteria (e.g., no mobile phone).

gender. A total of 300 people were contacted to participate in the study. After excluding those who did not answer or participate (104) and those who did not answer all the questions (8), we obtained a total of 188 participants, of whom 20 did not match the quotas, thus resulting in a final sample size of 168 people, with an effective response rate of 168/300 = 56%. In general, a probability sample is better than a convenience sample (e.g., students) whose homogeneous characteristics (e.g., age and education level) sometimes lead to questions regarding the generalizability of results. However, because this study employed randomized assignment to treatment conditions, the most salient question regarding sampling is whether the treatments are germane from the subjects' perspective. While generalizability is enhanced in this study through the broad characteristics of the sample, it is in fact achieved through randomized treatments and their understandable context. Sample characteristics are presented in Appendix 4.

As the goal of this paper is to investigate Internet users' privacy-related decision-making processes, the online survey presented participants with sample screen pages describing the scenario and showing the online forms to be filled in. Screen page samples are effective for eliciting perceptions in typical situations such as those associated with online data disclosure decision-making (Finch 1987). To make the situation as realistic as possible, participants were asked the name of their mobile provider at the beginning of the experiment, and the screenshot presented to them later in the survey included the logo of their mobile provider to make participants feel as if they were really on their mobile provider's website. Because of the policies of the access panel provider, especially the assured anonymity of the participants, subjects could not be asked to provide actual data in the experimental form and could only be asked whether they would provide the requested data elements as shown in the screenshots. Although this might be viewed as a limitation of this study (discussed later in this paper), this approach reduced the time required for administrating the survey, thus diminishing participants' fatigue and ensuring their willingness to participate.

As a manipulation check, we conducted an ANOVA and found that the data requested were perceived as less relevant in the long form condition (3.30) as compared to the short form (4.58, F =

17.98 and p = 0.000 for relevance), but no difference in terms of relevance appeared when the homepage (4.07) and lottery contexts were compared (3.81, F = 0.746 and p = 0.389).

Measurement

This study measured the model constructs using single or multi-item scales. In accordance with our agreement with the access panel company, the survey could require no more than 10 min in total to be answered, including the time necessary to read the scenario and the forms. Therefore, when possible, we adapted scales from the orthodox literature while choosing the more parsimonious alternatives to reduce the time necessary to answer the questionnaire and to minimize participants' cognitive burden. As a consequence, we employed no more than four items per scale, and we favored single-item scales whenever possible, especially for easy-to-understand constructs (Bergkvist et al. 2007) such as perceived relevance.

Specifically, withholding was measured by three items assessing whether the participants would have (or not) disclosed the requested data (see Appendix 5). One of the items was measured on a 1 to 5 scale from 1 ("not at all") to 5 ("yes certainly"), while the two others were measured using a more classical 1 to 7 Likert format. The falsification item asked whether the participants would have lied about at least one item, a scale adapted from Malhotra et al. (2004) and Son and Kim (2008). This item was measured on a 1 to 5 scale from 1 ("not at all") to 5 ("yes certainly"). Perceived relevance and trust were both measured with one item each assessing, respectively, whether the data requested were perceived to fit the situation and how much the participants trusted the firm requesting those data. Both were measured using a classic Likert 1 to 7 format, with 1 meaning "not agree at all" and 7 "totally agree." On the basis of the literature in the corresponding research streams and one qualitative study, we developed multi-item scales for perceived risks (3 items) and benefits (4 items). Both used a semantic differential scale in a 1 to 7 format (i.e., 1 ''not at all useful'' to 7 ''very useful'').

As the scales corresponding to the risks and the benefits were self-developed, we conducted a pilot test with a convenience sample of 53 students to assess the dimensionality and internal reliability

of the scales. The EFA concluded that all measures loaded distinctly on their corresponding factors. All Cronbach's alphas were also above 0.7, thus indicating that the scales were reliable (see Appendix 6).

ANALYSIS AND RESULTS

We used a structural equation modeling (SEM) approach to validate the measures and test the relationships between all the constructs. A variance-based partial least squares (PLS) method, using Smart PLS 3.2.6 (Ringle et al. 2015), offers greater benefits than covariance-based methods, because a least-squares estimation procedure avoids restrictive assumptions such as multivariate normality and residual distributions (Chin 2010). As PLS does not generate an overall goodness-of-fit index, model validity is assessed by examining the structural paths and R^2 values (Chwelos et al. 2001). In addition, following recent advice from Henseler et al. (2013), we provide the standardized root mean square residual (SRMR) for the PLS estimation (0.055), which is below the cutoff value of 0.08 suggested by Hu and Bentler (1998) by applying covariance-based SEM . Although it has recently been argued that the cutoff value of 0.08 is probably too low for PLS-SEM (Hair et al. 2017), the fact that our SRMR is well below that value shows that our model has a good fit.

Test for Common Method Bias

We checked for common method bias (CMB) in our data. First, we used *a priori* methods to limit CMB as suggested by Podsakoff et al. (2012). For example, we controlled *a priori* for item ambiguity through the feedback received from the pilot study with students. The choice of single-item or small number-item scales also reduced the chance that the participants "become fatigued by a seemingly unending stream of questions" (Podsakoff et al. 2012, p. 561). We additionally reduced proximity effects by separating items related to a same construct. Finally, we included some reversed items (for example, for the perceived relevance construct) in the survey.

In addition to these *a priori* methods that should have limited the chances for CMB to occur, we also addressed CMB *a posteriori* by statistical analysis (Podsakoff et al. 2003). There is some controversy in the current literature concerning the efficiency of currently available *a posteriori* methods in detecting and correcting CMB. Given our restrictions regarding the number of questions to be included in our questionnaire, we could not include a marker variable in addition to the constructs already present. We therefore used two other methods to statistically assess CMB *a posteriori*—methods that have been widely used in previous studies published in IS journals. We first employed Harman's single-factor test as suggested by Podsakoff et al. (2003). All the variables were loaded into an EFA, and the un-rotated factor solution was examined. CMB may exist if (1) a single factor emerges from the un-rotated factor solution or (2) one general factor accounts for the majority of the covariance in the variables (Podsakoff et al. 2003). Neither occurred here, suggesting that the CMB is not an issue in this study. We also followed the approach used by Liang et al. (2007). Using SmartPLS, we specified a method factor together with the original latent variables in the measurement model, and we calculated the squared factor loadings for both the method factor and the substantive factors (i.e., original latent variables). The average variance explained by the substantive factors was approximately 0.92, while that explained by the method factor was approximately 0.10, thus confirming that CMB is not a major concern in our study (see Appendix 7).

Measurement model: Instrument validation

PLS models require a two-stage analysis: the measurement model and the structural model (Chin 2010). The measurement model, which consists of the relationships between the constructs and the measurement indicators, assesses the psychometric properties of the scales through item loadings, internal consistency (reliability), and convergent and discriminant validity. The bootstrap sampling procedure enables testing of the magnitude and significance of the loadings using well-established guidelines regarding the convergent validity, discriminant validity, and reliability of the scales (Chin 2010; Gefen et al. 2005).

In our model, each construct exhibits consistent positive loadings (see Table 3), indicating general convergence. The standardized item-construct loadings are high (> 0.761) and statistically significant at the 0.001 level, with t-statistics well above 1.96. The results also indicate satisfactory item reliability for all the measures. The CRs, similar to Cronbach's alpha but considering actual factor loadings instead of assuming that each item is equally weighted, range from 0.858 to 0.924 (excluding constructs measured with a single item), thus above 0.70, indicating good internal consistency. In addition, all average variance extracted (AVE) values are greater than the suggested 0.50, thus indicating good convergent validity for the measurement model (Fornell et al. 1981).

We assessed the discriminant validity of reflective scales (Gefen et al. 2005; Lowry et al. 2009) by comparing the AVE of each construct with the shared variances between a single construct and all the other constructs (Fornell et al. 1981). Comparison of the square root of the AVE (figures on diagonal, Appendix 8) with the correlations among the constructs indicates that the items load higher on their intended construct than on any other construct, thus indicating satisfactory discriminant validity of all the constructs (Hair et al. 2011). Recent research on discriminant validity assessment has shown that a new approach based on the multitrait-multimethod matrix, called the heterotrait-monotrait (HTMT) ratio of correlations, provides a performance superior to the previously adopted methods such as the Fornell-Larcker criterion (Henseler et al. 2015). In our case, the HTMT assessment indicates that all the construct correlations hold for the most conservative criterion HTMT (0.85) (see Appendix 9), thus suggesting that discriminant validity is established also using this more conservative criterion (Henseler et al. 2015). Overall, these results suggest sufficient reliability and convergent/discriminant validity, which allow an interpretation of the structural parameters.

Structural Model Assessment and Hypothesis Testing

Following the structural model assessment procedure (Hair et al. 2017), we first need to check the structural model for collinearity issues by examining the VIF values of all sets of predictor constructs in the structural model. All our VIF values (which range between 1.321 and 1.375) are clearly below the threshold of 5, which indicates that collinearity is not a critical issue in the structural model, and we can accordingly assess the structural model results. The examination of the path coefficients and the variance explained (R^2) in the endogenous variables enables the assessment of the structural model. Following the recommendation of Chin (2010), we tested for the statistical significance of each path coefficient by t-tests using bootstrapping with 500 subsamples. Only one control variable out of the four that were included in the model (i.e., age, gender, education, and Internet experience) is significant: gender significantly influences falsification as women tend to falsify less often than men (beta = - 0.252, $p < 0.001$).

Table 3, Table 4, and Figure 3 show the results of the structural model estimation, including the significant standardized path coefficients, t-statistics, and the significance and amount of variance explained (R^2) along with the model-predictive relevance with regard to each endogenous construct (Q^2). The model explains a substantial amount of variance for both intentions to disclose ($R^2 = 0.546$) and falsify ($R^2 = 0.210$), greater than the recommended level of 0.10 (Falk and Miller 1992). The model also explains a great part of the variance of all mediating variables, with R^2 ranging from 0.187 (for trust) to 0.349 (for perceived risks). We also assess the structural model's predictive validity by the Q^2 value of the predictive relevance (Geisser 1974; Stone 1974). After running the blindfolding procedure (Chin 1998; Henseler et al. 2015) in SmartPLS, we obtain the Q^2 values well above zero for all DVs, indicating the predictive relevance of the PLS path model (see Table 3).

