Strangers on a Plane: Context-Dependent Willingness to Divulge Sensitive Information

LESLIE K. JOHN
ALESSANDRO ACQUISTI
GEORGE LOEWENSTEIN

New marketing paradigms that exploit the capabilities for data collection, aggregation, and dissemination introduced by the Internet provide benefits to consumers but also pose real or perceived privacy hazards. In four experiments, we seek to understand consumer decisions to reveal or withhold information and the relationship between such decisions and objective hazards posed by information revelation. Our central thesis, and a central finding of all four experiments, is that disclosure of private information is responsive to environmental cues that bear little connection, or are even inversely related, to objective hazards. We address underlying processes and rule out alternative explanations by eliciting subjective judgments of the sensitivity of inquiries (experiment 3) and by showing that the effect of cues diminishes if privacy concern is activated at the outset of the experiment (experiment 4). This research highlights consumer vulnerabilities in navigating increasingly complex privacy issues introduced by new information technologies.

The Internet, according to a *New York Magazine* article, has made privacy “an artifact—quaint and naïve.” “We are all,” as the author expressed it, “eternally onstage” (Nussbaum 2007). Public records that were previously difficult to access, such as home tax information, are now readily available and searchable. Many online retailers record customers’ click streams and have developed sophisticated techniques to make educated guesses about the customers’ characteristics. Moreover, beyond these data, which are collected covertly and without explicit consent, many people voluntarily post vast amounts of information on the Web, on tweets, on social network pages, and even on personal blogs. Yet, people also report being deeply concerned about their privacy and are uncomfortable with firms knowing intimate details of their lives (Hoffman and Novak 1997; Taylor 2003), often with good reason (Odlyzko 2003). Periodic explosions of concern about privacy, for example, in response to news stories about identity theft and highly publicized changes in social network sites’ privacy policies, show that public concern about privacy is latent, if not always activated.

How can we make sense of the contradictory attitudes that individuals display toward privacy—from the seemingly reckless willingness of some to post personal and even incriminating information on social networks to the deep concern people express over the range of information being collected about them and the way it is used (Westin 1991)? In this article we argue, and present experimental evidence to support, that moment-to-moment concerns about privacy are responsive to contextual cues that often bear little con-

---

Leslie K. John is a doctoral candidate in behavioral decision theory, 208 Porter Hall, Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA 15213 (lkjohn@andrew.cmu.edu). Alessandro Acquisti is an associate professor of information technology and public policy, Heinz College, Hamburg Hall, Carnegie Mellon University, Pittsburgh, PA 15213 (acquisti@andrew.cmu.edu). George Loewenstein is Herbert A. Simon Professor of Economics and Psychology, 208 Porter Hall, Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA 15213 (gl20@andrew.cmu.edu). Address correspondence to Leslie John. This article is based in part on the first author’s doctoral dissertation, which was supported by a fellowship from the Social Sciences and Humanities Research Council of Canada. The authors would like to thank Kinshuk Jerath, Howard Seltman, John Tierney, Joachim Vosgerau, and Roberto Weber for invaluable help with this project. The authors are especially grateful to the editor, the area editor, and the reviewers for their guidance and constructive comments.

John Deighton served as editor and Gal Zauberman served as associate editor for this article.

Electronically published August 31, 2010
nection to the objective dangers and benefits of divulging information.

In four experiments, we show that specific configurations of contextual cues can give rise to different levels of disclosure across situations characterized by equivalent disclosure dangers (and benefits; experiment 1) and that cues that are objectively associated with greater disclosure danger can, if they suppress privacy concerns, have the perverse effect of increasing disclosure (experiments 2–4). Beyond documenting these effects, two of the studies provide explicit support for this cue-based account of privacy concerns, described in greater detail in the next subsection of the article, and rule out alternative explanations. Experiment 3 shows that cues affect not only divulgence of information but judgments of the sensitivity of inquiries. Experiment 4 shows that when privacy concerns are evoked at the outset of the experiment, the impact of cues on divulgence is eliminated. These findings provide support for the idea that cues affect divulgence by rousing, or downplaying, privacy concerns.

CONCEPTUAL BACKGROUND

A considerable body of academic research on privacy is premised on the assumption of rational choice. This work has been characterized by the assumption that (1) people make sensible and consistent trade-offs between privacy and other concerns (Derlega et al. 1993; Posner 1981; Rosenfeld 2000; Stigler 1980) and (2) there are reliable differences between individuals in concern for privacy (Laudon 1996; Westin 1991). For example, it has been argued that disclosure decisions are made by balancing “the usefulness of privacy with the utility of openness” (Petronio 2000, 37) and that people engage in “disclosure management,” such that they disclose information only when they expect a “net benefit” (White 2004, 48).

Even research that has not explicitly adopted a rational choice perspective has often done so implicitly. For example, several researchers have attempted to measure the monetary value that people place on privacy (Danezis, Lewis, and Anderson 2005; Hann et al. 2007). In a conjoint analysis, Hann and colleagues (2002) analyzed the trade-offs people made between privacy dangers and disclosure benefits and concluded that “among U.S. subjects, protection against errors, improper access, and secondary use of personal information is worth $30.49–$44.62” (14).

Marketers and other social scientists (Jourard and Lasakow 1958; Milberg et al. 1995; Westin 1991) have also constructed individual difference measures of concern for privacy, an endeavor premised on the assumption that there are stable individual differences to be measured. Some researchers have recommended that marketers segment consumers into privacy types and use this classification to tailor their services and products to consumers in distinct segments (Hann et al. 2007; Milberg et al. 1995). Westin, for example, categorizes individuals into three privacy types: fundamentalists, pragmatists, and unconcerned (1991).

