

Running head: The best of strangers

The Best of Strangers:  
Context-dependent willingness to divulge personal information

Leslie K. John<sup>1</sup>, Alessandro Acquisti<sup>1</sup> and George Loewenstein<sup>1</sup>

Word count: 2499 (including abstract, text, figure legends, references, and notes)

*Publish Version Can Be Found:  
Journal of Consumer Research  
37(5), 858-873 (2011)  
<http://doi.org/10.1086/656423>*

---

<sup>1</sup> Carnegie Mellon University

## Abstract

Contrary to the assumption in much social science research that people have stable, coherent, preferences with respect to privacy, we find that concern about privacy, measured by divulgence of private information, is highly sensitive to contextual factors. We report results from 3 experiments, one of which was designed to elevate privacy concerns (paradoxically through assurances) and two of which were designed to suppress privacy concerns. This research raises serious questions about whether individuals will be able to navigate the increasingly complex issues of privacy in a self-interested fashion.

Individuals can be remarkably cavalier about divulging private information. The U.S. federal government has enhanced patient privacy through HIPAA, yet patients routinely sign consent forms without reading them. On the web, users routinely check the “I agree” box following privacy disclosures without reading them. In other cases, people disclose private information with little apparent awareness that they are doing so: physical locations can be tracked through electronic toll collection systems and cell phones, often unbeknownst to consumers. And going far beyond a lack of concern for privacy, many people actively seek out opportunities to divulge. On online social networking web sites such as Facebook, people routinely broadcast sensitive information – from suggestive photographs to home addresses – despite, or perhaps because of, the medium’s status as one of the most identifiable forms of communication.

Yet, in other contexts, people are extremely – and perhaps overly – privacy conscious. During doctor’s visits, individuals become acutely aware of the sensitivity of their health-related behavior, leading them to withhold information (such as risky sexual behavior) integral to the diagnosis and treatment of health problems. Others resist answering government-sponsored censuses, despite their known value to society.

The topic of privacy has not only captured media attention, but also that of social scientists, who have tended to view privacy through the lens of the rational choice perspective. A considerable body of academic research on privacy has assumed that people have a stable preference for privacy, and that they make sensible and coherent tradeoffs between privacy and other desired goals (Posner 1981; Stigler 1980). Consistent with such a perspective, several economists have attempted to measure the monetary value people place on privacy (Danezis et

al. 2005; Hann et al. 2002). Implicitly adopting a similar perspective, political scientists (Westin 1991) and psychologists (Jourard et al. 1958) have constructed individual difference measures of concern for privacy.

In contrast to the assumption of stable, coherent preferences, a large literature in behavioral decision research and behavioral economics documents systematic inconsistencies in people's preferences, in domains ranging from consumer choice (Simonson et al. 1992) to the valuation of environmental amenities (Kahneman et al. 1993). This research shows that preferences are often influenced by contextual factors that are difficult to justify on a normative basis (Slovic 1995). For example, preferences depend on how choice alternatives are framed (Tversky et al. 1974) and how preferences are elicited (Tversky et al. 1990). Research has further identified a range of mechanisms through which contextual factors influence decision making, including altering the salience of information (Tversky et al. 1995), the types of comparisons evoked (Hsee et al. 1999), and the types of memories brought to mind (Weber et al. 2006).

Such 'context effects' tend to be especially pronounced when people are uncertain about their own values, which is likely to be the case for privacy. If the *material* value of privacy is already extremely difficult to estimate, the psychological value is likely to be even less well-defined, creating the kind of 'preference uncertainty' in which inconsistencies in judgment and decision making commonly emerge.

In fact, some privacy-related research has already identified phenomena inconsistent with a coherent valuation of privacy (Acquisti 2004). For example, in a phenomenon dubbed the "privacy paradox" (Norberg et al. 2007), people report that privacy is important to them, yet engage in behaviors that indicate a remarkable lack of concern. It is perhaps not surprising, then,

that individual difference measures of privacy preference have generally failed to reliably predict privacy-related behavior (Himmelstein et al. 1963; Jourard et al. 1958).

In three experiments, we tested whether concern about privacy, as measured by propensity to divulge private information, is influenced by subtle contextual factors. We used contextual manipulations designed to differ in the degree to which they make privacy concerns salient. In each experiment, participants provided identifying information (email addresses) and then indicated whether they had engaged in a series of sensitive, and in some cases illegal, behaviors.

The first experiment's manipulation was intended to make privacy concern more salient than it otherwise would be, with the prediction that this would decrease willingness to admit to having engaged in sensitive behaviors. The second and third experiments, in contrast, attempted to distract subjects from privacy concerns with the prediction that people would, as a result, be more likely to divulge personal information.

Experiment 1 investigates the impact of confidentiality assurances on responses to sensitive questions. In contrast to the natural intuition that assurances would increase disclosure, we predicted, instead, that their effect would depend on whether a person was already primed to think about privacy.

Prior to completing a survey ostensibly about “students’ attitudes toward schoolwork,” university students were randomly assigned to receive an initial consent warning designed to make privacy concerns salient, or no such warning. They were also randomly assigned to receive either a confidentiality assurance, or no such assurance (supporting text). Within-subjects, we manipulated question intrusiveness: half of the questions were relatively innocuous, while the others were intrusive, inquiring about potentially punishable violations of academic integrity.

We predicted that when privacy concern is made salient (through a consent warning), people would react negatively to the subsequent confidentiality assurance, leading them to disclose less. In contrast, when participants receive no such consent warning, we anticipated that privacy concern would not be primed, and hence confidentiality assurances would, if anything, increase disclosure. The hypothesis was restricted to the academic integrity questions because, unlike the innocuous questions, they elicited information that people typically prefer to keep private. This study complements research showing that assurances impact people's willingness to *participate* in surveys (Singer et al. 1992), by hypothesizing that they can affect *what* people divulge, and that their influence is limited to privacy-relevant (intrusive) topics.