TABLE 3. STATISTICS FOR CONSTRUCTS

	Number of items	Loadings (min - max)	CR	AVE	Cronbach's alpha	R^2	Q^2
Relevance	1	1.000	1.000	1.000			
Benefits	4	0.798 - 0.866	0.897	0.686	0.848	0.192	0.114
Risks	3	0.761 - 0.862	0.858	0.669	0.752	0.349	0.204
Trust	1	1.000	1.000	1.000		0.187	0.177
Withholding	3	0.842 - 0.928	0.924	0.802	0.876	0.546	0.392
Falsification	1	1.000	1.000	1.000		0.210	0.162

FIGURE 3. RESEARCH MODEL RESULTS

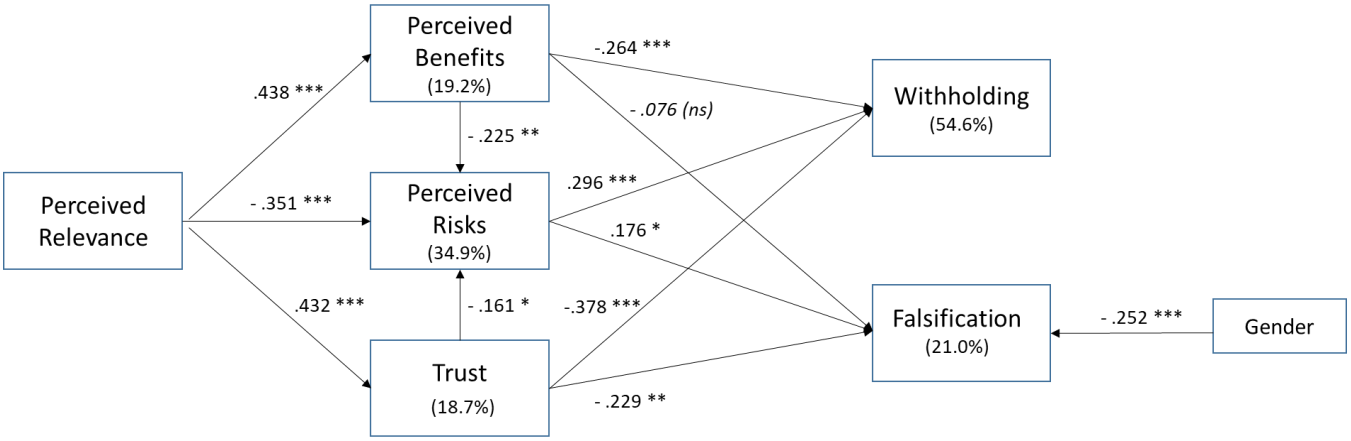


TABLE 4. TESTS OF HYPOTHESES

Hyp	Paths (from -> to)	Beta	St Dev	T Stat.	P-value	Result
H1a	Perceived Benefits -> Withholding	-0.264	0.067	3.932	***	**Supported**
H1b	Perceived Benefits -> Falsification	-0.076	0.089	0.854	> .05	Not Supported
H2a	Perceived Risks -> Withholding	0.296	0.063	4.714	***	**Supported**
H2b	Perceived Risks -> Falsification	0.176	0.085	2.063	*	**Supported**
H3a	Trust -> Withholding	-0.378	0.067	5.609	***	**Supported**
H3b	Trust -> Falsification	-0.229	0.092	2.485	**	**Supported**
H4	Perceived Benefits -> Perceived Risks	-0.225	0.075	3.000	**	**Supported**
H5	Trust -> Perceived Risks	-0.161	0.078	2.069	*	**Supported**
H6a	Perceived Relevance -> Perceived Benefits	0.438	0.073	5.990	***	**Supported**
H6b	Perceived Relevance -> Perceived Risks	-0.351	0.071	4.948	***	**Supported**
H6c	Perceived Relevance -> Trust	0.432	0.069	6.236	***	**Supported**
	Gender -> Falsification	-0.252	0.061	4.152	***	

LEGEND (two-tailed tests): * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

The results support 10 out of our 11 hypotheses (as H1 to H3 are decomposed in two for each DV and H6 has 3 subhypotheses). As we proposed, the perceived relevance of the data requested influences the perceived risks associated with the disclosure of those data (-0.351, $p < 0.001$), the perceived benefits to do so (0.438, $p < 0.001$) and the trust in the company asking for those data (0.432, $p < 0.001$), thus supporting H6 (a, b, and c). As also anticipated, both perceived benefits (- 0.225, $p < 0.01$) and trust (- 0.161, $p < 0.05$) significantly influence the perceived risks associated with data disclosure, thus supporting both H4 and H5.

In turn, all three mediating variables (perceived benefits, perceived risks, and trust) significantly influence the decision to withhold personal data (-0.264, $p < 0.001$, 0.296, $p < 0.001$, -0.378, $p < 0.001$, respectively), thus supporting H1a, H2a, and H3a. Trust is the most influential driver of withholding behavior followed by perceived risks (positive impact).[9] Regarding the second DV, contrary to our expectations, perceived benefits have no significant effect (-0.076, $p > 0.05$) on the tendency to falsify the data. However, both trust and perceived risks have significant relationships with falsification, with trust having a stronger negative effect (-0.229, $p < 0.01$) than the positive effect of perceived risks (0.176, $p < 0.05$). Taken together, these findings provide no support for H1b but validate both H2b and H3b.

Tests for Moderating Effects (P1)

To test P1 (moderating effects), we conducted two post hoc multigroup analyses (MGA) to compare our model based on the amounts of data requested to the applicants (low vs. high) and on the incentive (pecuniary vs. nonpecuniary) behind the data request (i.e., lottery vs. homepage) (see Appendix 10 for details on this analysis).

[9] Notably, in a workshop paper from a related but not isomorphic research domain, Treiblmaier and Chong (2007) found on the contrary that perceived risks were a much stronger predictor of willingness to disclose than trust.

TABLE 5. MGA RESULTS FOR BOTH CONTEXTUAL CUES (SIZE OF REQUEST AND INFERRED INCENTIVE OF THE REQUEST)

Hyp	Paths (from -> to)	MGA Results for Size of Request						MGA results for Inferred Incentive of the Data Request					
		β for Short Form	Sig	β for Long Form	Sig	p-value of the difference	Sig. diff. ?	β for Homepage Context	Sig	β for Lottery Context	Sig	p-value of the difference	Sig. diff. ?
H1a	Perceived Benefits -> Withholding	-0.197	*	-0.325	**	0.818	No	-0.256	**	-0.276	**	0.556	No
H1b	Perceived Benefits -> Falsification	-0.058		-0.067		0.472	No	-0.196		0.058		0.920	No
H2a	Perceived Risks -> Withholding	0.389	***	0.227	**	0.904	No	0.328	***	0.259	***	0.712	No
H2b	Perceived Risks -> Falsification	0.164		0.227	*	0.626	No	0.235	*	0.087		0.198	No
H3a	Trust -> Withholding	-0.303	**	-0.428	***	0.820	No	-0.332	**	-0.426	***	0.740	No
H3b	Trust -> Falsification	-0.165		-0.268	*	0.279	No	-0.062		-0.435	***	0.016	**Yes**
H4	Perceived Benefits -> Perceived Risks	-0.143		-0.320	**	0.119	No	-0.188		-0.246	*	0.342	No
H5	Trust -> Perceived Risks	-0.232	*	-0.096		0.805	No	-0.100		-0.212	**	0.241	No
H6a	Perceived Relevance -> Perceived Benefits	0.399	***	0.451	***	0.354	No	0.517	***	0.364	***	0.860	No
H6b	Perceived Relevance -> Perceived Risks	-0.384	***	-0.310	***	0.306	No	-0.413	***	-0.308	***	0.232	No
H6c	Perceived Relevance -> Trust	0.392	***	0.427	***	0.410	No	0.510	***	0.357	***	0.861	No
	Gender -> Falsification	-0.199	*	-0.329	***	0.153	No	-0.218	**	-0.315	***	0.211	No

LEGEND (two-tailed tests): * p < .05 ** p < .01 *** p < .001

Significant MGA difference (p < 0.05) when comparing two treatments corresponding to one contextual cue

Hypothesis verified for one treatment of the contextual cue but not the other although the MGA difference is not statistically significant (p > 0.05)

FIGURE 3

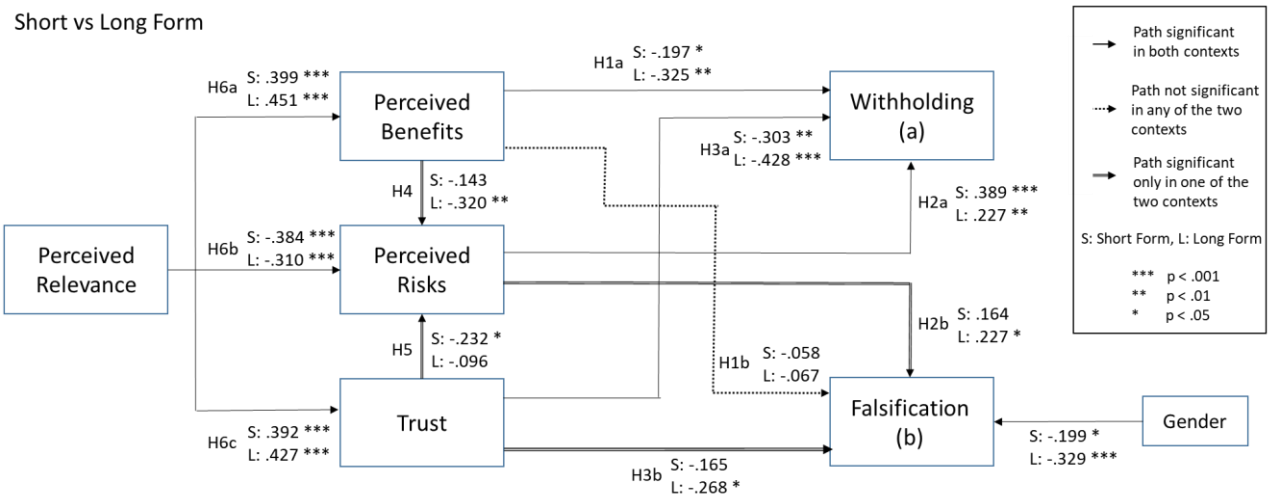
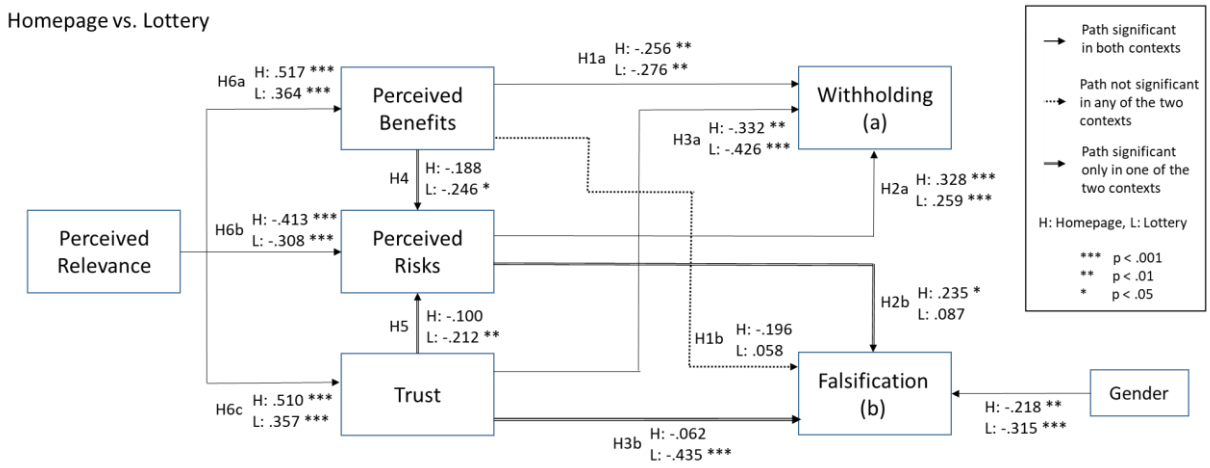


FIGURE 4



The results regarding the support for P1 are rather complex as 11 moderation effects are tested for two different contextual cues. The bottom-line results can be seen in Table 5. Figures 3 and 4 show the full model for each of the manipulated conditions in our 2x2 experiment. Globally, P1 is partially supported as we found some differences between the different treatments (i.e., hypotheses validated in one context but not the other).

According to the MGA results (Table 5), one statistically significant difference (underlined in dark gray) is noted. Interestingly, the effect of trust on falsification is significantly different when considering the incentive behind the data request ($p = 0.016$, dark gray) so that the effect holds in the lottery (i.e., pecuniary) context (-0.435, $p < 0.001$) but not in the homepage (i.e., nonpecuniary) context (-0.062, $p > 0.05$). In the lottery context, therefore, trust has a strong negative impact on falsification (i.e., the more people trust the company asking them personal details in order for them to participate in a lottery, the less they will tend to falsify). This does not seem to be the case in the homepage context where trust does not result in a reduced tendency to falsify.

Also noteworthy is the difference in the effect of trust on falsification when considering the amount of the data requested, although this difference is not statistically validated in the MGA results ($p = 0.279$, light gray). Here, the effect of trust holds when the form to be filled in is long (-0.268, $p < 0.05$) but not when only a small quantity of information is asked (- 0.165, $p > 0.05$). Therefore, when the amount of information to provide is high, trust has an important impact on the tendency to falsify, but this is no longer the case when the amount of requested information is small.