Challenging the assumptions of both perfect rationality and stable preferences, the field of behavioral decision theory has documented that preferences are often influenced by factors that are difficult to justify on a normative basis, for example, by elicitation method (Tversky, Slovic, and Kahneman 1990) and by the framing of alternatives (Dhar and Simonson 1992; Tversky and Kahneman 1974). These effects tend to be especially pronounced when people are uncertain about their own values (Fox and Tversky 1995; Griffin, Liu, and Khan 2005; Hsee et al. 1999, 2003), which is often the case for privacy (Acquisti 2004). If the material value of privacy is already extremely difficult to estimate, the psychological value is likely to be even less well defined.

The privacy literature provides strong hints that this is the case. For example, a closer look at scales designed to measure individual differences in privacy concerns (mean apprehension over the security of one’s personal information; Altman 1975; Culnan and Armstrong 1999; Margulis 2003; Smith et al. 1996; Stone et al. 1983) suggests that such differences may not be so stable. Although many of these scales have impressive psychometric properties (Smith, Milberg, and Burke 1996), the few studies that have assessed predictive validity have found that such scales, at best, only weakly predict actual disclosure (Lubin and Harrison 1964; Marshall 1974). Drawing both on the characteristics of privacy concern and on these specific findings in the privacy literature, we propose:

**Proposition 1.** Privacy is a domain in which preference uncertainty is pronounced.

Beyond the idea that privacy concern is subject to preference uncertainty, we further posit that people are likely to seek to resolve such uncertainty by relying on contextual cues. This intuition is based on research showing that when individuals are uncertain of their preferences, their decisions can be influenced, often powerfully, by contextual cues (Simonson and Tversky 1992). However, the contextual cues guiding decision making often provide a misleading indication of the prevailing costs and benefits, which can cause people to behave in counterproductive ways. For example, cues that signal decreases in objective dangers of disclosure can lead people to be less forthcoming with information: individuals given assurances of confidentiality are less willing to provide a questionnaire than those receiving no assurance (Frey 1986; Singer, Hippler, and Schwarz 1992).

Drawing again both on general findings in the decision-making literature and on specific findings relating to privacy, we propose, additionally:

**Proposition 2.** Momentary privacy concern can be driven by cues that rouse or downplay privacy concern, thereby affecting individuals’ willingness to disclose.

These propositions, in turn, lead to the main predictions tested in the four experiments:

**H1:** Situations that differentially activate privacy concern will lead to different levels of disclosure, even if they are equivalent with respect to the objective costs and benefits of disclosure (experiment 1).
H2: Contextual cues that lower privacy concern but signal higher objective disclosure dangers will increase disclosure (experiments 2–4).

By “disclosure danger,” we mean the potential for divulgence to result in negative outcomes (examples include receiving spam e-mails as a result of divulging one’s e-mail address and having one’s identity stolen as a result of divulging one’s social security number).

The notion that the cues that affect privacy concern do not necessarily coincide with the objective dangers of disclosure in a given situation raises two important questions: the first is what the impact of a particular cue will be on disclosure, and the second is whether the cue moves a person closer or further from the ideal level of disclosure. The four experiments in this article provide answers to the first question; in the discussion section, we discuss implications of our findings for the second question.

An important feature of our conceptual framework is that these cues—which are hypothesized to affect privacy concern—should only affect the divulgence of privacy-relevant information, which we operationalize by the sensitivity of the information that participants are asked to divulge. Information on one’s food preferences, for example, is inherently less sensitive, and hence less privacy relevant, than information on one’s sexual preferences. It is worth noting that the sensitivity of divulged information is one determinant of disclosure danger (others include the way divulged information is transmitted or stored). To test this implication of the framework—that sensitive inquiries should be particularly responsive to our manipulations—we test the impact of cues on the divulgence of both innocuous and sensitive (i.e., privacy-relevant) information. We predict that

H3: Cues that affect privacy concern will have a greater effect on willingness to divulge sensitive rather than benign information (experiments 1 and 3).

OVERVIEW

In four experiments, participants indicated whether they had engaged in a series of sensitive, and in some cases illegal, behaviors. Between subjects, we manipulated a factor designed to affect privacy concern. Consent forms were not used, out of concern that they might contaminate our manipulations by universally cueing privacy concerns.

The primary dependent measure was the proportion of questions answered affirmatively (i.e., affirmative admission rate; AAR). People who, because of privacy concerns, chose not to provide information on the item might either fail to answer the question or respond negatively (deny that they had engaged in the behavior). It is, however, unlikely that they would admit to having engaged in the behavior. Admissions rates thus reflect the complement of the sum of people who (1) did not engage in the behavior (which can be assumed to be, on average, equal across conditions), (2) reported that they did not, although they did, and (3) failed to answer the questions. This data analysis approach is conservative because, to detect an effect of the manipulation, the impact of cues has to rise above the noise (error variance) produced by differences in true rates of engaging in the behavior across conditions.

To make the admissions more relevant to marketers, we asked participants to provide identifying information in the form of e-mail addresses (at the beginning of the survey in experiment 1 and at the end in experiments 2–4). To encourage people to provide e-mail addresses and motivate truthful admissions, all participants could request to “receive personalized results, including where [they] fall relative to others on the traits and attitudes the survey measures.” They were also told (at the outset of the survey) that they would be able to access the survey’s moment-to-moment aggregate results. Demographic questions appeared on the same page on which participants were asked to supply their e-mail addresses.

EXPERIMENT 1

Experiment 1 was a $2 \times 2$ mixed design. Between subjects, we manipulated the method of inquiry: participants were asked either directly or indirectly whether they had engaged in each of 16 behaviors. We predicted that AARs would be lower in the direct-inquiry condition, which we hypothesized would rouse privacy concerns relative to the indirect-inquiry condition. Within subjects, we varied the intrusiveness of the questions (seven tame and nine intrusive).

Although the inquiry conditions were equivalent with respect to the dangers and benefits of disclosure (in the sense that both inquiry methods solicited the same information—whether the respondent had engaged in the behavior—from Web sites that were comparable except for the subtle difference in inquiry), the indirect-inquiry condition was designed to make admissions seem secondary—almost an afterthought—which we predicted would increase self-revelation by keeping privacy concern latent (hypothesis 1). However, this hypothesis is restricted to the intrusive questions—the tame questions are arguably not particularly privacy relevant, and therefore they should not be affected by a manipulation designed to affect privacy concern (hypothesis 3).