The 3-way interaction between the consent warning (present vs. absent), assurance (present vs. absent), and question intrusiveness (innocuous vs. academic integrity) manipulations was significant ( $F(1,194) = 4.73, p = 0.031$ ). As predicted, the combination of the consent and assurance affected admissions to the intrusive questions ( $F(1,93) = 3.87, p = 0.05$ ), but not to the tame questions. Assurance participants who had been given a consent warning made significantly fewer affirmative admissions than those who had not ( $t(96) = 2.70, p = 0.008$ ). Specifically, the average proportion of intrusive questions answered affirmatively was 0.18 for the former group, and 0.32 for the latter. For example, of participants who received a consent warning, only 8.0% of assurance subjects admitted to having “cheated on an exam,” while the admission rate was higher, at 35.4% for those who had received no such assurance ( $\chi^2(2) = 10.93, p = 0.001$ ) (supporting text).

-----

Insert Figure 1 about here

-----

At the end of the survey, participants indicated whether they were concerned about their privacy. Consistent with admission rates, there was a significant interaction between the consent warning and assurance manipulations ( $F(1,192) = 3.93, p = 0.049$ ) (supporting text).

Specifically, when given the warning, assurance participants were significantly more concerned about their privacy than those who received no such warning ( $t(96) = -2.08, p = 0.040$ ; No warning:  $M=2.79$  out of 7; Warning:  $M=3.66$ ).

In contrast to Experiment 1, which sought to activate latent concern about privacy through a consent warning, Experiment 2's contextual manipulation attempted to reduce such concern. The study was inspired by news stories about postings on the Facebook group "20 reasons why a girl should call it a night" in which young women voluntarily post compromising pictures of themselves – pictures that, in most other contexts, they would be mortified to share. We inferred that the frivolous nature of the site encouraged self-revelation and suppressed concern for privacy.

University students were recruited on campus and completed a survey. Between-subjects, the survey's title and interface were manipulated: In the frivolous condition (intended to downplay privacy concerns), the survey was called "How BAD are U??", and was light-hearted-looking. In the baseline condition, the survey was set within a professional context and was entitled "Carnegie Mellon University Survey of Ethical Standards" (supporting text). In contrast to Experiment 1, all of the 15 behaviors were sensitive (e.g. "Have you ever tried cocaine?"). We hypothesized admission rates to be higher in the frivolous condition.

Relative to baseline, participants in the frivolous condition were on average 1.7 times more likely to admit to having engaged in behaviors than those in the baseline condition ( $t(135)$

= 2.83,  $p = 0.005$ ). For example, a participant in the frivolous condition was on average 2.03 times more likely to admit to having “ever taken nude pictures of [him]self or a partner ” (affirmative admission rate was 15.7% in baseline versus 31.8% in frivolous condition;  $\chi^2(1) = 4.90$ ,  $p = 0.022$ ). People, it seems, feel more comfortable providing personal information on unprofessional sites that are arguably particularly likely to misuse it.

The contextual manipulation of Experiment 3, like that of Experiment 2, attempted to suppress concern about privacy by asking questions in a fashion that was expected to not bring issues of privacy to mind. It pertains to situations in which people “leak” private information without being aware that they are doing so.

Participants were directed to the study by a link titled “Test your ethics” in the on-line version of the New York Times, and randomly assigned to one of three inquiry conditions. The control condition asked participants, point-blank, whether they had engaged in a series of behaviors. In the two covert inquiry conditions, they were not asked directly whether they had engaged in the behaviors, but were asked to rate how unethical they judged them to be. However, with the explanation that people can only evaluate objectively behaviors that they have or have not engaged in, subjects were instructed to rate the ethicality of the behavior only if they *had* (commission condition) or *had not* (omission condition) engaged in it. We also manipulated the intrusiveness of the questions, within-subjects.

The omission and commission questions were designed to make the issue of whether one had done the behavior seem secondary – almost like an afterthought – which we predicted would increase self-revelation. We also predicted that, because the omission condition is even more indirect than the commission condition, with respondents only responding for behaviors they have *not* engaged in, this condition would elicit even greater admission rates than the



commission condition. We anticipated these effects only for the sensitive questions; by contrast, admissions to the innocuous questions were expected to be similar across conditions.

The interaction between intrusiveness and inquiry condition was significant ( $F(4,1256) = 12.0, p < 0.0005$ ), and admission rates were significantly different between conditions for the sensitive questions ( $F(2, 628) = 8.05, p < 0.0005$ ). Relative to point blank, participants were 1.80 and 2.21 times more likely to make affirmative admissions in the commission and omission conditions, respectively. Post-hoc tests revealed the point blank admission rate to be significantly different from the commission ( $p = 0.003$ ) and the omission ( $p = 0.002$ ).

-----

Insert Figure 2 about here

-----

Admission rates to “making a false insurance claim,” for example, were 20.2% in the omission condition, 9.8% in the commission condition, and 5.2% in the point blank condition ( $\chi^2(2) = 29.13, p < 0.0005$ ). Likewise, admissions to “neglecting to tell a partner about a sexually transmitted disease from which one is currently suffering” were 13.5% in omission, 7.7% in commission, and 1.9% in point blank ( $\chi^2(2) = 24.74, p < 0.0005$ ).

In Experiment 3, people admitted to having engaged in more sensitive behaviors when the method of inquiry was covert. We replicated this effect in a follow-up study with a different covert inquiry manipulation (supporting text). Covert inquiries, it seems, do not trigger concerns about privacy, and hence promote disclosure.