Three other hypotheses involving perceived risks as a driver or a consequence also yield some interesting results, although they do not show statistical significance in the MGA results (light gray). In particular, when looking at the influence of risks on falsification, we can see that this effect holds only for the long form (0.227, $p < 0.05$) and in the homepage context (0.235, $p < 0.05$). In both conditions, perceived risk results in a stronger tendency to falsify the data, which does not seem to happen in a lottery context or when the amount of requested information is small.

In addition, the impact of perceived benefits on perceived risks is validated only in the long form (- 0.320, $p < 0.01$) and in the lottery (- 0.246, $p < 0.05$) conditions. Finally, trust translates into reduced perceived risks only when the number of details requested is low (- 0.232, $p < 0.05$) and in the lottery (i.e., pecuniary) context (- 0.212, $p < 0.01$).

Globally, the drivers of withholding are affected by neither the amount of data requested nor the incentive (pecuniary vs. nonpecuniary) behind the data request, whereas the two drivers of falsification (i.e., trust and perceived risks) are impacted by both contextual cues (i.e., data quantity and inferred incentive). In addition, both antecedents of perceived risks are also significantly affected by the amount of information asked and the incentive (pecuniary vs. nonpecuniary) behind the data request.

DISCUSSION

This study makes three important contributions to the literature stream, each of which can be linked to one of research questions of the study as stated at the outset.

The first research question was "To what extent does a privacy tradeoff – based on perceived benefits, perceived risks, and trust – explain a) individuals' data disclosure decisions (i.e., withholding) and b) the phenomenon of individuals providing falsified data in privacy-related decisions?" First, it was seen that this model explained a significant portion (54.6%) of the variance in withholding. In addition, context matters in explaining withholding: with respect to the amount of data requested, in the "short form" condition, a reduction of risks is most likely to reduce withholding, and this is followed by an increase in trust. In the "long form" condition, trust is most important in reducing withholding, with perceived benefits following closely behind. With respect to the inferred incentive, in the lottery (pecuniary) context, trust is the most important driver, followed by perceived benefits. In the personalized homepage (nonpecuniary) context, trust and a reduction in perceived risks are the most important factors in reducing withholding. Second, as one of the first studies within the privacy domain to consider the notion of falsification as a DV, our explanation of

21% of the variance in that construct is notable and represents one of the main contributions of this study. We offer potential explanations for this phenomenon that will be worth considering in future privacy-related studies. In this area also, context matters, with the influence of trust on falsification being very different depending on the context of the data request (amount of data requested and incentive behind the request).

Of particular interest are the findings related to the influence of benefits on each of our DVs, especially their non-significant effect on falsification. In our model, we incorporated trust as an antecedent to privacy-related decisions, in addition to the classical privacy-calculus elements (benefit/risk trade-off). Our findings confirm trust as the most important driver of both withholding and falsification decisions. Specifically, falsification seems to be driven by a trust/risk trade-off, as opposed to the classical privacy calculus (benefit/risk) paradigm. Interestingly, when trust is taken out of the model, benefits, in turn, have a significant effect on falsification[10]. Trust thus plays a suppressor effect: when people trust how firms are handling their data, no matter how much benefit they will be given in exchange of the data, it will significantly reduce the chances that they provide falsified information. This finding complements previous literature on privacy, especially the work by Norberg et al. (2007) which already discussed the corresponding influence of trust and risks on disclosure behavior. Contrary to the popular paradigm that consider the calculus perspective of privacy (benefit/risk trade-off) as 'the most useful framework for analyzing contemporary consumer privacy concerns' (Culnan and Bies, p. 326; Xu et al. 2009, p. 139), this framework may only or mainly be useful when there is no (or not enough) trust in the firm collecting the data.

The second research question was "To what extent can individuals' perceptions of risks, benefits, and trust in privacy-related decisions be explained by the contextual factor of perceived relevance?" In contrast with many privacy studies that view perceived benefits, perceived risks, and even trust as purely exogenous variables, this study indeed considered a contextual antecedent to

[10] We sincerely thank the reviewing team for their relevant comment and helpful guidance in emphasizing this finding.

each of these constructs (i.e., perceived relevance) (see Figure 2). Because all these relationships proved to be significant, we confirm the interest of considering contextual antecedents to these key privacy-related decision drivers. In addition, we found that our context-specific antecedent (perceived relevance) influenced perceived benefits, risks, and trust in the four treatment conditions of our manipulated context, thus providing some generalizability to this context-specific driver.

The third, exploratory research question was "How does the context of a data request impact individuals' decisions regarding disclosure and falsification?" We found that our manipulated contextual cues had a moderating effect on four of the 11 hypotheses in our model. This provides suggestive evidence regarding the importance of contextual factors in decisions associated with data disclosure and falsification and indicates the need for additional research in this area.

Researchers should find that their future efforts can be enlightened by these findings regarding our three research questions.

Implications for Research and Future Research Avenues

As noted above, this study is the first to look across a nomological model of important drivers of privacy-related decisions while considering the fact that individuals may respond to data requests in various ways: by honestly supplying the requested information, by refusing to provide the information, or by providing false information. As such, it makes a substantive contribution to the literature stream on its own. However, its greatest contribution may be as a motivator for future research that builds on this nomological model and, especially, considers how other contextual factors may come into play. Additionally, future researchers can benefit by considering other cognitive processing models that go well beyond those considered in this study.

Context

Because this is one of the first studies to consider the option of falsification in individuals' response patterns, we intentionally limited our consideration of contextual factors to an exploratory

zone (Research Questions 2 and 3). Following the path outlined by Hong et al. (2014), we considered one contextual antecedent (perceived relevance) and the potential moderating effects of two manipulated variables (size and inferred incentive of the request). Even with this modest objective, we found the following:

- The context-specific antecedent (perceived relevance) has a very significant impact on perceived benefits, perceived risks, and trust in all manipulated conditions.
- For four of the 11 hypotheses in our model, relationships were moderated by manipulated contexts, although in only one of these conditions (inferred incentive of the data request as a moderator of the relationship between trust and falsification), the difference between both treatments (i.e., between the lottery [pecuniary incentive] and the homepage [nonpecuniary incentive]) was significant at the 0.05 level.

These findings suggest that in our exploratory contextual initiative, the contextual antecedent (perceived relevance) has proven to have a significant role in explaining individuals' response patterns. We therefore recommend that future researchers be especially mindful of its role when crafting their own models.

The findings from our work also suggest that although our initial foray into the moderating effects should be viewed as exploratory, we already show some interesting results that will need to be confirmed in future research. As can be seen in the "Moderators" column of Appendix 1, we are obviously engaged in a domain that has received little previous research attention. Of the 42 studies listed in the appendix, in our judgment, only six considered any contextual factors (based on the definition proffered earlier) as moderators (Anderson et al. 2011; Bansal et al. 2016; Karwatzki et al. 2017; Li et al. 2012; Li 2014; Zimmer et al. 2010b). Most importantly, only one of these six studies (Bansal et al. 2016) found evidence of significant moderating effects based on the contextual factors they tested. This points to a need for a deep focus in future research. Future studies could benefit from both additional attention to contextual antecedents and moderators ("Levels 2a and 2b," respectively, per Hong et al. (2014). We discuss each.

Contextual antecedents. As was noted earlier, this study has gone further than most studies in the privacy domain by considering one contextual antecedent to both perceived risks and perceived benefits in our model. While our approach was obviously a fruitful one (our model explained 35% and 19% of the variance in perceived risks and perceived benefits, respectively), it is obvious that additional work could be done to strengthen the explanatory structure associated with these two important privacy calculus inputs. For example, it might be possible to provide an even stronger explanation for perceived risk by considering other factors that simultaneously were being processed by a decision-maker. One might conjecture that all other things being equal, an individual might perceive more risk in a certain data-seeking situation if (s)he was already in an anxious affective state or under some level of time pressure (Petty et al. 1986). Regarding perceived benefits, we also embraced a somewhat simplistic approach in light of the paucity of past research in this domain, by including only one antecedent variable (perceived relevance) in our model. An expanded consideration of perceived benefits might well include an additional level of calculus that addressed both the size of probable benefits as well as the likelihood of those benefits accruing, thus inducing a stochastic modeling process. Such an approach would obviously deepen the model and also enable interesting experimental treatments in which, for example, decisions could be manipulated to test the impact of both size and probability in impacting benefits directly and other DVs indirectly.

Contextual moderators. In looking across the research stream as depicted in Appendix 1 "Moderators" column and considering this study, one can see that only a few contextual moderators have been tested to date:

- Industry/type of data (Anderson et al. 2011; Bansal et al. 2016; Li et al. 2012; Zimmer et al. 2010b)
- Website attributes (Karwatzki et al. 2017; Li 2014)
- Requester properties (Anderson et al. 2011)
- Amount of data – this study
- Intended use of data/disclosure incentives – (Anderson et al. 2011), this study.

Obviously, the most frequently considered moderator has been the type of information (sometimes implied by the industry). This is indeed an important area for consideration, as it is certainly conceivable that an individual's perceptions and responses across an entire research model could differ based on such a context. For example, consider again our introductory scenario regarding the Pizza Shack and a data request for consumers' birthdays. It is quite conceivable that, had a financial institution (rather than a pizza delivery service) been the requesting entity, individuals might have responded differently to a request for their birthday or, even more importantly, more sensitive information such as their income. Driving such a difference in response patterns could well have been different perceptions of benefits/risks, different levels of trust, etc.—that is, many of the components that comprised our basic decision-making model.[11]

Although these five areas seem to be those on which researchers have focused to date in assessing contextual moderating effects, the list is obviously far from an exhaustive one. Future research would benefit from a broader consideration of such potential moderators, which could well include constructs such as government regulation, industry norms, and technological alternatives (e.g., differing platforms and different devices).

Having noted the extreme importance of addressing context in future research, we now turn to some specific areas in which our present model—whether embraced as a base for future studies on its own or incorporated into broader models—merits attention: the role of trust, additional drivers for falsification, and additional consideration of demographic and experiential relationships.

Focal areas in the present model

The role of trust. Smith et al. (2011) noted that, while trust is widely viewed as an important construct in privacy research, its specific role in terms of its antecedents and direct/moderating effects has not been clearly established in the past. Based on some previous (in some cases,

[11] We gratefully acknowledge an anonymous reviewer for this insight regarding our opening scenario.

conflicting) findings and our own argumentation, we posited four hypotheses that included trust. For all four, our study did confirm the postulated direction, and in all cases, the relationship exhibited the strength associated with statistical significance. This suggests that our model's treatment of the trust construct was generally correct, but obviously much additional work is required to fully develop the relationships.

We recommend, therefore, that additional work be done to clarify the precise nature of the trust construct(s) that are salient in privacy-related decisions. We noted earlier that trust is usually viewed as a dyadic construct in psychological research, and most theoretical development therefore relies on relationships from that perspective. However, in IS privacy research, trust is usually conceptualized as an individual's trust in an entity that is requesting data. We suggest that this distinction needs even further clarification, as the components of privacy-related trust may be even more complex than in other areas addressed in the general IS research stream. For example, there may be some parameters associated with data types or combinations thereof that would be reflected in privacy-related trust but that would not be salient in other contexts. We can only speculate regarding the nature of all such dimensions, but the meager support for the trust-related hypotheses in this study – an outcome consistent with the mixed findings in the overall IS privacy research domain – suggests that much additional attention is needed.