Method

Participants. Eight hundred and ninety New York Times Web site visitors participated ($M_{age} = 41$ years, SD = 14.3; 59.3% male; 81.3% Caucasian, 7.8% Asian, 3.4% Hispanic, 1.8% African American, 5.6% other ethnicities). Five percent of participants were excluded from the sample because they did not finish the survey (NS differences in completion rate between conditions).

Procedure. A link to the survey titled “Test Your Ethics” was posted on TierneyLab—a New York Times science columnist’s official blog. The ethics cover story served to distract participants from the survey’s actual focus on privacy.
TABLE 1
EXPERIMENT 1: AFFIRMATIVE ADMISSION RATES TO THE INTRUSIVE ITEMS, BY QUESTION AND CONDITION
(LISTED IN ORDER OF PRESENTATION)

<table>
<thead>
<tr>
<th>Item</th>
<th>Affirmative admission rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Letting a friend drive after thinking he or she had had too much to drink</td>
<td>44.0 50.0</td>
</tr>
<tr>
<td>4. Neglecting to tell a partner about a sexually transmitted disease from which one is currently suffering</td>
<td>3.5 5.5</td>
</tr>
<tr>
<td>5. Lying about one’s income or that of one’s family to someone**</td>
<td>28.9 41.3</td>
</tr>
<tr>
<td>7. Having sex with the current husband, wife, or partner of a friend</td>
<td>15.5 11.4</td>
</tr>
<tr>
<td>8. Cheating on one’s tax return**</td>
<td>12.9 21.7</td>
</tr>
<tr>
<td>10. Having a fantasy of doing something terrible (e.g., torturing) to someone</td>
<td>49.4 55.5</td>
</tr>
<tr>
<td>12. Viewing pornography when unsure whether the subjects are underage*</td>
<td>11.1 16.4</td>
</tr>
<tr>
<td>13. Knowing about or witnessing a serious crime and failing to report it or stop it*</td>
<td>4.7 8.3</td>
</tr>
<tr>
<td>15. Making a false insurance claim**</td>
<td>2.6 6.2</td>
</tr>
</tbody>
</table>

*Chi-square test significant at p < .05 (two sided).
**Chi-square test significant at p ≤ .01 (two sided).

Upon clicking the link, participants were randomly assigned to one of two inquiry conditions.

Participants were told that they would be asked to rate the ethicality of a series of different behaviors and that “because people’s judgments of ethicality are influenced by whether or not they have personally engaged in the behavior, we will also be asking you whether or not you have done the behaviors.”

Participants were presented with 16 pairs of questions; in each pair, they were presented with a behavior and rated its ethicality on a scale labeled Not at all unethical/Somewhat unethical/Quite unethical/Extremely unethical/Nothing to do with ethics. Participants also indicated whether they had engaged in the behavior, although the way they did this varied between subjects.

Inquiry Method Manipulation. In the direct-inquiry condition, participants were asked, point-blank, “Have you done this behavior?” and responded on a response scale labeled Yes/No. In the indirect-inquiry condition, admissions were tightly coupled with ethicality ratings. Participants were first presented with the behavior; below it, the ethicality rating scale appeared twice. First, it was preceded by the question, “If you have EVER done this behavior, how unethical do you think it was?” The second version of the scale was labeled “If you have NEVER done this behavior, how unethical do you think it would be, if you were to do it?” It was only possible to answer using one of the two rating scales (fig. A1).

Pretest. Within subjects, we varied the intrusiveness of the behaviors, as determined by a separate sample of New York Times Web site visitors who rated the intrusiveness of the behaviors on a 4-point scale (Not at all intrusive/Mildly intrusive/Intrusive/Very intrusive). On the basis of these ratings, we chose a subset of items—those judged to be relatively tame (e.g., “Littering in a public place”) and those judged to be relatively intrusive (e.g., “Cheating on one's tax return”). In the main study, the items were presented in a pseudorandom order of intrusiveness (i.e., we made sure that there were never more than two questions of the same intrusiveness level presented in a row), which was the same for all participants.

Results

E-mail Addresses. Most participants (88.0%) gave an e-mail address (NS differences between conditions).

Ethicality Ratings. Ethicality ratings were negatively correlated with affirmative admissions (Spearman r ranged from −.17 to −.52, depending on the item; all p < .0005).

Affirmative Admission Rates. A 2 × 2 mixed ANOVA revealed a significant interaction between question intrusiveness and inquiry method (F(1, 843) = 28.24, p < .0005). In support of hypotheses 1 and 3, follow-up simple effect testing revealed that for the intrusive items, AARs were significantly higher in the indirect-inquiry condition relative to point-blank (M_{direct} = 0.24, M_{indirect} = 0.19; t(843) = −4.19, p < .0005). Specifically, indirect-inquiry participants were on average 1.48 times more likely to admit to having engaged in the intrusive behaviors than were those in the point-blank condition (table 1; fig. 1). The AARs to the tame items were not statistically different between conditions (t(843) = 1.93, NS), which is also consistent with hypothesis 3.

Experiment 1: Discussion

Individuals were more likely to admit to having engaged in sensitive behaviors and, hence, were more willing to divulge private information, when asked indirectly. Experiment 1 provides preliminary evidence that situations with objectively equivalent dangers and benefits of disclosure can elicit different degrees of disclosure. Moreover, the interaction between inquiry method and intrusiveness—the fact that only the privacy-relevant items were affected by the inquiry manipulation—is consistent with the notion that the different forms of inquiry lead to differences in privacy concern, a symptom of which was differences in disclosure.
EXPERIMENTS 2–4

In experiments 2–4, we test hypothesis 2 by introducing a contextual cue that downplays privacy concern, leading to increased disclosure (experiments 2–4) and lowered judgments of the sensitivity of inquiries (experiment 3), even though it is indicative of elevated disclosure danger. Specifically, we vary the look and feel of the Web pages on which participants are asked to disclose sensitive information. This hypothesis builds on previous research showing that the look of a Web site can influence consumers’ product preferences and purchase decisions (Mandel and Johnson 2002).