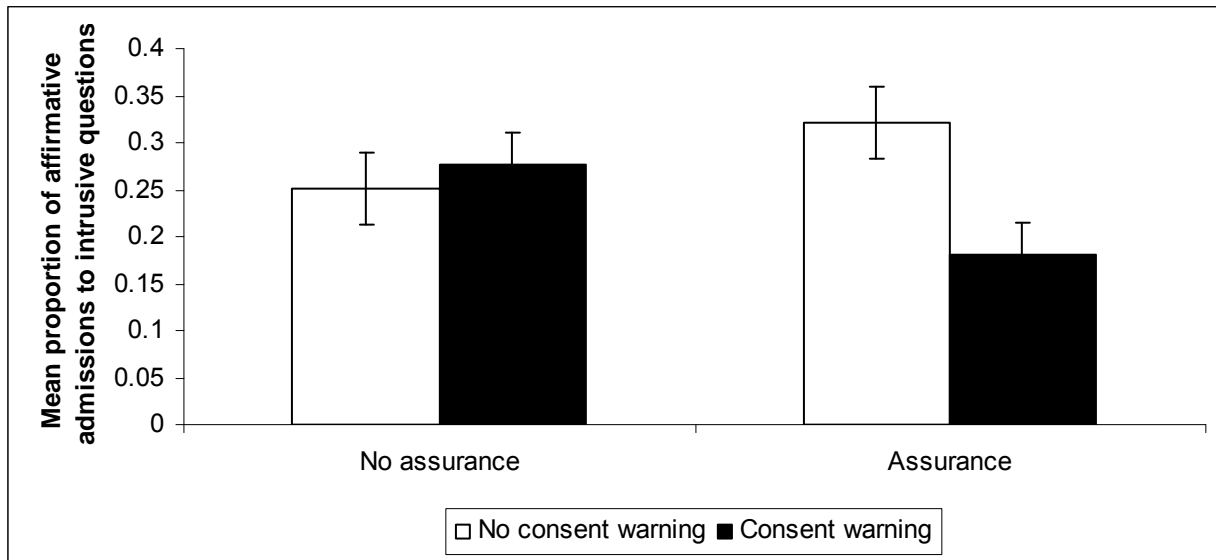
These experiments are consistent with the hypothesis that disclosure is responsive to contextual factors that have little to do with the actual costs and benefits of divulging information. In situations in which privacy concerns are activated, as was the case in

Experiment 1, it is likely that people will fail to divulge information even when the risks of doing so are low (as is the case under conditions of anonymity). However, when privacy concerns are suppressed, as was the case in Experiments 2 and 3, the same logic implies that people are prone to divulging information when it may not be in their self-interest to do so.

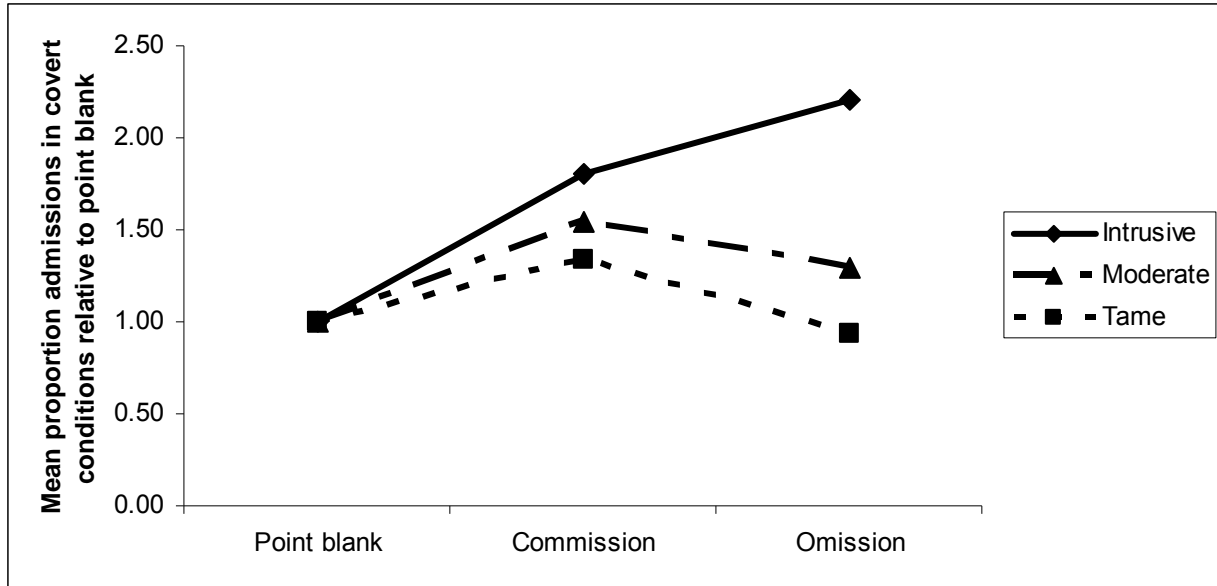
Issues of privacy are becoming increasingly prominent and complex given technological advances that have vastly expanded the capabilities for information sharing. By illustrating that disclosure of private information is influenced by contextual factors that have little, if any, normative justification, the current research casts serious doubt on whether individuals may be able to navigate this complexity according to their best interests

## Acknowledgments

We thank John Tierney, Howard Seltman, Sam Issacharoff, and Roberto Weber..



**Fig. 1.** The mean proportion of intrusive questions to which participants responded affirmatively (Experiment 1). Bars indicate standard error of the mean.



**Fig. 2.** The mean proportion of affirmative admissions in covert conditions relative to the point blank condition, by question intrusiveness (Experiment 3). Admission rates in the covert conditions indicate, on average, how much more likely people are to admit to having engaged in the behaviors relative to the point blank (hence point blank admission rates have been converted to 1).