Additional drivers for falsification. The importance of the falsification construct in privacy research cannot be overstated, as one phenomenon often cited by privacy observers is that of a "privacy paradox," in which individuals' stated that privacy concerns are inconsistent with their actual behavior (Smith et al. 2011). As was aptly noted by Keith et al. (2013, p. 1171), "[r]esearchers may mistakenly identify the privacy paradox phenomenon in their studies if their…design does not assess the accuracy of the data provided." In that spirit, we note that much more could be done to further our understanding of falsification behavior in response to data requests. Although the vast majority of research has been in an organizational (rather than online or consumer) context, the discipline of business ethics has devoted some attention to the factors associated with this

phenomenon (e.g., Acke et al. 2011; Bowie 2012; Grover et al. 1994; Takala et al. 1999). Additionally, some applied psychologists have considered some contextual factors associated with falsification through various communication media (Naquin et al. 2010). Obviously, much additional investigation would be required to expand our theoretical modeling of falsification behavior as a response to data requests, and even more work would be needed to test those models. We certainly acknowledge the nontrivial nature of further investigation of this phenomenon, but we also highlight the relatively unexplored nature of this important context in privacy research.

Additional consideration of demographic and experiential relationships. A few prior studies (e.g., Culnan et al. 1999; Sheehan et al. 1999) have noted demographic differences in individuals' perceptions and responses in privacy-related contexts. In this study, we found that one demographic variable (gender) had a significant effect on falsification. We are unaware, however, of any prior theoretical development that would explain such a difference. One positive interpretation of the very limited impact of demographic variables could be that our model is indeed a robust one. However, the development of strong theory to explain potential differences associated with gender and other demographic variables could prove a significant contribution to the research stream.

These three areas— the role of trust, additional drivers for falsification, and additional consideration of demographic and experiential relationships—are ripe for additional consideration in all situations in which our current model is embraced as a baseline. However, we also wish to highlight and question a basic assumption that undergirds not only most research to date in this particular domain but also in many other IS contexts: the level of cognitive effort that individuals devote to their decision-making processes.

Consideration of low-effort cognitive processing

Although not overtly mentioned in the published privacy articles to date, a covert assumption is that individuals who are engaging in privacy calculus are embracing what is viewed as "high effort" cognitive processing, in which they attempt to rationally evaluate the benefits and risks.

Indeed, a perusal of the article in Appendix 1 reveals no examples of studies that challenge this basic premise (Note that the level of effort in cognitive processing constitutes a decision-making state at the individual level. It is *not* an external contextual factor as defined earlier). Yet there are many situations in which individuals instead embrace "low effort" cognitive processing routes; for example, individuals in happy moods are much more likely to expend low amounts of cognitive effort and instead to rely on heuristics and biases in making decisions (Petty et al. 1986). Within the larger privacy research stream, a few authors (e.g., Acquisti et al. 2005; Li et al. 2010; Li et al. 2008; Wakefield 2013; Yu et al. 2015) have begun to investigate privacy-related decisions under conditions associated with low-effort processing (Dinev et al. 2015).

Chaiken (1978; 1980)'s heuristic-systematic model and Petty and Wegener (1998)'s elaboration likelihood model (ELM) both distinguish more and less effortful information processing and decision-making—with the ELM terming the high-effort route as "central" and the low-effort route as "peripheral." For any given decision, an individual can be expected to rely on a mixture of high and low effort cognitive processing. In a privacy-related context, high-effort processing is characterized by "responses to external stimuli result[ing] in deliberate analyses, which lead to fully informed privacy-related attitudes and behaviors" (Dinev et al. 2015, pp. 641-642). Low-effort processing is characterized by "relatively little cognitive effort or conscious awareness" and a "relatively greater reliance on automatic heuristics" (Dinev et al. 2015, p. 642). Such heuristics may be grounded in, or accompanied by, peripheral cues, biases, and misattributions.[12]

The relative apportionment of cognitive resources to the low- and high-effort routes is determined by factors such as affect (mood and emotion), cognitive resources, motivation, and time constraints. This suggests that researchers could plow a fruitful research path by conducting a series of experiments that manipulate subjects' relative allocation of resources to low- and high-effort processes, with measurement of cognitive, perceptual, and behavioral responses associated with

[12] Such heuristics, peripheral cues, biases, and misattributions may also be factors in high-effort processing, but they are much more likely to drive decision-making in low-effort situations.

each. This could be done within the context of basic model of this study, although it may well be found that, especially for decisions that are driven primarily by low-effort processes, some of the constructs in our model may become inoperative. For example, rather than form conscious perceptions of risks and benefits, individuals who are driven primarily by low-effort processing may instead make disclosure/falsification decisions based primarily on heuristics or biases, thus bypassing several components of the basic model.[13] Obviously, future studies focused on disclosure and/or falsification would benefit greatly from a consideration of both low-effort and high-effort cognitive processing.

Implications for Practice

In addition to the implications for the research stream (above), this study has significant implications for practice in several areas. We highlight five.

First, those who request personal data from individuals should be aware that many of the same factors that drive data disclosure also—in converse—drive falsification of data. Thus, data requesters should be wary of accepting self-reported data from individuals at face value. In some cases, it may be important to cross-validate certain data elements (for example, some credit reporting firms are now offering instantaneous identity verifications for a small fee). It is also worth noting, however, that falsification of data was seen to decline significantly when individuals trust the firm requesting the data and perceive that the risks have been mitigated. This suggests that entities may profitably invest both in education (regarding risks) and in data protection (along with marketing thereof). Our study on drivers of data falsification provides a meaningful way for organizations to be aware of the possibility of false data during data collection along with some factors that might mitigate it (e.g., enhanced trust and reduced risks). Interestingly, our results show that the impact of

[13] Although a description of the precise mechanics of such experiments goes well beyond the scope of the present discussion, researchers who are interested in manipulation of affect may wish to consult resources such as Martin (1990), Schwarz et al. (1983), Sutherland et al. (1982), Velten (1968), Westerman et al. (1996), and researchers who wish to consider manipulation of cognitive resources may wish to consult resources such as Ainsworth et al. (2014), Gino et al. (2011), Pohl et al. (2013), and Schmeichel (2007).

each of those two factors might differ significantly depending on the situation in which data are requested (size of request and inferred incentive in our case). Organizations should therefore adapt their investments and communicate differently depending on the contexts in which they request data from their customers.

Second, the fact that trust affects privacy-related decisions such as withholding and falsification, although not entirely new (at least regarding its influence on disclosure), is an important finding. This suggests that firms should be encouraged to invest resources in initiatives that lead to a general enhancement of trust across society. Such initiatives may include industry codes of conduct, privacy seals, consumer education, or even lobbying for regulations that provide a baseline level of protection for all consumers. To the extent that trust is enhanced across the consumer spectrum, all organizations that rely on consumer data will benefit.

As the European Union's General Data Protection Regulation (GDPR) took effect in mid-2018, many of its legally mandated actions for data collection and handling are forcing European firms that handle citizens' data to embrace steps that will improve trust levels (Tankard 2016). Even U.S. firms that handle Europeans' data (e.g., credit card transaction processors) are being impacted (Gilbert 2016) and will need to consider registering under the EU-U.S. "Privacy Shield" program, which will dictate most of the same provisions (Downes 2016).

Third, as can be seen in Table 5, the effect of trust on falsification depends both on the size of the data request and on the inferred incentive for the request, such that higher trust levels are associated with reduced falsification only in a lottery context and when the quantity of requested data asked is large. The influence of perceived risks on falsification shows a similar although slightly different pattern: a reduced level of perceived risks is associated with reduced falsification when personalization is offered in exchange for the data and when the quantity of requested data is large. The same is true for withholding, with its drivers having different weights depending on the context. For example, in the lottery and long form contexts, withholding is driven first by trust followed by benefits, whereas for a short form, a reduction of risks is the most likely to reduce withholding. These

complex relationships suggest that initiatives to maintain or enhance trust or to reduce the perceived risks associated with data disclosure must be formulated carefully; practitioners should consider these boundary conditions to maximize their initiatives' efficiency.

Fourth, the (non)influence of perceived benefits is noteworthy. One might conjecture that people falsify data because they want to reap the benefits out of disclosure while minimizing the associated risks. This does not seem to be totally the case, however, as perceptions of more benefits do not lead to a reduction in falsification. Therefore, although proposing interesting benefits to consumers can certainly increase their disclosures, it will not reduce the risk of data falsification. Practitioners' initiatives to enhance benefits should thus be considered carefully to avoid misleading returns.

Fifth, relevance of the data matters. Big data principles could lead practitioners to try to collect as much data as possible from their customers, but consumers care about the legitimacy of each granular data request. In many cases, less data may lead to more disclosure and, more importantly, to a higher level of accuracy.

Potential limitations

While this study makes significant contributions to the literature, there are three potential limitations, though we argue that none of them stands as a significant threat to the validity of the study.

First, because our sampling relied on a panel research firm, we were constrained from gathering identifiable data–and thus real disclosure–from the subjects. We therefore measured their stated intentions, rather than their actual behavior, in our model. This approach is fairly common in the privacy research stream, and the distinct differences that were observed in stated intentions suggest that findings of this study are robust ones (see Table 5).

Second, in an attempt to minimize the length of our data collection time with subjects, we relied on single-item measures for some of the constructs in our model. Although there are support

(Bergkvist et al. 2007) and precedent (e.g., certain constructs in Culnan et al. 1999; Hui et al. 2007; Smith et al. 1996) for this approach, we do acknowledge that multi-item measures are frequently used by many IS researchers. This choice was constrained by our partnership with the access panel company and our desire to capture participants' behavior and decision-making process in a real-life situation. Given the observed validity of our measurement model, this is not a significant limitation, but we do recommend that researchers who extend this research stream consider devoting additional attention to the measurement of the constructs in the model.

Third, the solicitation of our sample from a single country (France) may be argued by some to limit the generalizability of the results. However, the fact that the subjects were recruited through an online panel, that they exhibited broad demographic dispersion, and that they were randomly assigned to different treatment conditions, all lend credence to the generalizability of the results. While it might be fruitful for the study to be replicated with a worldwide sample, the strong relationships that were demonstrated in the model, coupled with the fact that objectives of the study did not include an evaluation of cross-cultural differences, suggest that the use of subjects from a single country was not a substantive limitation.

CONCLUSION

In this study, we attempted to move the privacy research domain forward by demonstrating the importance of considering not only individuals' data disclosures but also falsification thereof. We hope that other researchers will benefit from our steps and will now work with us to extend this important subdomain of IS privacy research.

SPECIAL NOTE

This article is dedicated to H. Jeff Smith and his family.

REFERENCES

Acke LF, Chu BK, Kuang X, and Qi L (2011) Lying: An experimental investigation of the role of situational factors. *Business Ethics Quarterly* **21**(**4)**, 605-632.

Acquisti A and Grossklags J (2005) Privacy and rationality in individual decision making. *IEEE Security & Privacy* **3**, 26-33.

Acquisti, A., John, L. K., & Loewenstein, G. (2013). What is privacy worth?. *The Journal of Legal Studies*, *42*(2), 249-274.Ainsworth SE, Baumeister RF, Ariely D, and Vohs KD (2014) Ego depletion decreases trust in economic decision making. *Journal of Experimental Social Psychology* **54**, 40-49.

Anderson CL and Agarwal R (2011) The digitalization of healthcare: Boundary risks, emotion, and consumer willingness to disclose personal health information. *Information Systems Research* **22**(**3)**, 469-490.

Angst CM and Agarwal R (2009) Adoption of electronic health records in the presence of privacy concerns: The elaboration likelihood model and individual persuasion. *MIS Quarterly* **33**(**2)**, 339-370.

Awad NF and Krishnan MS (2006) The personalization privacy paradox: An empirical evaluation of information transparency and the willingness to be profiled for personalization. *MIS Quarterly* **30**(**1)**, 13-28.

Bansal G, Zahedi FM, and Gefen D (2010) The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decision Support Systems* **49**(**2)**, 138-150.

Bansal G, Zahedi FM, and Gefen D (2015) The role of privacy assurance mechanisms in building trust and the moderating role of privacy concern. *European Journal of Information Systems* **24**, 624-644.

Bansal G, Zahedi FM, and Gefen D (2016) Do context and personality matter? Trust and privacy concerns in disclosing private information online. *Information & Management* **53**(**1)**, 1-21.