We compare disclosure as a function of whether the Web site looks professional versus unprofessional. Prior research has established a link between the look of a Web site and its security; professional Web sites (i.e., sites from reputable companies and institutions) are more likely than unprofessional ones to have P3P (a privacy preserving technology; Cranor 2002; Cranor et al. 2008; Grimm and Rossnagel 2000; Turner and Dasgupta 2003). Further research suggests that a Web site’s look is indicative of its reputability—for example, whereas professional Web sites use color and fonts sparingly and rarely misspell words, unprofessional Web sites are opposite on these characteristics (Ivory and Hearst 2002a, 2002b; Ivory, Sinha, and Hearst 2001). Therefore, unprofessional-looking Web sites are less likely to offer privacy protection relative to professional-looking ones, making them, if anything, higher in disclosure danger. Yet, as we will show, the characteristics that signal that a site is unprofessional can also suppress privacy concern and facilitate disclosure. This occurs even though observers (i.e., individuals asked only to rate the Web sites’ security) perceive the unprofessional site to have higher disclosure danger.

Experiment 2

Experiment 2 was a three-condition between-subjects design in which we manipulated the survey’s interface (professional vs. baseline vs. unprofessional).

Method

Participants. Two hundred students (which excludes the 1.5% of people who started but failed to complete the study) participated ($M_{age} = 21$ years, SD = 3.1; 58.6% male; 37.3% Asian, 37.3% Caucasian, 3.5% African American, 8.5% Indian, 13.5% other ethnicities).

Procedure. Laptops were placed on tables in buildings on the Carnegie Mellon University campus; students were asked to complete a brief “Web survey about student behaviors,” as they walked by. The first screen explained that the survey was about college students and that the experimenters were “interested in the types of behaviors that college students engage in.” Participants were also informed, on both this first screen of the survey and the sign advertising the experiment, that the survey was being conducted by researchers at Carnegie Mellon.

Participants were asked, using a Yes/No response scale, whether they had engaged in each of 15 intrusive behaviors (table 2). To directly link between-condition differences in disclosure to different levels of privacy concern, the survey
concluded with a 4-item privacy scale. Thus, whereas in experiment 1, the effect of the manipulation on privacy concern was demonstrated in an indirect way (i.e., through its impact on intrusive, but not tame, questions), in experiment 2, we demonstrate this sensitivity directly, through a self-report measure. Responding on a 5-point scale with endpoints “strongly disagree” and “strongly agree,” participants rated the extent to which, as they completed the survey, they were concerned about “incriminating myself,” “whether my answers would truly be private,” “who might have access to my answers,” and “whether the survey was truly anonymous.” This was the first time in the study that privacy-related words (e.g., “anonymous,” “private”) appeared.

On the next page, using the same response scale as for the privacy questions, participants indicated the extent to which they had been “tempted to say ‘No’ to behaviors in which I have actually never engaged” and “to say ‘Yes’ to behaviors in which I have actually engaged.” Each participant was given a candy bar after completing the survey.

Survey Interface Manipulation. Participants were randomly assigned to one of three conditions, differing only in interface and title (fig. A2). In designing the interfaces, we used elements identified in the human–computer interaction literature as indicative of unprofessional versus professional sites (Ivory and Hearst 2002a, 2002b; Ivory, Sinha, and Hearst 2001). The unprofessional version (intended to downplay privacy concerns) was titled “How BAD Are U???” in red font and included a cartoon devil logo. The professional version was titled “Carnegie Mellon University Executive Council Survey on Ethical Behaviors” in black font and displayed the university’s official crest. Finally, to make the case that the unprofessional condition downplays privacy concerns, rather than that the professional condition heightens such concerns, we also included a baseline condition that was designed to be as neutral as possible—it was called “Survey of Student Behaviors” in black font against a white background (fig. A2).

Pretest. To verify that perceived disclosure danger differed between the interfaces, 30 students were shown a screen shot of both the unprofessional and the professional Web sites and asked to rate how secure they perceived each to be, using a 5-item scale (α = .88; adapted from Salisbury et al. 2001). The items were “I would feel secure sending sensitive information over this website,” “This website is a secure means through which to send information,” “I would have concerns about giving out sensitive information on this website” (reverse scored), and “Overall, this website is a safe place to transmit sensitive information.” The items were answered on a 7-point response scale with endpoints labeled “strongly disagree” and “strongly agree.” The order in which participants rated the interfaces was counterbalanced between subjects.

Participants in the pretest rated the unprofessional site as significantly less secure than the professional one ($M_{unpro} = 7.5, M_{pro} = 20.9; F(1, 28) = 116.27, p < .0005$). There was no main effect of the order of presentation of the interfaces ($F(1, 28) = 2.40, NS$), nor was there an interaction between order and interface ($F(1, 28) = 1.32, NS$).

Results and Discussion

E-mail Addresses. At the end of the survey, 32% of participants gave their e-mail addresses. Participants in the baseline condition were less likely to give their e-mail addresses than those in the other conditions (professional = 38.5%, baseline = 18.8%, unprofessional = 38.0%; $\chi^2(2)$
= 7.60, p < .05), and surprisingly, the mean AARs were higher among participants who gave their e-mail (M\_mail = 0.29, M\_no\_mail = 0.19; t(198) = 3.90, p < .0005). This may reflect individual differences (whether momentary or enduring) between respondents; those who were less concerned about privacy may have divulged more and been more willing to provide e-mail addresses. Because of the between-condition difference in propensity to supply an e-mail address, we controlled for this variable in the analyses we report below.