## Supporting Text

### Materials and Methods

#### Experiment 1

##### *Method*

Prior to completing a survey ostensibly about “students’ attitudes toward schoolwork,” 199 university students were randomly assigned to receive an initial consent warning, or no such warning. They were also randomly assigned to receive either a confidentiality assurance, or no such assurance (Appendices S1, S2). Participants were presented with the consent warning (if applicable), followed by an assurance (if applicable) on the next page. Next, they were asked to provide identifying information in the form of a school (.edu) email address. Such an email address is both account traceable (the student’s account can be identified) and domain traceable (the student’s school can be identified). Most (81.3%) participants recorded their .edu email address. There was no significant difference between conditions in propensity to disclose email addresses, nor was there in admissions between those who did and did not provide their email addresses.

The survey consisted of thirteen questions varying in sensitivity (Appendix S3). All but one of the questions was answered on a dichotomous, yes/no response scale. The exception was “What is your GPA?” which was answered on a continuous scale. Seven of the questions were relatively tame, including “Are you currently taking at least four courses?” and “do you regularly attend classes?”. The six target sensitive items assessed

issues of academic integrity, including “Have you ever plagiarized text for any kind of assignment?” and “Have you ever cheated on an exam?”

A series of demographic questions was presented on the final page of the survey. There were no significant differences between conditions on any of the demographics.

Immediately following data collection, a research assistant blind to the hypotheses of the experiment coded the email addresses for identifiability. The email addresses were then separated from the data and destroyed.

We interpreted blank questions as missing data (i.e. neither admissions nor denials). However, the results did not change substantively when we re-ran the analyses with blank responses coded as denials.

See Table S1 for admission rates to intrusive items, by condition.

See Figure S1 for figure of self-reported concern for privacy, by condition (significant interaction and follow-up test reported in main paper).

## **Experiment 2**

### *Method*

Tables were set up in buildings across the Carnegie Mellon University campus; 137 students were asked to complete a brief web-survey in exchange for a chocolate bar as they walked by. The first screen of the survey informed participants that the survey was about college students, and that the experimenters were “interested in the types of behaviors that college students engage in.” They were also informed that at the end of the survey, they would be able to access its real-time, aggregate results, and that they “would also be given the opportunity to receive personalized results, including where you fall

relative to others on the traits and attitudes the survey measures.” These feedback opportunities created an incentive for the participants to respond truthfully.

Next, participants were randomly assigned to one of the two conditions: frivolous (designed to downplay privacy concerns) or baseline. The two versions differed only in their title and interface: In the frivolous condition, the survey was called “How BAD are U??”, and had a fun and light-hearted-looking interface, including a cartoon devil as the logo (Appendix S4). In the baseline condition, the survey had a more professional-looking interface: it was entitled the “Carnegie Mellon University Survey of Ethical Standards” and the university’s logo was displayed on each page (Appendix S4).

Participants were first asked a series of demographic questions. Two of the demographic variables were statistically different between conditions. There were significantly more males (64.3%) in the baseline condition than in the frivolous condition (47.0%); however, there was no difference in the number of affirmative admissions made between genders. There were also significantly more Asians (61.4%) in the baseline condition than the frivolous condition (42.4%), and Asians admitted to having engaged in fewer behaviors than other races (M number of affirmative admission for Asians = 2.7; for other races = 4.0;  $t(122) = 2.96$ ,  $p = 0.004$ ); however, the results we report in the main paper hold when race is included as a covariate.

On the subsequent pages, participants were presented with a series of sensitive behaviors and asked to indicate whether or not they had engaged in each one (Table S2).

At the end of the survey, participants were asked for their email address, and were given the opportunity to request personalized results of the survey. 48% of participants gave their email address. The proportion of participants who gave their email address in



Experiment 2 was much lower than in Experiments 1 and 3. This is likely because unlike the latter experiments, in which email addresses were requested at the beginning of the survey, in Experiment 2, they were requested at the end of the survey – after participants had admitted to engaging in deviant behaviors., and there were no significant differences between conditions in propensity to give an email address. Moreover, affirmative admission rates were not statistically different between participants who gave versus did not give their email address.

The aggregate, real-time results appeared on the last page of the survey. After completing the survey, each participant was given a chocolate bar.

Email addresses were immediately separated from the data and personalized results were emailed to participants who had chosen this option. Email addresses were then coded with respect to identifiability and then destroyed.

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

Consistent with the results we report in the main paper, it is also worth noting that participants' comments at the end of the survey were consistent with our hypothesis. Remarks in the frivolous condition included: "Fun stuff!!" and "That was hilarious." Conversely, none of the participants in the baseline condition commented on the fun nature of the survey.

### **Experiments 3 and 4**

### *Overview*

Experiments 3 and 4 examine how reports of sensitive behaviors are affected by how the questions are asked. Previous research has attempted to elicit estimates of the prevalence of sensitive behaviors through indirect methods of inquiry such as randomized response techniques (RRT) (see (Tourangeau et al. 2007) for an excellent review). RRTs typically add a random component to admission rates, and can only estimate aggregate responses. Such techniques protect individual privacy by creating ambiguity in the interpretation of an individual subject's responses.

In the present experiments, we investigate the idea that to the extent that a given method of inquiry makes privacy concerns salient, participants will be less forthcoming in responding to sensitive questions. To that end, Experiments 3 and 4 test the impact of a novel type of indirect inquiry on responses to sensitive questions. Unlike RRTs, which can only estimate the prevalence of the target behaviors in *aggregate*, the present method makes it possible to infer whether the *specific individual* has engaged in the behavior in question.