Bergkvist L and Rossiter JR (2007) The predictive validity of multiple-item versus single-item measures of the same constructs. *Journal of Marketing Research* **44**(**2)**, 175-184.

Bowie NE (2012) Lying and deception. *Business Ethics Quarterly* **22**(**3)**, 579-585.

Buttner OB and Goritz AS (2008) Perceived trustworthiness of online shops. *Journal of Consumer Behavior* **7**(**1)**, 35-50.

Cavuscoglu H, Phan TQ, Cavuscoglu H, and Airoldi EM (2016) Assessing the impact of granular privacy controls on content sharing and disclosure on Facebook. *Information Systems Research* **27**(**4)**, 848-879.

Chai S, Sanjukta D, and Rao HR (2011-2012) Factors affecting bloggers' knowledge sharing: An investigation across gender. *Journal of Management Information Systems* **28**(**3)**, 309-341.

Chaiken S (1978) *The use of source versus messaeg cues in persuasion: An information processing analysis*, University of Massachusetts Amherst, MA.

Chaiken S (1980) Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology* **39**, 752-766.

Chakraborty R, Vishik C, and Rao HR (2013) Privacy preserving actions of older adults on social media: Exploring the behavior of opting out of information sharing. *Decision Support Systems* **55**, 948-956.

Chen R (2013) Living a private life in public social networks: An exploration of member self-disclosure. *Decision Support Systems* **55**, 661-668.

Chen R and Sharma SK (2015) Learning and self-disclosure behavior on social networking sites: The case of Facebook users. *European Journal of Information Systems* **24**, 93-106.

Chin WW (1998) Commentary: Issues and opinion on structural equation modeling. *MIS Quarterly* **22**(**1)**, vii-xvi.

Chin WW (2010) How to write up and report PLS analyses in *Handbook of partial least squares*. V. E. Vinzi, W. W. Chin, J. Henseler and H. Wang (eds.), Springer, Berlin, 655-690.

Choi BCF and Land L (2016) The effects of general privacy concerns and transactional privacy concerns on Facebook apps usage. *Information & Management* **53**(**7)**, 868-877.

Chwelos P, Benbasat I, and Dexter AS (2001) Research report: Empirical test of an EDI adoption model. *Information Systems Research* **12**(**3)**, 304-321.

Crossler R and Posey C (2017) Robbing Peter to pay Paul: Surrendering privacy for security's sake in an identity ecosystem. *Journal of the Association for Information Systems* **18**(**7)**, 487-515.

Culnan MJ and Armstrong PK (1999) Information privacy concerns, procedural fairness and impersonal trust: An empirical investigation. *Organization Science* **10**(**1)**, 104-115.

Culnan MJ and Bies RJ (2003) Consumer privacy: Balancing economic and justice considerations. *Journal of social issues* **59(2)**, 323-342.

Dinev T and Hart P (2006) An extended privacy calculus model for e-commerce transactions. *Information Systems Research* **17**(**1)**, 61-80.

Dinev T, Hart P, and Mullen MR (2008) Internet privacy concerns and beliefs about government surveillance - an empirical investigation. *Journal of Strategic Information Systems* **17**, 214-233.

Dinev T, McConnell AR, and Smith HJ (2015) Informing privacy research through information systems, psychology, and behavioral economics: Thinking outside the "APCO" box. *Information Systems Research* **26**(**4)**, 639-655.

Dinev T, Xu H, Smith HJ, and Hart P (2013) Information privacy and correlates: An empirical attempt to bridge and distinguish privacy-related concepts. *European Journal of Information Systems* **22**, 295-316.

Downes L (2016) The business implications of the EU-US. "Privacy shield". https://cb.hbsp.harvard.edu/cbmp/product/H02NUX-PDF-ENG accessed on December 8, 2017.

Finch J (1987) The vignette technique in survey research. *Sociology* **21**(**1)**, 105-114.

Fornell C and Larcker DF (1981) Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research* **18**(**1)**, 39-50.

Gefen D and Straub D (2005) A practcial guide to factorial validity using PLS-Graph: Tutorial and annotated example. *Communications of the Association for Information Systems* **16**(**5)**, 91-109.

Geisser S (1974) A predictive approach to random effect model. *Biometrika* **61**(**1)**, 101-107.

Gerlach J, Wiljaja T, and Buxmann P (2015) Handle with care: How online social network providers' privacy policies impact users' information sharing behavior. *Journal of Strategic Information Systems* **24**(**33-43)**.

Gilbert F (2016) EU general data protecton regulation: What impact for businesses established outside the European Union. *Journal of Internet Law* **19**(**11)**, 3-8.

Gino F, Schweitzer ME, Mead NL, and Ariely D (2011) Unable to resist temptation: How self-control depletion promotes unethical behavior. *Organizational Behavior and Human Decision Processes* **115**(**2)**, 191-203.

Goes P (2013) Information systems research and behavioral economics. *MIS Quarterly* **37**(**3)**, 3-8.

Grover SL and Hui C (1994) The influence of role conflict and self-interest on lying in organizations. *Journal of Business Ethics* **13**(**4)**, 295-303.

Hair JF, Hult GTM, Ringle CM, and Sarstedt M (2017) *A primer on partial least squares structural equation modeling.* (2nd ed.) Sage Publications,Thousand Oaks, CA.

Hair JF, Ringle CM, and Sarstedt M (2011) PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice* **19**(**2)**, 139-152.

Henseler J, Dijkstra TK, Ssarstedt M, Ringle CM, Diamantopoulos A, Straub DW, and Calantone RJ (2013) Common beliefs and reality about PLS: Comments on Ronkko and Evermann *Organizational Research Methods* **17**(**2)**, 182-209.

Henseler J, Ringle CM, and Sarstedt M (2015) A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science* **43**(**1)**, 115-135.

Hodder M, Churchill E, and Cobb J (2013) Lying and hiding in the name of privacy. customercommons.org/research, accessed on December 5, 2017.
Hong W, Chan FKY, Thong JYL, Chasalow LC, and Dhillon G (2014) A framework and guidelines for context-specific theorizing in information systems research. *Information Systems Research* **25**(**1)**, 111-136.
Hu LT and Bentler PM (1998) Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecificiation. *Psychological Methods* **3**(**4)**, 424.
Hui KL, Teo HH, and Lee SYT (2007) The value of privacy assurance: An exploratory field experiment. *MIS Quarterly* **31**(**1)**, 19-33.
James TL, Warkentin M, and Collignon SE (2015) A dual privacy decision model for online social networks. *Information & Management* **52**, 893-908.
Jiang Z, Heng CS, and Choi BCF (2013) Research note - privacy concerns and privacy-protective behavior in synchronous online social interactions. *Information Systems Research* **24**(**3)**, 579-595.
Kanuk, L., & Berenson, C. (1975). Mail surveys and response rates: A literature review. *Journal of Marketing Research*, 440-453. Karwatzki S, Dytynko O, Trenz M, and Veit D (2017) Beyond the personalization-privacy paradox: Privacy valuation, transparency features, and service personalization. *Journal of Management Information Systems* **34**(**2)**, 369-400.
Kehr F, Kowatsch T, Wentzel D, and Fleisch E (2015) Blissfully ignorant: The effects of general privacy concerns, general institutional trust, and affect in the privacy calculus. *Information Systems Journal* **25**, 607-635.
Keith MJ, Babb JS, Lowry PB, Furner CP, and Abdullat A (2015) The role of mobile-computing self-efficacy in consumer information disclosure. *Information Systems Journal* **25**, 637-667.
Keith MJ, Thompson SC, Hale J, Lowry B, and Greer C (2013) Information disclosure on mobile devices: Re-examining privacy calculus with actual user behavior. *International Journal of Human-Computer Studies* **71**(**12)**, 1163-1173.
Knijnenburg BP, Kobsa A, and Jin H (2013) Counteracting the negative effect of form auto-completion on the privacy calculus. Thirty Fourth International Conference on Information Systems Milan, 2013.
Kordzadeh N and Warren J (2017) Communicating personal health information in virtual health communities: An integration of privacy calculus model and affective commitment *Journal of the Association for Information Systems* **18**(**1)**, 45-81.
Krasnova H, Veltri NF, and Gunther O (2012) Self-disclosure and privacy calculus on social networking sites: The role of culture. *Business & Information Systems Engineering* **4**(**3)**, 1.
Lancelot Miltgen C and Peyrat-Guillard D (2014) Cultural and generational influences on privacy concenrs: A qualitative study in 7 European countries. *European Journal of Information Systems* **23**(**1)**, 103-125.
Li H, Luo X, and Xu H (2017) Resolving the privacy paradox: Toward a cognitive appraisal and emotion approach to online privacy behaviors. *Information & Management* **In press**.
Li H, Sarathy R, and Xu H (2010) Understanding situational online information disclosure as a privacy calculus. *Journal of Computer Information Systems* **51**(**1)**, 62-71.
Li H, Sarathy R, and Zhang J (2008) The role of emotions in shaping consumers' privacy beliefs about unfamiliar online vendors. *Journal of Information privacy and Security* **4**(**3)**, 36-62.
Li K, Lin Z, and Wang X (2015) An empirical analysis of users' privacy disclosure behaviors on social network sites. *Information & Management* **52**, 882-891.
Li T and Unger T (2012) Willing to pay for quality personalization? Trade-off between quality and privacy. *European Journal of Information Systems* **21**, 621-642.
Li Y (2014) The impact of disposition to privacy, website reputation and website familiarity on information privacy concerns. *Decision Support Systems* **57**, 343-354.
Liang H, Saraf N, Hu Q, and Xue Y (2007) Assimilation of enterprise systems: The effect of institutional pressures and the mediating role of top management. *MIS Quarterly* **31**(**1)**, 59-87.

Lima J (2015) Lying consumers could destroy big data. June 10, 2015 (available at http://www.cbronline.com/news/internet-of-things/consumer/lying-consumers-could-destroy-big-data-4597794).
Liu C, Marchewka JT, Lu J, and Yu CS (2004) Beyond concern: A privacy-trust-behavioral intention model of electronic commerce. *Information & Management* **42**(**1)**, 127-142.
Lobel B (2015) Quality of data suffers as consumers are reluctant to disclose personal information. *smallbusiness.co.uk*.
Lowry PB, Cao J, and Everard A (2011) Privacy concerns versus desire for interpersonal awareness in driving the use of self-disclosure technologies: The case of instant messaging in two cultures. *Journal of Management Information Systems* **27**(**4)**, 163-200.
Lowry PB, Romano NC, Jenkins JL, and Guthrie RW (2009) The CMC interactivity model: How interactivity enhances communication quality and process satisfaction in lean-media groups. *Journal of Management Information Systems* **26**(**1)**, 155-196.
Malheiros M, Preibusch S, and Sasse MA (2013) "Fairly truthful": The impact of perceived effort, fairness, relevance, and sensitivity on personal data disclosure. *Trust and Trustworthy Computing: Lecture Notes in Computer Science* **7904**, 250-266.
Malhotra KN, Kim SS, and Agarwal J (2004) Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research* **15**(**4)**, 336-355.
Martin M (1990) On the induction of mood. *Clinical Psychology Review* **10**(**6)**, 669-697.
McKnight DH, Choudhury V, and Kacmar C (2002) Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research* **13**(**3)**, 334-359.
Metzger MJ (2006) Effects of site, vendor, and consumer characteristics on web site trust and disclosure. *Communication Research* **33**(**3)**, 155-179.
Naquin CE, Belkin LY, and Kurtzberg TR (2010) The finer points of lying online: E-mail versus pen and paper. *Journal of Applied Psychology* **95**(**2)**, 387-394.
Nicolaou AI and McKnight DH (2006) Perceived information quality in data exchanges: Effects on risk, trust, and intention to use. *Information Systems Research* **17**(**4)**, 332-351.
Nissenbaum H (2014) Respect for context as a benchmark for privacy online: What it is and isn't in *The futures of privacy*. C. Dartiguepeyrou (ed.), Fondation Télécom, Institut Mines-Télécom, 19-30.
Norberg PA, Horne DR, and Horne DA (2007) The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs* **41(1)**, 100-126.
Ozdemir ZD, Smith HJ, and Benamati JH (2017) Antecedents and outcomes of information privacy concerns in a peer context: An exploratory study. *European Journal of Information Systems* **26**, 642-660.
Petty R and Cacioppo JT (1986) *Communication and persuasion: Central and peripheral routes to attitude change.* Springer-Verlag, New York.
Petty R and Wegener D (1998) *Attitude change: Multiple roles of persuasion variables.* (4 ed.) McGraw-Hill, New York.
Podsakoff P, MacKenzie S, Lee J, and Podsakoff N (2003) Common method bias in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology* **88**(**5)**, 879-903.
Podsakoff PM, MacKenzie SB, and Podsakoff NP (2012) Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology* **63**, 539-569.
Pohl RF, Erdfelder E, Hilbig BE, Liebke L, and Stahlberg D (2013) Effort reduction after self-control depletion: The role of cognitivie resources in use of simple heuristics. *Journal of Cognitive Psychology* **25**(**3)**, 267-276.
Posey C, Lowry PB, Roberts TL, and Ellis TS (2010) Proposing the online community self-disclosure model: The case of working professsionals in France and the U.K. who use online communities. *European Journal of Information Systems* **19**, 181-195.
Ringle CM, Wende S, and Becker JM (2015) Smartpls 3. SmartPLS GmbH, Boenningstedt.