**Affirmative Admission Rates.** We coded questions left blank as missing (i.e., neither admissions nor denials). The results do not change substantively (if anything, they are more supportive of our hypothesis) when the data are analyzed with blank responses coded as denials.

AARs were significantly different between conditions (M\_unprof = 0.28, M\_base = 0.21, M\_prof = 0.17; F(2, 196) = 6.82, p = .001). Follow-up pair-wise comparisons revealed the AARs to be significantly higher in the unprofessional condition relative to the baseline (F(1, 132) = 4.89, p < .05) and professional (F(1, 133) = 13.34, p < .0005) conditions. Participants in the unprofessional condition were on average 1.74 and 1.98 times more likely to admit to having engaged in the behaviors relative to the baseline and professional conditions, respectively (table 2). Disclosure rates were similar in the professional and baseline conditions (F(1, 126) = 1.10, NS), suggesting that the effect is driven by facilitation of disclosure in the unprofessional condition.

**Privacy Concern.** The privacy scale was reliable (a = .93) and revealed significant differences between conditions (M\_unprof = 1.8, M\_base = 2.3, M\_prof = 2.3; F(2, 196) = 6.20, p < .005). Follow-up pair-wise comparisons indicated that participants in the unprofessional condition were significantly less concerned about their privacy relative to those in the baseline (F(1, 132) = 11.79, p = .001) and professional (F(1, 133) = 9.20, p < .005) conditions (baseline vs. professional F(1, 126) = 0.55, NS).

**Truthfulness of Responses.** There were significant differences between conditions in the degree to which respondents were tempted to deny the truth (M\_unprof = 2.0, M\_base = 2.6, M\_prof = 2.5; F(2, 196) = 3.37, p < .05). Follow-up pair-wise comparisons revealed that participants in the unprofessional condition were significantly less likely to indicate that they had said no to behaviors in which they had actually engaged, relative to both the baseline (F(2, 132) = 6.51, p = .01) and the professional conditions (F(2, 133) = 3.92, p = .05). There were no significant differences in the self-reported propensity to say yes to behaviors in which they had not engaged (F(2, 196) = 0.35, NS).

The results of experiment 2 are consistent with the theory that privacy concerns can be either roused or, in this case, downplayed by contextual cues (e.g., the Web interface), thereby affecting disclosure. The unprofessional-looking Web site both suppressed privacy concern and facilitated disclosure. Participants in the unprofessional condition were less likely to deny involvement in behaviors in which they had actually engaged. The unprofessional interface did not seem to cause participants to admit to behaviors in which they had not engaged; rather, it appeared to facilitate admissions to behaviors in which they had engaged.

However, an alternative and more mundane explanation of experiment 2 is that the unprofessional condition caused participants to perceive the behaviors to be more socially desirable, making them more willing to respond affirmatively. Contrary to this interpretation, we argue that, as a symptom of suppressed privacy concern, the unprofessional interface caused participants to perceive the questions to be less intrusive relative to the professional condition. In experiment 3, we test these two competing accounts.

**Experiment 3**

In experiment 3, in addition to asking participants to divulge sensitive information, we measured participants’ perceptions of (a) the intrusiveness of the questions and (b) the social desirability of the behaviors. We predicted that the survey interface manipulation would only affect judgments of question intrusiveness.

Similar to experiment 1, we included items varying in intrusiveness to test hypothesis 3. We also conducted a pre-test to ensure that the social desirability of the tame versus intrusive questions was roughly equivalent. (A potential confound would have existed if the intrusive questions were also more socially desirable than the tame ones.) For example, the questions “Do you know who your state senators are?” and “Have you ever used sex toys?” were both judged to refer to socially desirable behaviors, yet the former is tame, whereas the latter is intrusive. Likewise, the questions “Have you ever eaten so much so as to feel sick afterward?” and “Have you ever watched someone while they undressed, without their knowledge?” were both judged to refer to socially undesirable behaviors, yet the former is tame, whereas the latter is intrusive.

Participants were presented with 12 questions. Between subjects, we manipulated the survey interface (unprofessional vs. professional); within subjects, we varied the intrusiveness of the questions (tame vs. intrusive). For each question, participants indicated whether they had engaged in the behavior, rated the question’s intrusiveness, and rated the social desirability of the behavior. We counterbalanced the order in which these three pieces of information were elicited. The experiment was therefore a 2 × 2 × 6 condition mixed design, with survey interface and order manipulated between subjects and question intrusiveness varied within subjects. There were three dependent measures: AARs, intrusiveness ratings, and social desirability ratings.

We hypothesized, first, that AARs would be higher in the unprofessional condition relative to the professional condition (hypothesis 2); however, as in experiment 1, this hypothesis is restricted to the intrusive questions (hypothesis 3). We also hypothesized that participants would judge the questions to be less intrusive in the unprofessional condition relative to the professional condition; by contrast, there
should be no difference between these conditions with respect to the social desirability ratings.

Method

Participants. One hundred and eighty students (which excludes the 1.6% of people who started but failed to complete the survey) participated for course credit ($M_{age} = 20$ years, SD = 1.2; 54.5% female; 29.7% Asian, 54.5% Caucasian, 3.4% African American, 6.9% Indian, 3.4% other ethnicities; NS differences between conditions). Twenty participants (10.9%) were excluded from the sample because they had taken the survey before (see procedure below). Excluded participants were no different from those included in the sample with respect to any of the dependent measures. The proportion of participants who were excluded was not significantly different between conditions.

Procedure. Participants signed up to participate in a “survey of behaviors” and were subsequently e-mailed the link to the survey. The e-mails were sent through an automated system that blinded us from knowing the participants’ names or e-mail addresses and that made the participants aware of this fact. The initial page of the survey, on which random assignment occurred, was the same as in experiment 2.