### *Method*

#### Manipulations

The experiment was a 3x3 condition mixed design, with question intrusiveness as the within-subjects factor, and mode of inquiry as the between-subjects factor.

*Question intrusiveness.* The intrusiveness of the 34 questions was determined by a separate sample of New York Times readers' ratings of question intrusiveness. Based on these ratings, we created three clusters of questions, representing different gradients of

intrusiveness: tame (10 questions), moderate (11 questions), and intrusive (10 questions). A list of the questions in the order in which they were presented is in Table S3. The items were presented in a pseudo-random order of intrusiveness.

*Method of inquiry.* In the overt, point blank condition, participants were asked to indicate the extent to which they had engaged in each behavior. For example, participants were presented with a question such as “Have you ever cheated on your tax return?” and indicated whether or not they had done this (yes/no).

In the two covert inquiry conditions, the items were phrased as statements rather than questions, and participants were asked to rate how unethical (if at all) they judged each item to be. For example, participants were presented with an item such as “Cheating on one’s tax return” and chose from four response options: not at all unethical, somewhat unethical, unethical, and very unethical. In the commission condition, participants were asked to only rate the behaviors in which they *had* engaged. Conversely, in the omission condition, they were asked only to rate the behaviors in which they had *not* engaged. Thus, in all three inquiry conditions, participants were presented with the same list of behaviors, and were asked – whether overtly or covertly – to indicate if they had engaged in each one (Table S4).

## Procedure

A link to the survey was posted on a New York Times Science columnist’s blog page. The link was titled “Test your Ethics” and was accompanied by the following description: “An online survey being conducted by social scientists at Carnegie Mellon University who are asking people’s views on what constitutes ethical behavior.” This

cover story served to distract participants from thinking of privacy from the outset, and to provide a realistic rationale for the experiment, thus helping to establish its credibility. Upon clicking the link, 1029 participants were randomly assigned to one of four conditions.

Participants were asked to provide their email address and then to indicate whether they would like to receive “personalized results of the experiment, including where you fall relative to others on the traits and attitudes the survey measures.” This served as an incentive for participants to respond truthfully.

Participants were first presented with a series of demographic questions. There were no significant differences between conditions with respect to any of the demographics.

Next, all participants were presented with an instruction page and were required to indicate that they had read and understood the instructions prior to continuing with the survey.

In the point-blank condition, subjects were further told that in an effort to get a baseline measure of the frequency of each behavior, they would be asked to indicate if they had (i.e. at least once) engaged in each behavior.

In the covert conditions, participants were told that in an effort to determine which types of behaviors are seen as more or less ethical, they would be asked to rate the extent to which they believed each behavior is unethical. In the commission condition, they were further informed that “because people are sometimes not objective about behaviors in which they *have not* personally engaged, we are only interested in your ratings of behaviors in which you *HAVE* engaged.” Conversely, in the omission

condition, they were informed that “because people are sometimes not objective about behaviors they *have* personally engaged in, we are only interested in your ratings of behaviors in which you *HAVE NOT* engaged.” (Appendix S5).

In both covert conditions, participants were then given two sample items – one in which they probably had not engaged – murdering someone – and one in which they probably had engaged – telling a white lie – and instructed how to respond appropriately. Additionally, a notice was visible at the top of each page of the survey, reminding commission participants to “please rate the behavior ONLY if you HAVE engaged in it AT LEAST ONCE” and omission participants to “please rate the behavior ONLY if you have NEVER engaged in it yourself”.

Personalized results were emailed to participants who had chosen this option. Email addresses were coded for identifiability (by a research assistant blind to the hypotheses of the experiment) and separated from the data. The email addresses were then destroyed.

The vast majority (93.7%) of participants gave an email address and there were no significant differences between conditions in propensity to give email address. The email addresses did not differ in identifiability between conditions and participants who provided highly identifiable email addresses made just as many affirmative admissions as those who had not ( $F(1,824) = .586, p = 0.44$ ). Moreover, such participants were no less likely to drop out of the survey than those who had not provided such identifying information ( $\chi^2 = 0.084, p = 0.959$ ).

Data analysis

In analyzing the data from an experiment of this type, it is critical to distinguish between items left blank to denote denials (in the commission condition) or admissions (in the omission condition), from those left blank for reasons unrelated to the construct of interest (e.g. dropping out of the survey).

Failure to make these distinctions could lead to erroneous conclusions. For example, an omission participant who left questions blank because he dropped out of the survey would appear as if he had admitted to having engaged in each of the behaviors, and the same behavior in the commission condition could be potentially misinterpreted as denial. This situation is particularly problematic, since it biases the results in the direction of our hypothesis (i.e. by artificially increasing intrusive admissions of omission subjects). Appendix S6 describes the two-step procedure we developed to correct for this problem; the results we report in the paper have been adjusted using this procedure.

#### ***Experiment 4: Follow-up to Experiment 3***

##### *Overview*

Experiment 4 replicates the results of Experiment 3 within a design that a) makes nonresponse completely independent from admissions/denials, precluding the need to adjust admission rates and b) rules out an alternative explanation pertaining to a desire to rate items (described below).

Though the results of Experiment 3 supported our hypotheses, an additional, unexpected, pattern emerged. In addition to the general finding that covert inquiry facilitates admissions to intrusive items, when collapsing across question type, admission rates tended to be the highest in the commission condition (Table S4). This pattern can be

explained by a desire to rate items: Such a desire would a) tempt omission and commission subjects to rate the items *regardless* of whether they had done them and b) occur across all intrusiveness levels. In the commission condition, rating an item is interpreted as an affirmative admission, while in the omission condition, it is interpreted as a denial. Thus, rating items indiscriminately with respect to whether one has actually done them would overestimate admission rates (regardless of intrusiveness level) in the commission, and underestimate admission rates in omission – exactly the observed pattern.