Robert LP, Dennis AR, and Hung Y-TC (2009) Individual swift trust and knowledge-based trust in face-to-face and virtual team members. *Journal of Management Information Systems* **26**(**2)**, 241-279.

Schmeichel BJ (2007) Attention control, memory updating, and emotion regulation temporarily reduce the capacity for executive control. *Journal of Experimental Psychology: General* **136**(**2)**, 241-255.

Schwaig KS, Segars AH, Grover V, and Fiedler KD (2013) A model of consumers' perceptions of the invasion of information privacy. *Information & Management* **50**, 1-12.

Schwarz N and Clore GL (1983) Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology* **45**(**3)**, 513-523.

Sheehan KB and Hoy M (1999) Flaming, complaining, abstaining: How online users respond to privacy concerns. *Journal of Advertising* **28**(**3)**, 37-51.

Shih H-P, Lai K-H, and Cheng TCE (2017) Constraint-based and dedication-based mechanisms for encouraging online self-disclosure: Is personalization the only thing that matters? *European Journal of Information Systems* **26**, 432-450.

Smith H, Dinev T, and Xu H (2011) Information privacy research: An interdisciplinary review. *MIS Quarterly* **35**(**4)**, 989-1015.

Smith HJ, Milberg JS, and Burke JS (1996) Information privacy: Measuring individuals' concerns about organizational practices. *MIS Quarterly* **20**(**2)**, 167-196.

Son JY and Kim SS (2008) Internet users' information privacy-protective responses: A taxonomy and a nomological model. *MIS Quarterly* **32**(**3)**, 503-529.

Spiekermann S, Krasnova H, Koroleva K, and Hildebrand T (2010) Online social networks: Why we disclose. *Journal of Information Technology* **25**(**2)**, 109-125.

Stone M (1974) Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B (Methodological)* **36**(**2)**, 111-147.

Suh B and Han I (2003) The impact of customer trust and perception of security control on the acceptance of electronic commerce. *International Journal of Electronic Commerce* **7**(**3)**, 135-161.

Sutherland G, Newman B, and Rachman S (1982) Experimental investigations of the relations between mood and inensive unwanted cognitions. *British Journal of Medical Psychology* **55**(**2)**, 127-138.

Takala T and Urpilainen J (1999) Managerial work and lying: A conceptual framework and an explorative case study. *Journal of Business Ethics* **20**(**3)**, 181-195.

Tankard C (2016) What the gdpr means for businesses. *Network Security* **2016**(**6)**, 5-8.

Tow WN-FH, Dell P, and Venable J (2010) Understanding information disclosure behaviour in australian facebook users. *Journal of Information Technology* **25**, 126-136.

Treiblmaier H and Chong S (2007) Antecedents of the intention to disclose personal information on the internet: A review and model extension Sixth Annual Workshop on HCI Research in MIS, Montreal, 2007, 30-34.

Velten E (1968) A laboratory task for induction of mood states. *Behaviour Research and Therapy* **6**(**4)**, 473-482.

Wakefield R (2013) The influence of user affect in online information disclosure. *The Journal of Strategic Information Systems* **22**(**2)**, 157-174.

Westerman R, Spies K, Stahl G, and Hesse FW (1996) Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology* **26**(**4)**, 557-580.

Wirtz J, Lwin MO, and Williams JD (2007) Causes and consequences of consumer online privacy concern. *International Journal of Service Industry Management* **18**(**4)**, 326-348.

Xie E, Teo HH, and Wan W (2006) Volunteering personal information on the internet: Effects of reputation, privacy notices, and rewards on online consumer behavior. *Marketing Letters* **17**(**1)**, 61-74.

Xu H, Teo H-H, Tan BCY, and Agarwal R (2009) The role of push-pull technology in privacy calculus: The case of location-based services. *Journal Of Management Information Systems* **26**(**3)**, 135-174.

Yu J, Hu PJH, and Cheng TH (2015) Role of affect in self-disclosure on social network websites: A test of two competing models. *Journal of Management Information Systems* **32**(**2)**, 239-277.

Zimmer JC, Arsal R, Al-Marzouq MM, D., and Grover V (2010a) Knowing your customers: Using a reciprocal relationship to enhance voluntary information disclosure. *Decision Support Systems* **48**, 395-406.

Zimmer JC, Arsal RE, Al-Marzouq M, and Grover V (2010b) Investigating online information disclosure: Effects of information relevance, trust and risk. *Information & Management* **47**, 115-123.

APPENDIX 1

Articles in Top IS Journals with Disclosure/Falsification to an Organizational Entity as Dependent Variable[14]

Reference	**Salient dependent variables (DVs)**	**Antecedents to DVs (sig. indicated by *): n-1:[15] Immediate; n-2: Secondary (if any); n-3: Tertiary (if any); n-4: Quaternary (if any)**	**Moderators considered (if any)**	**Theoretical framework(s) relied upon (if any)**
(Anderson et al. 2011)	Willingness to provide access to personal health info (implied disclosure)	n-1: Health status emotion *	Type of info, intended purpose, requesting stakeholder [16]	Privacy boundary theory, risk as feelings
(Angst et al. 2009)	Opt-in intention to adopt electronic health records (implied disclosure)	n-1: Concern for Information Privacy *, post-attitude * n-2: Argument frame *, issue involvement *	Concern for information privacy	Elaboration likelihood model

[14] The 5-year citation counts for each journal in the categories "Computer Science, Information Systems" and "Information Systems & Library Science" for 2016 were downloaded from the Web of Science, Social Science Citation Index (SSCI) on November 14, 2017. Each list was sorted by the 5-year citation counts, largest to smallest. The top quintile of each such list was then considered for the "top" journals for that list. From each of those lists, the authors excerpted those journals that, in their opinion, are customarily targeted by IS researchers and combined these into the list above. This resulted in a "Top 10" list of journals, as follows*: Decision Support Systems* (DSS), *European Journal of Information Systems* (EJIS), *Information & Management* (I&M), *Information Systems Journal* (ISJ), *Information Systems Research* (ISR), *Journal of AIS* (JAIS), *Journal of Information Technology* (JIT), *Journal of MIS* (JMIS), *Journal of Strategic Information Systems* (JSIS), and *MIS Quarterly* (MISQ). Within these journals, we searched for articles (with no limitation on publication dates) that included the words "privacy"/"disclosure" and/or "falsification"/"falsify"/"misrepresent"/"misrepresentation"/"lying"/"lie" in the title, abstract, or (when available) author-supplied keywords. Salient articles that resulted from these searches are included in this table; we have included only those articles that documented empirical studies related to data subjects' disclosure of information about themselves to organizational entities and that derived and/or tested models related thereto.

[15] The terms "n-1," "n-2," etc., refer to the logical distance from the dependent variable. n-1 variables have direct relationships to the DV, which in some cases (when n-2 variables are also included in the model) has them serve as mediators. Similar relationships hold at the n-2 and n-3 levels.

[16] Underlined items are those that, in our estimation, represent contextual factors as defined herein. Specifically, they are largely grounded in the "environment external to the individual" (Bansal et al. 2016, p. 2).

Reference	Salient dependent variables (DVs)	Antecedents to DVs (sig. indicated by *): n-1:[15] Immediate; n-2: Secondary (if any); n-3: Tertiary (if any); n-4: Quaternary (if any)	Moderators considered (if any)	Theoretical framework(s) relied upon (if any)
(Awad et al. 2006)	Intention – willingness to be profiled online for personalized service/advertising (implied disclosure)	n-1: Previous online privacy invasion (mixed *), privacy concern, importance of privacy policies *, importance of information transparency * n-2: Demographics (gender, education, income), previous online privacy invasion *, privacy concern *, importance of privacy policies *		Utility maximization (akin to privacy calculus) [a]
(Bansal et al. 2010)	Intention to disclose health information	n-1: Health info privacy concern *, trust in the health website [b] *, prior positive experience with the website * n-2: Health info privacy concern, prior positive experience with the website *, risk beliefs health info * n-3: Perceived health info sensitivity *, previous online privacy invasion* n-4: Personality (5 factors) (mixed *), poor health status *		Utility theory (akin to privacy calculus) [a]
(Bansal et al. 2016)	Intention to disclose information	n-1: - Trust in the website [b] * - Positive experience with the website * - Internet privacy concerns * n-2: - Positive experience with the website * - Internet privacy concerns (mixed *) - Previous online privacy invasion * - Personality (Big 5) (mixed *)	Industrial context (health, finance, e-commerce website)	The contextualization of the theory of reasoned action and its synthesis with prospect theory

Reference	Salient dependent variables (DVs)	Antecedents to DVs (sig. indicated by *): n-1:[15] Immediate; n-2: Secondary (if any); n-3: Tertiary (if any); n-4: Quaternary (if any)	Moderators considered (if any)	Theoretical framework(s) relied upon (if any)
(Bansal et al. 2015)	Intention to disclose	n-1: Trust (in the website) [b] * n-2: - Argument quality/Adequacy of a privacy policy statement (regarding collection, errors, secondary use, improper access) * - Peripheral cues (website info quality, availability of company info, design appeal, reputation) *	Privacy concerns	Elaboration likelihood model
(Cavuscoglu et al. 2016)	Facebook content sharing and disclosure (through wall posts, private messages)	n-1: Facebook privacy policy change *, growth in Facebook friendship network (mostly *)		Privacy calculus, communication privacy management [a]
(Chai et al. 2011-2012)	Self-reported knowledge-sharing behaviors on blogs (disclosures)	n-1: Social ties *, reciprocity *, trust [b] *, information privacy concerns n-2: reciprocity (on social ties) *, social ties on trust *, trust on info privacy concerns *		Social capital theory
(Chakraborty et al. 2013)	Privacy-preserving and sharing actions on Facebook	n-1: Gender (mixed *), friends' behaviors *		Social capital theory, activity theory of aging, social role theory
(Chen 2013)	Privacy self-disclosure behaviors (self-reported)	n-1: Attitude * n-2: Extroversion *, Perceived critical mass *, Perceived Internet risk *	Privacy	Information disclosure behavior, expectancy-value theory
(Chen et al. 2015)	Self-disclosure extent (declared behavior)	n-1: SNS usage rate * n-2: Attitude toward using the SNS * n-3: Extroversion *, Perceived Networking Assistance *, Perceived Cyber Risk *, Social Influence *	Gender	Learning theories