Participants were presented with the same 12 behaviors three times. One of the times, they indicated whether they had engaged in the behavior (Yes/No response scale); another time, they rated the intrusiveness of the behavior (“rate how intrusive (if at all) you would find it, to be asked the following questions”; response options: Not at all intrusive/Mildly intrusive/Intrusive/Very intrusive); another time, they rated its social desirability (“we would like to know how ‘cool’ (or not) you think it is to do each of the following behaviors”; response options: Not at all cool/Somewhat cool/Cool/Very cool). The 12 behaviors were presented in a pseudorandom order with respect to intrusiveness and social desirability, which was held constant across the three tasks (i.e., admissions, intrusiveness ratings, social desirability ratings).

Because recruitment took place at the same institution, but at a later date, as in experiment 2, a question also asked participants whether they had taken the survey before. Finally, the aggregate results collected to that point in time appeared on the last page of the survey.

Pretest. We conducted a pretest to choose questions that varied in intrusiveness and to ensure that the chosen tame versus intrusive questions did not differ with respect to social desirability. Forty-four pilot participants were presented with 60 questions describing various behaviors. They were asked to rate (a) the intrusiveness of each question and (b) how socially desirable it is to engage in the behavior described in each question. The participants in the pilot study were from the same participant pool as those who would participate in the actual study but were prevented from participating in the actual study. On the basis of these ratings, we chose 12 questions to be used in the actual study—six tame and six intrusive. The intrusive questions that we chose for the main study were judged in this pretest to be significantly more intrusive than the tame questions ($M_{intr} = 13.3, M_{tame} = 5.0; t(43) = 15.35, p < .0005$); there were no differences with respect to social desirability.

Results and Discussion

E-mail Addresses. At the end of the survey, 47% of participants gave an e-mail address (NS differences between conditions).

Affirmative Admission Rates. We coded questions left blank as missing (i.e., neither admissions nor denials). The results do not change substantively when the data are analyzed with blank responses coded as denials.

A 2 (interface) × 2 (intrusiveness) × 6 (order) mixed ANOVA revealed a significant two-way interaction between interface and question intrusiveness ($F(1, 147) = 4.13, p < .05$). Follow-up pairwise comparisons revealed that the AARs were significantly higher in the unprofessional condition for the intrusive questions ($M_{unpro} = 33, M_{pro} = .26$; $t(158) = 2.55, p = .01$). Participants in the unprofessional condition were 1.52 times more likely to admit to having engaged in the intrusive behaviors, relative to those in the professional condition (table 3; fig. 2). There was no difference between conditions for the tame questions ($t(157) = 0.97, NS$). There was also a main effect of intrusiveness; the AARs were lower for the intrusive items than for the tame items.

![Table 3](image-url)

<table>
<thead>
<tr>
<th>Item</th>
<th>Professional (%)</th>
<th>Unprofessional (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Have you ever watched someone while they undressed, without their knowledge?</td>
<td>9.2</td>
<td>17.9</td>
</tr>
<tr>
<td>4. Have you had more than five sexual partners?</td>
<td>18.4</td>
<td>16.7</td>
</tr>
<tr>
<td>6. Have you ever failed to pay back a loan?</td>
<td>10.5</td>
<td>19.0</td>
</tr>
<tr>
<td>8. Have you ever used sex toys?</td>
<td>17.1</td>
<td>21.4</td>
</tr>
<tr>
<td>10. Have you ever fantasized about having violent nonconsensual sex?**</td>
<td>14.5</td>
<td>31.0</td>
</tr>
<tr>
<td>12. Have you ever illegally downloaded music or software from the Internet?</td>
<td>85.5</td>
<td>91.7</td>
</tr>
</tbody>
</table>

**Chi-square test significant at $p \leq .01$ (two sided).
tame items ($M_{\text{tame}} = .64, M_{\text{intr}} = .29; F(1, 147) = 259.16, p < .0005$).

**Intrusiveness Ratings.** A 2 (interface) $\times$ 2 (intrusiveness) $\times$ 6 (order) mixed ANOVA revealed a significant main effect of interface ($F(1, 144) = 4.75, p < .05$). Specifically, the average question intrusiveness rating (on a 4-point scale) was lower in the unprofessional condition relative to the professional condition ($M_{\text{unpro}} = 2.1, M_{\text{pro}} = 2.2$). There was also a main effect of intrusiveness; the intrusive items were rated as more intrusive than the tame items ($M_{\text{intr}} = 2.8, M_{\text{tame}} = 1.5; F(1, 144) = 428.61, p < .0005$).

**Social Desirability Ratings.** A 2 (interface) $\times$ 2 (intrusiveness) $\times$ 6 (order) mixed ANOVA of the social desirability ratings revealed no significant differences. Therefore, the social desirability explanation of our findings was not supported, as there was no effect of survey interface on participants’ perceptions of the social desirability of the behaviors.

Experiment 3 shows that, relative to a professional-looking interface, an unprofessional-looking interface leads people to divulge more private information, even when this information is socially undesirable. The unprofessional interface also caused participants to judge the questions to be less intrusive. A competing interpretation of experiment 2—that the unprofessional interface increases disclosure because it leads people to perceive the behaviors to be socially desirable—was not supported.

**Experiment 4**

In experiment 4, we test whether the effect of the survey interface on disclosure disappears when privacy concern is roused at the outset of the experiment. Experiment 4 was a $2 \times 2$ between-subjects design in which we orthogonally manipulated (a) whether participants were cued to think of privacy from the outset (privacy vs. control) and (b) the interface of the subsequent survey (unprofessional vs. professional). To reduce complexity in the experimental design, only intrusive questions were used. We hypothesized an interaction: in the absence of privacy cueing, we expected a replication of experiments 2 and 3 such that participants in the unprofessional condition would disclose more than those in the professional condition. However, when privacy concern is roused from the outset of the study, we expected the difference between the unprofessional and professional conditions to disappear. We expected that heightening privacy concern from the outset of the study would buffer participants against the unprofessional interface’s tendency to lower privacy concern and elicit divulgence.

**Method**

The method and questions were the same as experiment 2, except for the removal of the baseline condition from the interface manipulation, the addition of the privacy cueing manipulation, and a few other minor changes, all of which are described below.