Moreover, this tendency may have been exaggerated by our subject recruitment method: Prospective participants were told that they would be (and in fact were) sent feedback on how their answers compared to those of other respondents. Thus, it seems that the recruitment process specifically targeted individuals who may be particularly driven to rate items.

## *Method*

### Manipulations

In a 2x2 mixed design, participants were asked whether they had engaged in a series of behaviors (Table S5). Similar to Experiment 3, the within-subjects factor was question type (tame versus intrusive), and the between-subjects factor was method of inquiry. All participants were asked to rate how unethical they deemed each behavior to be, and to indicate whether they had engaged in it.

*Question intrusiveness.* The intrusiveness of the questions was determined by a separate sample of New York Times readers' ratings of question intrusiveness. We used a

subset of 16 questions from Experiment 3. The items were presented in a pseudo-random order of intrusiveness.

*Method of inquiry.* In the overt condition, participants were first presented with the behavior, then asked, point blank, whether they had engaged in it, and then rated how unethical they deemed it to be (Appendix S7).

In the covert inquiry condition, the ethicality rating scale was presented twice. In the first case, it was labeled as: “If you have EVER done the behavior, how unethical do you think it was?” The second 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 place an answer in one of the two rating scales; hence, this method of covert inquiry did not rely on non-response (Appendix S7). The rating task was also completely independent from admissions/denials because all participants were asked to rate the items irrespective of whether they had engaged in the behavior.

## Procedure

Subjects were recruited through the same link used in Experiment 3. Upon clicking the link, 867 participants were randomly assigned to one of the two conditions. The procedure was the same as that of Experiment 3, except for the new method of inquiry manipulation as described above.

## Results

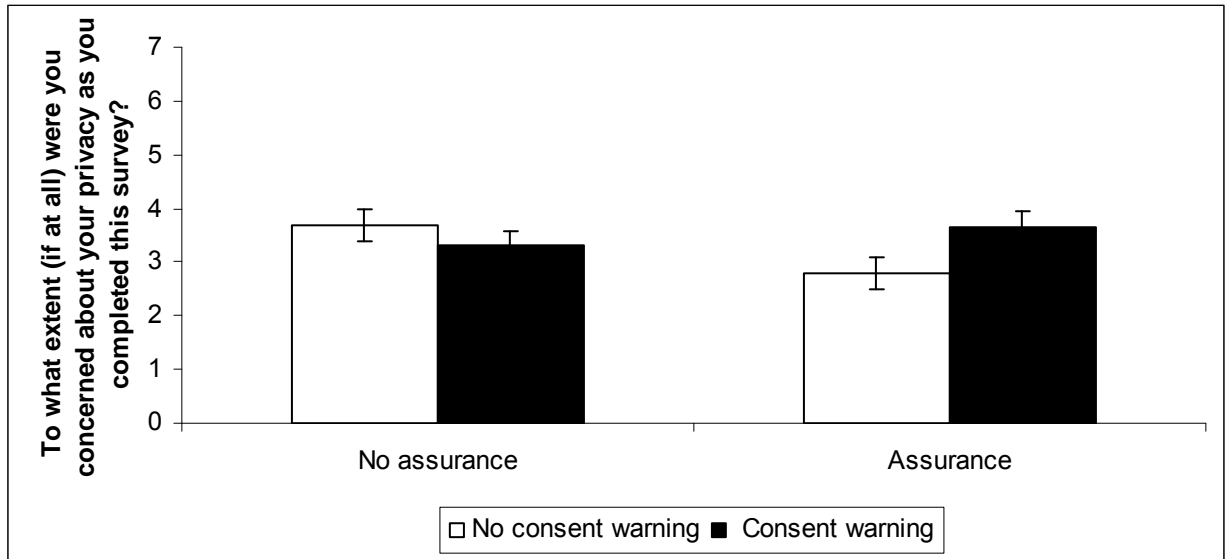
A 2x2 mixed ANOVA with question intrusiveness as the within-subjects factor and mode of inquiry as the between-subjects factor revealed a significant interaction of



method of inquiry and question intrusiveness ( $F(1, 864) = 26.21, p < 0.0005$ ). Follow-up analyses revealed that participants in the covert condition were significantly more likely to admit to having engaged in the sensitive items relative to those in the overt condition ( $t(827.83) = -3.20, p = 0.001$ ) (see Table S6 for admission rates to each intrusive question by condition) We used the t-test for heterogenous variances, since Levene's test for equality of variances was significant.. Specifically, participants in the covert condition were, on average, 1.48 times more likely to admit to engaging in the sensitive behaviors, relative to those in the overt condition. Although the follow-up analysis revealed a statistically significant difference between method of inquiry conditions for the tame questions, the effect is much less dramatic. Specifically, relative to the overt condition, the mean proportion of affirmative admissions in covert condition is 0.95.

This replication is all the more striking given the use of a weakened manipulation – after all, part of the potency of Experiment 3's manipulation is the facilitation of admission by omission (i.e. nonresponse).

**Figure S1.** Self-reported concern for privacy measured at the end of the survey (Experiment 1). Bars indicate standard error of the mean.



**Appendix S1.** Experiment 1: Consent warning

## **NOTICE OF CONSENT!**

**In compliance with informed consent procedures, this is a message to warn you about the nature of this questionnaire.**

You will be asked a series of questions, some of which pertain to breaches of academic integrity, and hence, are quite serious. Therefore, we are asking for your consent, but not for your signature. Please check the box below to indicate your consent to participate.

Please check the box below **only if you have read the above.**

I have read the above notice of consent.