Reference	Salient dependent variables (DVs)	Antecedents to DVs (sig. indicated by *): n-1:[15] Immediate; n-2: Secondary (if any); n-3: Tertiary (if any); n-4: Quaternary (if any)	Moderators considered (if any)	Theoretical framework(s) relied upon (if any)
(Choi et al. 2016)	Willingness to delegate (information disclosure to apps through Facebook)	n-1: Transactional privacy concerns * n-2: Privacy attributes (info collection and profile control) *	General privacy concerns	Communication privacy management theory
(Crossler et al. 2017)	Intentions to use identity ecosystem (disclosures)	n-1: Intrapersonal characteristics (censorship attitude, self-efficacy, behavioral-based inertia, previous similar experience), perceptions of the controlling agent (trustworthiness [b], privacy concerns), perceived system characteristics (granularity, efficacy, inconvenience) n-2: Reputation		Privacy calculus, social exchange theory [a] *(Note: Model tested across three domains, so significance levels not captured here)*
(Dinev et al. 2006)	Willingness to provide personal info	n-1: Internet privacy concerns *, Internet trust [b] *, personal Internet interest *, perceived Internet privacy risk * n-2: Perceived Internet privacy risk *		Privacy calculus [a]
(Dinev et al. 2008)	Willingness to provide personal information to transact on the Internet (PPIT)	n-1: Privacy concerns related to information abuse (PCIA) *, privacy concerns related to information finding (PCIF) *, perceived need for governmental surveillance (PNGS) *, government intrusion concerns (GIC) n-2: Linkages between PCIA and PNGS *, PCIF and GIC *, PCIA and GIC *		Privacy calculus, asymmetric information theory [a]

Reference	Salient dependent variables (DVs)	Antecedents to DVs (sig. indicated by *): n-1:[15] Immediate; n-2: Secondary (if any); n-3: Tertiary (if any); n-4: Quaternary (if any)	Moderators considered (if any)	Theoretical framework(s) relied upon (if any)
(Gerlach et al. 2015)	Willingness to disclose	n-1: Perceived privacy risk *, privacy policy permissiveness n-2: Privacy policy permissiveness (on perceived privacy risk) *		
(Hui et al. 2007)	Actual disclosure	n-1: Manipulations: no privacy assurance; assurance through privacy statement *; assurance through privacy statement and seal; variable money incentives *. Measured: propensity to trust *, prior experience with info misuse *, prior Internet shopping experience *, privacy concerns		Choice theory using utility function (akin to privacy calculus) [a]
(James et al. 2015)	Intentions to disclose (data items and parties who gain access)	n-1: - Interpersonal privacy identity (info and interaction management) * - Privacy calculus/benefits (socialization, self-expression, pleasing others) *	None	Privacy calculus [a]
(Karwatzki et al. 2017)	Willingness to share information	n-1: Disposition to value privacy (DTVP) *, level of personalization *, transparency features	DTVP, transparency features	Information boundary theory
(Kehr et al. 2015)	Intention to disclose	n-1: Perceived privacy *, general privacy concerns *, general institutional trust [b] n-2: Perceived risks of information disclosure *, perceived benefits of information disclosure * n-3: Information sensitivity *	Affect	Privacy calculus [a]

Reference	Salient dependent variables (DVs)	Antecedents to DVs (sig. indicated by *): n-1:[15] Immediate; n-2: Secondary (if any); n-3: Tertiary (if any); n-4: Quaternary (if any)	Moderators considered (if any)	Theoretical framework(s) relied upon (if any)
(Keith et al. 2015) – Study 2	Disclosure	n-1: Privacy concern, privacy settings * , age *, gender, ethnicity, perceived risk *, perceived benefit *, mobile computing self-efficacy (MCSE) * n-2: privacy concern *, MCSE *		Privacy calculus, trust theory [a]
(Kordzadeh et al. 2017)	Willingness to communicate personal health info	n-1: Expected positive personal outcomes *, expected positive community outcomes *, privacy concern *, affective commitment	Affective commitment	Privacy calculus, affective commitment [a]
(Spiekermann et al. 2010)	Self-disclosure (self-reported)	n-1: Perceived privacy risk *, convenience *, relationship building *, self-presentation, enjoyment * n-2: Trust in other OSN members, perceived control **,** trust in OSN provider [b] * n-3: Perceived control		Privacy calculus [a]
(Lancelot Miltgen et al. 2014)	Disclosure behavior and protection behavior	n-1: Privacy concerns including 4 foci (Control, protection and regulation, trust, and responsibility)		(Note: Qualitative study through focus groups)
(Li 2014)	Behavioral intention	n-1: Perceived benefits *, site-specific privacy concerns * n-2: Website reputation [b] *, disposition to privacy *, website familiarity * n-3: Privacy experience *, gender, age, education	Website reputation, website familiarity	Privacy as control, levels of privacy, developmental theory of privacy

Reference	Salient dependent variables (DVs)	Antecedents to DVs (sig. indicated by *): n-1:[15] Immediate; n-2: Secondary (if any); n-3: Tertiary (if any); n-4: Quaternary (if any)	Moderators considered (if any)	Theoretical framework(s) relied upon (if any)
(Li et al. 2017)	Disclosure (intention)	n-1: PC *, Liking *, Motive consistency *, Perceived Privacy control *, Covariates: gender, shopping experience *, past invasion of privacy *, task type n-2: Motive consistency, Perceived Privacy control	- Perceived privacy control	Multidimensional development theory, cognitive appraisals, emotions
(Li et al. 2012)	Willingness to provide info	n-1: Likelihood of using online personalization * n-2: Privacy Concerns * and Privacy Protection, Perceived Quality of Personalization *	- Industry Domain - Past Experience	Privacy calculus [a]
(Li et al. 2015)	Disclosure breadth and depth, sensitive vs. less sensitive disclosure	n-1: Gender and age *, account rating *, number of friends *, number of blogs *, blog length	- Gender and age	Communication privacy management theory
(Lowry et al. 2011)	Use of instant messaging (implied disclosure)	n-1: Behavioral intention to use IM * n-2: Attitude toward IM technology * n-3: Information privacy concerns *, desire for awareness * n-4: Masculinity, uncertainty avoidance, power distance, collectivism (Mixed *– see Fig 4 in article)		Social exchange theory *(Note: US and China samples)*

Reference	**Salient dependent variables (DVs)**	**Antecedents to DVs (sig. indicated by *): n-1:[15] Immediate; n-2: Secondary (if any); n-3: Tertiary (if any); n-4: Quaternary (if any)**	**Moderators considered (if any)**	**Theoretical framework(s) relied upon (if any)**
(Malhotra et al. 2004)	Behavioral intention to disclose information	n-1: Trusting beliefs [b] *, risk beliefs *, type of info requested * n-2: Internet privacy concerns * n-3: Type of info requested *		
(McKnight et al. 2002)	Intention to share information with web vendor	n-1: Perceived web risk, trusting intention (willingness to depend on web vendor), trusting beliefs in web vendor [b] n-2: Perceived vendor information, perceived site quality, structural assurance of the web *(Note: Due to model's complexity, significance levels not shown here)*		
(Ozdemir et al. 2017)	Information disclosure (declared)	n-1: Benefits *, Trust *, PC * n-2: Risks *, Privacy Experiences * and Privacy Awareness *	None	Antecedents – privacy concerns – outcomes framework
(Posey et al. 2010)	Disclosure (declared)	n-1: - Social influence * - Perceived benefit (reciprocity) * - Perceived trust [b] * - Risk beliefs * - Perceived collectivism * - Education and Age	Culture (Fr vs. UK)	Social exchange theory, social penetration theory, cross-cultural theory (individualism-collectivism)

Reference	Salient dependent variables (DVs)	Antecedents to DVs (sig. indicated by *): n-1:[15] Immediate; n-2: Secondary (if any); n-3: Tertiary (if any); n-4: Quaternary (if any)	Moderators considered (if any)	Theoretical framework(s) relied upon (if any)
(Schwaig et al. 2013)	Behavioral intention	n-1: Attitude *, PC * n-2: PC, Individual differences (alienation, self-esteem, computer anxiety) *, Attributes of an information practice (permission, transfer, interaction with IT) *		
(Shih et al. 2017)	Online self-disclosure (intention)	n-1: Switching cost *, Dependency *, Cognitive Trust [b] *, Affective Trust [b] n-2: Cognitive social identity *, Affective social identity (mixed *), Evaluative social identity *		Constraint-based (lock-in) and dedication-based (trust-building) mechanisms, social identity theory
(Son et al. 2008)	Info privacy-protective responses, including refusal to disclose, misrepresentation [c]	n-1: Information privacy concerns (mixed *), perceived justice, societal benefits		Justice theory
(Tow et al. 2010)	Sharing on Facebook	n-1: Context, value		*(Note: Qualitative study based on interviews and observation)*
(Wakefield 2013)	Intentions to disclose	n-1: Website trust [b] *, positive affect *, negative affect *, website privacy * n-2: Internet security (mixed *) *(Note: Also some relationships between n-1 variables (all *))*		Cognitive consistency theory, motivational model

Reference	Salient dependent variables (DVs)	Antecedents to DVs (sig. indicated by *): n-1:[15] Immediate; n-2: Secondary (if any); n-3: Tertiary (if any); n-4: Quaternary (if any)	Moderators considered (if any)	Theoretical framework(s) relied upon (if any)
(Xu et al. 2009)	Intention to disclose personal information	n-1: Privacy benefits *, Privacy risks * n-2: Compensation (mixed *), Industry self-regulation *, government regulation (mixed *) *(Note: ran model twice – for pull and push; some control variables were significant, but not included here)*		Privacy calculus, justice theory [a]
(Yu et al. 2015)	Self-disclosures on social networking websites	n-1: Expression, Self-presentation *, Social acceptance *, Reciprocity *, Social rejection, Privacy risk n-2: Affect toward self-disclosures *, Affect toward SN websites (mixed *) *(Note: For "indirect" model. Study also tested a "direct" model.)*		Direct causation theory, affect heuristic theory
(Zimmer et al. 2010a)	Actual disclosure	n-1: Disclosure intent *, benefits n-2: Trust [b] *, privacy *	Dyadic condition	Social response theory, principle of reciprocity
(Zimmer et al. 2010b)	Disclosure intention (email address, postal address, weight, and medical history)	n-1: Usefulness, Attitude * n-2: Trust [b] *, Risk *, Relevance *	Type of data	Theory of reasoned action, transaction cost economics

[a] Studies (also underlined in light gray) that use the privacy calculus (or a closely related variant) as their main theoretical framework

[b] Studies that include Trust (or a trust-related construct) in their model

[c] Studies that include Falsification (or a closely related variant) in their model

Appendix 2. Sample Screenshots corresponding to Data Request Forms (Original version)

Short Form/Homepage/One of the main mobile service providers in France (Orange)

Créez votre ESPACE PERSONNEL
Vous y RETROUVEREZ toutes les informations sur VOTRE COMPTE,
des informations EXCLUSIVES et de nombreuses offres PERSONNALISEES
Inscrivez-vous dès maintenant en COMPLETANT le formulaire ci-dessous

orange

Vos coordonnées

les champs suivis d'un * sont obligatoires, vous devez les compléter pour valider votre inscription

Nom*
Adresse*
Code postal*
Ville*
Email *

Indiquez nous vos souhaits

Je désire recevoir vos offres par email — Oui Non
Je désire recevoir les bons plans de vos partenaires — Oui Non

Long Form/Lottery/Another main mobile service provider in France (SFR)

Pour participer à notre grand JEU-CONCOURS
Et peut être gagner un des nombreux LOTS en jeu
Dont un SAFARI au Kenya
COMPLETEZ le formulaire ci-dessous

SFR

Vos coordonnées

les champs suivis d'un* sont obligatoires, vous devez les compléter pour valider votre inscription

Civilité* — Mme Mlle M
Nom*
Prénom*
Adresse*
Code postal*
Ville*
Email*
Téléphone*