**Participants.** There were 769 participants ($M_{\text{age}} = 30$ years, $SD = 11.5$; $59.9\%$ female; $16.2\%$ Asian, $65.1\%$ Cau-
cian, 6.9% African American, 1.9% Indian, 3.9% other ethnicities; all NS between conditions), excluding the 1.2% of people who started but failed to complete the survey. Sixteen percent of participants were excluded from the sample because they had taken the survey before (assessed in the same way as in experiment 3). Although excluded participants were significantly younger than those who were included ($M_{ex} = 26.9$ vs. $M_{in} = 30.7$; $t(758) = 3.52$, $p < .0005$), they were the same with respect to all other dependent measures. The proportion of excluded participants was not significantly different between conditions.

Procedure. Participants were offered a candy bar in exchange for participating and completed the surveys on a laptop provided by the experimenter. Recruitment took place on one of two adjacent university campuses and differed slightly between campuses: potential participants were recruited as they walked by either tables set up in different buildings on the Carnegie Mellon University campus or a mobile lab (http://www.cbdr.cmu.edu/datatruck/) parked on a city street on the University of Pittsburgh campus. Within each location, participants were randomly assigned to one of the four conditions. To control for possible differences between the various locations, we included location as a covariate in our analyses. Participants were asked to complete two surveys; they were told that the first survey was a “photo identification task” (which served as the privacy cue manipulation, described below) and that the second was a “survey of behaviors” (which was the same as experiment 2).

Privacy Cue Manipulation. In the first survey, participants were asked to either “Phind the phishing e-mails” (privacy condition) or “Find the endangered fish” (control condition). First, the relevant terms were defined. In the privacy condition, participants were given the definition of phishing (an e-mail that attempts “to acquire sensitive information such as usernames, passwords and credit card details by masquerading as a trustworthy entity”; http://en.wikipedia.org/wiki/Phishing); in the control condition, participants were given the definition of an endangered species (http://en.wikipedia.org/wiki/Endangered). Participants were asked on the subsequent page to define either the term “phishing” (privacy condition) or “endangered species” (control condition).

On each of six subsequent pages, participants were presented with photos and asked to categorize them. In the privacy condition, participants were presented with screen shots of e-mail messages and indicated whether each constituted phishing or “just spam.” In the control condition, they were shown pictures of fish and indicated whether they thought the species was endangered. The answers were displayed at the end of the survey.

Results and Discussion

E-mail Addresses. At the end of the survey, 44% of participants gave their e-mail addresses (NS differences between conditions).

Affirmative Admission Rates. As predicted, there was a significant interaction between the two manipulations ($F(1, 632) = 3.91$, $p < .05$) and no main effects (fig. 3). Simple effect testing revealed that in the absence of privacy cueing (fishing condition), participants in the unprofessional condition admitted to having engaged in significantly more
behaviors, relative to those in the professional condition \( \bar{M}_{\text{unpro}} = 0.33, \bar{M}_{\text{pro}} = 0.29; F(1, 324) = 6.43, p < .05 \)—in other words, a replication of experiments 2 and 3. Relative to the professional condition, participants in the unprofessional condition were on average 1.17 times more likely to admit to having engaged in the behaviors. By contrast, when privacy concerns were cued from the outset of the experiment (phishing condition), there was no difference in AARs between unprofessional and professional conditions.

Discussion of Experiments 2–4

These three experiments show that people seem naturally more comfortable disclosing personal information on unprofessional sites—which are arguably more likely to misuse it. This occurs even though participants in a pilot study judged the unprofessional site to be higher in disclosure danger. Experiment 3 shows that participants in the unprofessional condition, whose privacy concern has been suppressed, perceive the questions to be less intrusive than do those in the professional condition. Experiment 4 replicates and extends these findings by showing that the effect of the survey interface on disclosure is eliminated when privacy concerns are evoked from the outset of the study.

GENERAL DISCUSSION

These experiments suggest that privacy concern and, in turn, willingness to divulge are influenced by contextual cues that are incommensurate with, or even negatively related to, the objective dangers of disclosure. This can lead to differences in disclosure between situations with the same objective disclosure dangers and benefits (experiment 1) and even to increased disclosure in situations indicative of greater disclosure danger (experiments 2–4). Suppressing privacy concern caused individuals to perceive questions to be relatively nonintrusive (experiment 3); activating privacy concern at the outset of experiment 4 buffered individuals against contextual cues that otherwise downplayed privacy concern when it was actually warranted.

Implications

Our results stand in contrast to the considerable body of privacy research that is premised on the assumption of rational choice and that has informed marketing recommendations on how to obtain personal information from consumers. For example, some have argued that consumers will be more likely to comply with a firm’s requests for personal information if the firm displays its privacy policy (Benassi 1999; Culnan and Armstrong 1999). Although this recommendation is intuitively sensible—privacy policies contain information relevant for making rational disclosure decisions—our research suggests that providing such information may backfire, by rousing privacy concern and therefore suppressing divulgence. Our results therefore imply that such recommendations could be improved by considering the effect of nonnormative contextual factors on privacy concern. Specifically, experiment 1 suggests that indirect attempts at obtaining sensitive information may be particularly fruitful. For better or for worse, numerous technologies already capitalize on the inadvertent disclosures people make through Internet searches; Google Flu Trends, for example, has been used to predict the spread of influenza.

Experiments 2–4 imply that individuals are prone to disclosing in contexts that downplay privacy concern—ironically, even when such contexts are likely higher in both objective and perceived disclosure danger. These experiments suggest that consumers will be especially forthcoming with information when sensitive questions are asked informally. Combined with research showing that assurances of anonymity and confidentiality can backfire (Singer et al. 1992), these results suggest, perhaps ironically, that marketers may be particularly successful in obtaining private information when they make the fewest promises to protect consumers’ privacy—enabling marketers to retain great flexibility in how they may use the disclosed information.