*“No consent warning” condition received no such warning*

**Appendix S2.** Experiment 1: Assurance*Assurance condition*

Concerning the confidentiality and anonymity of your responses:

Please be advised that maintaining the confidentiality and anonymity of your responses is of the utmost importance to us. The following procedure will be used to maintain your anonymity in analysis, publication, and presentation of any results. Anonymity will be maintained during data analysis and publication/presentation of results by any or all of the following means: (1) You will be assigned a number as names will not be recorded. (2) The researchers will save the data file by your number, not by name. (3) Only members of the research group will view collected data in detail. (4) Any recordings or files will be stored in a secured location accessed only by authorized researchers.

*No assurance condition received no assurance paragraph.*

**Appendix S3.** Experiment 1: Survey

	<b>Yes</b>	<b>No</b>
1. Since high school (this will not be repeated on future questions), have you ever handed an assignment in late?		
2. Are you currently taking at least four courses?		
3. Have you ever plagiarized text for any kind of assignment?		
4. Have you ever let a classmate copy from you during an exam?		
5. Do you arrive late to class more often than the majority of your classmates?		
6. On average, do you find the number of students in your classes to be conducive to learning?		
7. Have you ever copied a classmate's homework?		
8. What is your grade point average?	GPA: _____	
9. Have you ever cheated on an exam?		
10. Have you ever requested an extension for an assignment?		
11. Do you regularly attend classes?		
12. Have you ever lied to a teacher in order to avoid taking an exam or handing in a term paper on time?		
13. Have you ever lied about your grade point average?		

**Appendix S4.** Experiment 2: Screenshots of survey interface by condition  
Frivolous condition:



**How BAD Are U???**

44%

4. Have you ever smoked marijuana (i.e. pot, weed)?

Yes

No

5. Have you ever "cheated" while in a relationship?

Yes

No

6. Have you ever driven when you were pretty sure you were over the legal blood alcohol level?

Yes

No

Baseline condition:

**Carnegie Mellon**

Carnegie Mellon University Survey of Student Behaviors



44%

4. Have you ever smoked marijuana (i.e. pot, weed)?

- Yes
- No

5. Have you ever "cheated" while in a relationship?

- Yes
- No

6. Have you ever driven when you were pretty sure you were over the legal blood alcohol level?

- Yes
- No

## Appendix S5. Experiment 3: Instruction pages by condition

Point blank condition:

CarnegieMellon

2. Carnegie Mellon University Survey on Ethical Standards

22%

**PLEASE READ THIS NOTE! This is not the usual yada-yada!**

This is a study of ethical standards.  
 In the next pages, you will be presented with a series of statements describing various behaviors. We are trying to get a baseline measure of the frequency of each behavior. **As such, we would like you to indicate IF you HAVE engaged in the given behavior (i.e. at least once).** (If the behavior in question does not relate to you at all, choose the "No" option as your answer.)

\* 1. I have read and understand these instructions.

Yes.

Commission condition:

CarnegieMellon

2. Carnegie Mellon University Survey on Ethical Standards

22%

**PLEASE READ THIS NOTE! This is not the usual yada-yada!**

This is a study of ethical standards.  
 In the next pages, you will be presented with a series of statements describing various behaviors. We are trying to determine which types of behaviors are seen as more or less ethical. We would like you to rate the extent to which you think each behavior is unethical. (If you believe that the behavior has nothing to do with ethics, choose the "Not at all unethical" option as your answer.)

**\*\*NOTE: Because people are sometimes not objective about behaviors they have not personally engaged in, we are only interested in your ratings of behaviors in which you HAVE engaged. Therefore, ONLY IF you HAVE engaged in the given behavior (i.e. at least once), please RATE it. Otherwise, please leave all remaining items BLANK.\*\***

**Example:** Imagine that you are asked to judge how unethical it is to tell a white lie, and imagine that you have told at least one white lie in your life. You think it is only somewhat unethical. Then, in the following question, you would **click** on the "somewhat unethical" box

**Telling a white lie.**

Not at all unethical   
  Somewhat unethical   
  Quite unethical   
  Extremely unethical

However, let's take an act that you have probably never committed: murdering someone. You believe that this is very unethical. However, in the following question, you would **NOT click** on the "extremely unethical" box, since you have never performed that behavior. Ok?

**Murdering someone.**

Not at all unethical   
  Somewhat unethical   
  Quite unethical   
  Extremely unethical

\* 1. I have read and understand these instructions.

Yes.



Omission condition:

**Carnegie Mellon**

22%

**PLEASE READ THIS NOTE! This is not the usual yada-yada!**

This is a study of ethical standards.

In the next pages, you will be presented with a series of statements describing various behaviors. We are trying to determine which types of behaviors are seen as more or less ethical. We would like you to rate the extent to which you think each behavior is unethical. (If you believe that the behavior has nothing to do with ethics, choose the "Not at all unethical" option as your answer.)

**\*\*NOTE: Because people are sometimes not objective about behaviors they have personally engaged in, we are only interested in your ratings of behaviors in which you have NOT engaged. Therefore, ONLY IF you have NEVER (i.e. not even once) engaged in the given behavior, please RATE it. Otherwise, please leave all remaining items BLANK.\*\***

**Example:** Imagine that you are asked to judge how unethical it is to tell a white lie, and imagine that you have told at least one white lie in your life. You think it is only somewhat unethical. Then, in the following question, you would **NOT click** on the "somewhat unethical" box, since you have performed the behavior.

**Telling a white lie.**

Not at all unethical     Somewhat unethical     Quite unethical     Extremely unethical

However, let's take an act that you have probably never committed: murdering someone. You believe that this is extremely unethical. Then, in the following question, you would **click** on the "extremely unethical" box since you have never performed the behavior. Ok?