Votre profil

Date de naissance — / /
Nationalité*
Niveau d'études*
Profession*
Situation familiale*
Nombre d'enfants (si vous n'en avez pas, mettez 0)

Vos habitudes et préférences

Hobbies/ centres d'intérêt*
Combien de temps passez vous au téléphone par mois ?*
A quand remonte l'achat de votre téléphone actuel ?*
Utilisez-vous des SMS ?* — Oui Non
Pour proposer cette offre à un de vos proches, merci d'indiquer son email*
Revenus annuels en 2004 (si vous n'en avez pas, mettez 0)

Indiquez nous vos souhaits

Je désire recevoir vos offres par email — Oui Non
Je désire recevoir les bons plans de vos partenaires — Oui Non

Appendix 3. Data Request Forms translated in English

Short Form/Homepage

Logo of the participant's mobile phone supplier (e.g., Orange)

To get a personal space where you would find all details about your account along with exclusive information and personalized offers, Create your profile by filling in the form just below:

(* required information)

Name*: _____ Address*: _____
City*: _____ Zip Code*: _____
Email Address*: _____@_____

I wish to receive offers by email ◉ YES ○ NO

I wish to receive offers from partner companies ◉ YES ○ NO

Long Form/Lottery

Logo of the participant's mobile phone supplier (e.g., SFR)

To participate in our Lottery and be able to win several prizes including a Kenya Safari, Fill in the form just below:

(* required information)

Title *: _____
Last Name *: _____ First Name *: _____
Address *: _____
City*: _____ Zip Code*: _____
Email Address *: _____@_____
Phone Number *: _____
Country *: _____
Birth Date: _____
Education *: _____
Profession *: _____
Marital Status *: _____ Number of children: _____
Hobbies *: _____
Time spent phoning *: _____
Cell phone date of purchase *: _____
Use of SMS and MMS *: _____
To sponsor a friend, indicate his/her email address here*: _____@_____
Yearly income: _____

I wish to receive offers by email ◉ YES ○ NO

I wish to receive offers from partner companies ◉ YES ○ NO

Appendix 4. Sample characteristics

Variables	Values	%
Demographics		
Gender	M	51%
	F	49%
Age (years)	18-24	30%
	25-34	30%
	35-44	17%
	45-54	14%
	55+	9%
Profession	White collars	18%
	Blue collars	42%
	Inactive (incl. students)	40%
Education level	Less than high school	24%
	Graduate	50%
	Postgraduate	15%
	PhD	11%
Experiences		
Internet experience	Beginner	6%
	Familiar	64%
	Expert	30%
E-mail usage	More than once a day	69%
	Once a day	22%
	More than once a week	9%
Web usage	More than once a day	73%
	Once a day	14%
	More than once a week	11%
	Even less	2%
Online purchase experience (number of online purchases/year)	No	6%
	Less than 5	29%
	5 to 20	35%
	More than 20	30%

Appendix 5. Survey items and statistics

Measures			
For each of the following statements state if you tend to agree or not	*Scale*	*Mean*	*SD*
(REL1) I don't see why the company is asking me some of these details (*inv.)*	1 – 7	4.19	1.973
(TRU1) I trust the company asking these information	1 – 7	4.35	1.898
Do you consider filling in this form as …	*Scale*	*Mean*	*SD*
(RIS1) Secure (7) to Insecure (1) (*inv.*)	1 – 7	4.98	1.735
(RIS2) Unsafe (7) to Safe (1)	1 – 7	4.23	1.799
(RIS3) Risky (7) to Not risky (1)	1 – 7	3.89	1.613
(BEN1) Beneficial (7) to Not beneficial (1)	1 – 7	4.04	1.732
(BEN2) Useful (7) to Not useful (1)	1 – 7	4.39	1.672
(BEN3) Valuable (7) to Not valuable (1)	1 – 7	4.13	1.589
(BEN4) Appealing (7) to Unappealing (1)	1 - 7	3.64	1.754
For each of the following statements state if you tend to agree or not	*Scale*	*Mean*	*SD*
(WIT1) I am willing to answer those questions (*inv.*)	1 - 7	4.69	1.910
(WIT2) I see no issue in providing those details (*inv.*)	1 - 7	4.49	1.901
(WIT3) I would have filled in and validated this form (*inv.*)	1 - 5	3.54	1.303
(FAL1) I would have given false information to some of the items in the form	1 - 5	1.94	1.134

Appendix 6. Results of the Exploratory Factor Analysis (EFA) from the pilot test

	Items	Communalities	Loadings		Cronbach's alpha
			BEN	RISK	
BEN1	Beneficial (7) to Not beneficial (1)	0.734	0.897		0.80
BEN2	Useful (7) to Not useful (1)	0.731	0.887		
BEN3	Valuable (7) to Not valuable (1)	0.680	0.783		
BEN4	Appealing (7) to Unappealing (1)	0.678	0.767		
RIS1_i	Secure (7) to Insecure (1) (inv.)	0.662		0.727	0.76
RIS2	Unsafe (7) to Safe (1)	0.767		0.796	
RIS3	Risky (7) to Not risky (1)	0.737		0.776	
Percentage of Variance			51%	19%	

Appendix 7. Test for Common Method Bias (CMB)

Construct	Indicator	Item Factor Loadings	Variance explained by the Factors	Method Factor Loadings	Variance explained by the Methods
Perceived Relevance	REL1	1.000	1.000	-0.273	0.074
Perceived Benefits	BEN1	0.866	0.750	0.282	0.079
	BEN2	0.834	0.695	0.320	0.102
	BEN3	0.817	0.667	0.360	0.130
	BEN4	0.798	0.636	0.298	0.089
Perceived Risks	RIS1	0.761	0.579	-0.315	0.099
	RIS2	0.862	0.744	-0.334	0.112
	RIS3	0.829	0.687	-0.287	0.083
Trust	TRU1	1.000	1.000	0.273	0.075
Withholding	WIT1	0.928	0.862	0.319	0.102
	WIT2	0.914	0.836	0.298	0.089
	WIT3	0.842	0.708	0.339	0.115
Falsification	LY1	1.000	1.000	-0.296	0.088
Average		0.921	0.852	0.165	0.098

Appendix 8. Discriminant validity: Correlations and Squared roots of AVEs (Fornell et al. 1981)

	REL	TRU	RIS	BEN	WIT	FAL
REL	1.000					
TRU	0.438	0.828				
RIS	-0.519	-0.448	0.818			
BEN	0.432	0.428	-0.409	1.000		
WIT	-0.530	-0.558	0.568	-0.612	0.896	
FAL	-0.329	-0.259	0.294	-0.332	0.468	1.000

Appendix 9. Discriminant validity: THE Heterotrait-Monotrait (HTMT) ratio (Henseler et al. 2015)

	REL	TRU	RIS	BEN	WIT
TRU	0.432				
RIS	0.596	0.464			
BEN	0.472	0.464	0.545		
WIT	0.561	0.649	0.698	0.642	
FAL	0.329	0.332	0.342	0.273	0.501

LEGEND:
REL: Perceived Relevance of the data request
BEN: Disclosure Benefits
RIS: Perceived Risks of data disclosure
TRU: Trust in the Company requesting the data
WIT: Withholding Intention
FAL: Falsification Intention

Appendix 10. Multi-Group Analysis

The first step in comparing two groups is to establish measurement invariance (Henseler et al. 2015). We followed the three-step Measurement Invariance of Composite Models (MICOM) procedure suggested by Henseler et al. (2015) for each of the contextual cues separately (see Appendices APP10-1, APP10-2). By using identical indicators per model and treating the data in the same manner, we established the first necessary step, configural invariance, in the MICOM procedure, both for the size of the request (amount of data) and the inferred incentive. We also established the second step, compositional invariance, for both contextual cues, by observing that the composite score correlation mean is larger than the 5% quantile of the empirical distribution of correlation between scores (Henseler et al. 2015).

Continuing with the third step of the MICOM procedure, we checked the variances (MICOM step 3.a) and mean values (MICOM step 3.b) between the construct scores. We first checked, for each contextual cue, if the observations that belong to the first group (low quantity and lottery) have the same variance in their latent variable score as the observations of the second group (high quantity and homepage) (MICOM step 3.a). We confirmed that variances are equal in both cases, therefore establishing at least partial measurement invariance. The last step (MICOM step 3.b) requires testing if means are equal in both groups for each contextual cue. This step is verified for the inferred incentive (lottery versus homepage) but not for the size of the request. We have thus established full measurement invariance for the inferred incentive and partial measurement invariance for the size of the request, which allows us to conduct a MGA and compare standardized path coefficients across groups for both contextual cues. In particular, the group with low data quantity (n = 42) was compared to the group with high data quantity (n = 42), and separately, the group with the lottery context (n = 42) was compared to the group with the homepage context (n = 42) using the MGA function of SmartPLS (see Table 5).

Table APP10-1. MICOM Results for Data Quantity Groups (Low vs. High)

Composite	C value (=1)	95% Confidence Interval	Compositional Invariance?
Perceived Relevance	1.000	[1.000; 1.000]	Yes
Perceived Benefits	0.999	[0.994; 1.000]	Yes
Perceived Risks	0.988	[0.987; 1.000]	Yes
Trust	1.000	[1.000; 1.000]	Yes
Withholding	1.000	[0.998; 1.000]	Yes
Falsification	1.000	[1.000; 1.000]	Yes
Composite	**Logarithm of the composite's variances ratio (=0)**	**95% Confidence Interval**	**Equal variances?**
Perceived Relevance	-0.122	[-0.289; 0.260]	Yes
Perceived Benefits	-0.246	[-0.395; 0.384]	Yes
Perceived Risks	-0.129	[-0.440; 0.405]	Yes
Trust	0.069	[-0.359; 0.332]	Yes
Withholding	-0.119	[-0.337; 0.359]	Yes
Falsification	-0.210	[-0.595; 0.548]	Yes
Composite	**Difference of the composite's mean value (=0)**	**95% Confidence Interval**	**Equal mean values?**
Perceived Relevance	-0.626	[-0.301; 0.290]	No
Perceived Benefits	0.213	[-0.294; 0.290]	Yes
Perceived Risks	-0.264	[-0.308; 0.277]	Yes
Trust	0.310	[-0.323; 0.284]	No
Withholding	0.291	[-0.294; 0.274]	No
Falsification	-0.268	[-0.312; 0.312]	Yes

Table APP10-2. MICOM Results for Inferred Incentive Groups (Lottery vs. Homepage)

Composite	C value (=1)	95% Confidence Interval	Compositional Invariance?
Perceived Relevance	1.000	[1.000; 1.000]	Yes
Perceived Benefits	0.998	[0.994; 1.000]	Yes
Perceived Risks	0.999	[0.986; 1.000]	Yes
Trust	1.000	[1.000; 1.000]	Yes
Withholding	1.000	[0.998; 1.000]	Yes
Falsification	1.000	[1.000; 1.000]	Yes
Composite	**Logarithm of the composite's variances ratio (=0)**	**95% Confidence Interval**	**Equal variances?**
Perceived Relevance	0.172	[-0.290; 0.255]	Yes
Perceived Benefits	0.005	[-0.429; 0.408]	Yes
Perceived Risks	-0.217	[-0.415; 0.395]	Yes
Trust	-0.099	[-0.337; 0.344]	Yes
Withholding	-0.010	[-0.400; 0.373]	Yes
Falsification	-0.235	[-0.534; 0.548]	Yes
Composite	**Difference of the composite's mean value (=0)**	**95% Confidence Interval**	**Equal mean values?**
Perceived Relevance	-0.127	[-0.301; 0.313]	Yes
Perceived Benefits	0.124	[-0.306; 0.275]	Yes
Perceived Risks	-0.169	[-0.302; 0.313]	Yes
Trust	0.065	[-0.310; 0.284]	Yes
Withholding	0.115	[-0.322; 0.296]	Yes
Falsification	-0.111	[-0.268; 0.290]	Yes