Limitations

We attempt to draw general conclusions about how individual concern for privacy, and in turn self-disclosure, responds to contextual cues in normative ways. Yet, much as general conclusions about other topics (e.g., curiosity) are often drawn from research focusing on a subset of domains, the present research is limited to one domain of disclosure: the revelation of sensitive, and potentially incriminating, personal facts. Although we suspect that our results apply to a wide range of other disclosure domains, we do not test this assumption. Future research could therefore test the effect of contextual cues on other types of disclosure behavior—for example, on people’s propensity to divulge other people’s personal information.

Second, whereas much of the privacy literature to date has focused on individual differences, we failed to take account of such differences but focused exclusively on situational factors that we posited to drive momentary changes in privacy concern. Our research is therefore silent on how contextual factors might interact with individual differences to affect self-disclosure. For example, people may differ in both the extent to which they are generally concerned about their privacy and their general desire to divulge, which could affect the impact of contextual cues on self-disclosure: a cue signaling that public divulgence is likely might increase disclosure among individuals high in the desire to divulge but inhibit it by individuals who are high in privacy concern. Understanding how individual differences interact with contextual factors to affect self-disclosure might help marketers better predict consumers’ responses to requests for personal information and allow for more targeted, and ultimately more successful, attempts at obtaining consumers’ information.

Third, we cannot validate the truthfulness of our respondents’ disclosures; it is therefore possible that our manipulations simply affected people’s propensity to lie. In experiments 2–4, for example, the unprofessional Web site...
may have made individuals more likely to say yes to behaviors in which they had not actually engaged. Note, however, that if this were the case, one would have expected participants in the unprofessional condition to have judged the behaviors to be relatively socially desirable; they did not (experiment 3). In addition, a self-report measure found no differences in the extent to which participants were “tempted to say ‘Yes’ to behaviors in which I have actually never engaged” (experiment 2). But regardless of the truthfulness of participants’ admissions, we think the results have provocative implications—the mere claim that one has committed a crime, for example, can have serious consequences, irrespective of its validity. Nonetheless, validating the truthfulness of admissions is an important topic for future research because it could help to devise techniques to promote divulgence. For example, if it is determined that indirect inquiry facilitates truthful admissions, it could be used at blood donation clinics to ask potential donors important but sensitive information about their blood (e.g., HIV status). Beyond making an important practical contribution, such a finding would also contribute to the literature on eliciting truthful responses to sensitive questions (Lamb and Stern 1978; Lensvelt-Mulders et al. 2005; Tourangeau and Yan 2007).

Open Questions

The present research enables us to make inferences about when individuals are and are not concerned about privacy; further research is needed to identify conditions that promote good disclosure decisions. For example, although experiment 4 shows that the unprofessional-looking survey’s ability to facilitate disclosure is eliminated when participants have been cued to think of privacy, it would be wrong to conclude that cueing people to think about privacy concerns will necessarily make them disclose and withhold information when it is in their best interest to do so.

Whether a cue affecting privacy concern leads to more self-interested disclosure decisions depends on the relationship between privacy concern and objective disclosure danger before and after the introduction of the cue. For example, consider a cue (such as that introduced in experiments 2–4) that downplays privacy concern despite increased disclosure danger. If the underlying tendency is to be underconcerned about one’s privacy, such a cue would be detrimental by causing individuals to overdisclose even more than they normally would. This situation would present a tension between consumers’ best interests and marketers’ motives to obtain information from them. However, if the tendency is to be overly concerned about one’s privacy, such a cue would be beneficial, by bringing people’s disclosure closer to ideal levels. It may seem paradoxical that a cue that makes a person feel more safe but be less safe should increase disclosure, but if an individual is prone to underdisclose in the first place, such a cue would be mutually beneficial to both consumers and marketers. By the same token, a cue that makes a person feel less safe but be more safe should decrease disclosure, but, if the tendency is to be underconcerned in the first place, such a cue would be beneficial.

This framework can account for a wide range of factors that affect self-disclosure in seemingly nonnormative ways. For example, the finding that assurances of confidentiality can decrease people’s willingness to respond to surveys on sensitive subjects (Singer et al. 1992) fits with this perspective. Assurances serve as cues that rouse privacy concern; however, because they promise confidentiality, they also lower the objective dangers of disclosure. In this case, a mutually beneficial way for marketers to obtain information from people may be to protect the confidentiality of consumers’ data but to not inform the consumers of this protection.

Concluding Comment

What about the strangers on a plane? We suspect that the phenomenon of opening up to a complete stranger on a plane is sufficiently ubiquitous that most readers of this article will have instantly understood the allusion. Can our results shed light on this well-known, if anecdotal, phenomenon? Perhaps a stranger on the plane is at the “sweet spot” when it comes to the absence of cues that trigger concern about privacy. First, the individual is a stranger whom one is unlikely to encounter again. Second, the setting is divorced from normal daily life, as if it is happening in a parallel world. Finally, at least for the many of us with a fear of flying, part of the explanation may lie in the unconscious belief that the stranger will take our secrets to the grave, when the plane, inevitably, crashes.
APPENDIX

FIGURE A1
EXPERIMENT 1: QUESTION LAYOUT IN INDIRECT-INQUIRY CONDITION

The Behavior:
Cheating on one’s tax return.

A) If you have EVER done this behavior, how unethical do you think it was?
   - Not at all unethical
   - Somewhat unethical
   - Quite unethical
   - Extremely unethical
   - It depends
   - Nothing to do with ethics.

B) If you have NEVER done this behavior, how unethical do you think it would be if you were to do it?
   - Not at all unethical
   - Somewhat unethical
   - Quite unethical
   - Extremely unethical
   - It depends
   - Nothing to do with ethics

NOTE.—Color version available as an online enhancement.
FIGURE A2

EXPERIMENT 2: SURVEY INTERFACE MANIPULATION (IN ORDER: UNPROFESSIONAL, BASELINE, PROFESSIONAL)

NOTE.—Color version available as an online enhancement.
REFERENCES