**Murdering someone.**

Not at all unethical     Somewhat unethical     Quite unethical     Extremely unethical

**\* 1. I have read and understand these instructions.**

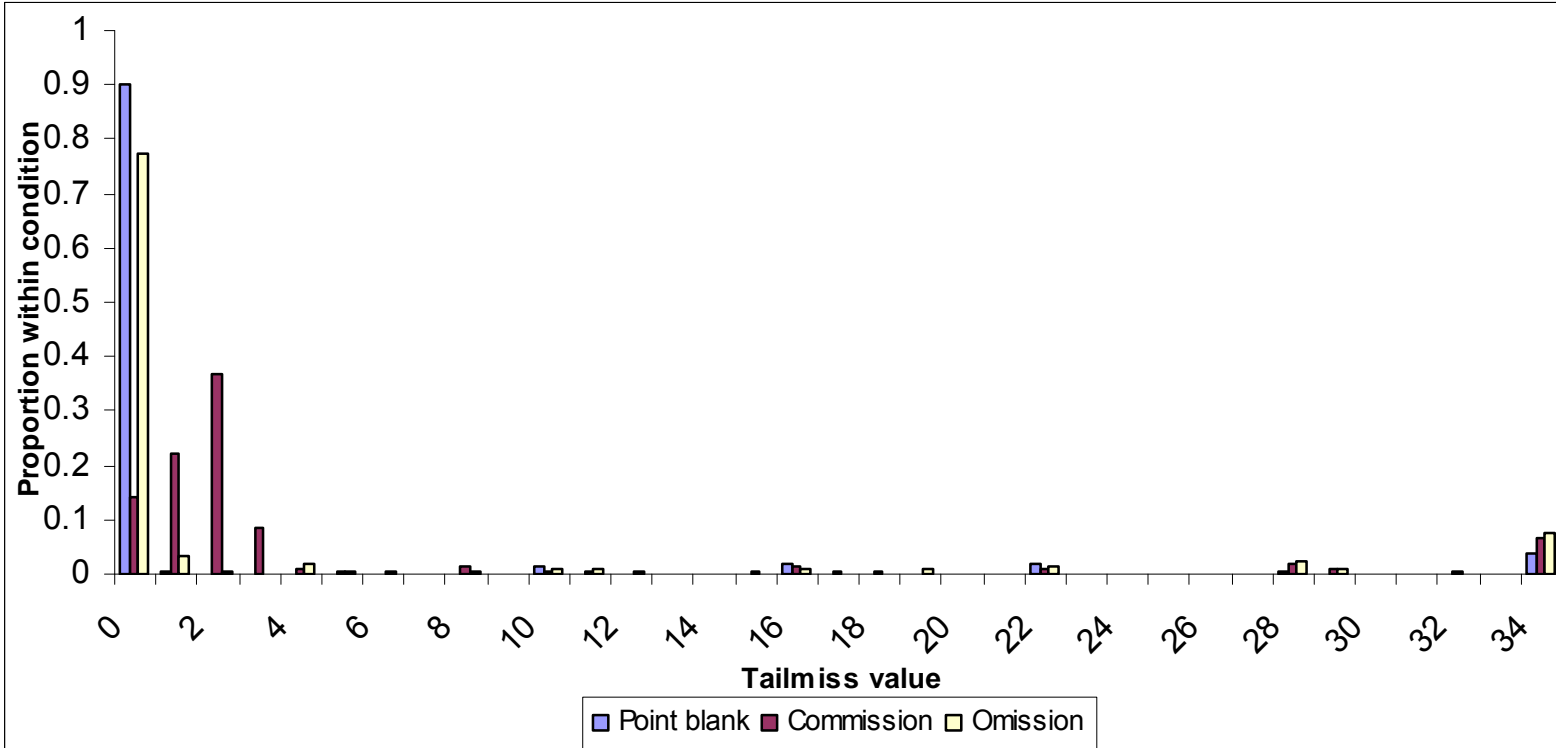
Yes.

**Appendix S6.** Experiment 3: Nonresponse adjustment procedure

*Step 1: Compare attrition rates between conditions*

In order to adjust admission rates in the covert inquiry conditions for those who dropped out of the survey, we first determined whether the attrition rates between conditions were equivalent. To do so, we calculated “tailmiss” – the number of consecutive questions a subject left blank, starting from the last one: question 34.

For point-blank participants, tailmiss represents the point at which a subject dropped out of the survey (i.e. stopped answering questions). In this condition, tailmiss took one of seven values: 34, 28, 22, 16, 10, 4, and 1 (see figure below). All but one of these values correspond to page breaks. The only exception is one participant who had a tailmiss value of 1.: there were six questions on the first five pages, and four on the last. A value of 28 for example, means that the participant left the last 28 questions blank, or in other words, answered the first page only. Thus, when point blank participants dropped out of the survey (with only one exception), they did so between pages.



For the covert inquiry conditions, the tailmiss measure of attrition is confounded by the method of inquiry, in which non-responses are interpreted as denials (commission condition) or admissions (omission condition). Despite this complication, reasonable guesses can be made to distinguish the covert condition participant leaving questions blank to denote denials (commission condition) or admissions (omission condition), from the participant who leaves the question blank because he has dropped out of the survey.

The critical logic of Step 1 is the following: the higher the tailmiss value, the more likely responses represent attrition, rather than non-response due to the construct of

interest. An omission participant with a tailmiss value of 28 for example, would appear to have performed all of the last 28 behaviors in the survey. It is more plausible that high tailmiss values represent non-response, rather than true denials (commission condition) or admissions (omission condition). In other words, long sequences of non-response, counting backward from the last question, are indicative of attrition.

We conducted chi-square analyses of the proportion of participants with tailmiss values corresponding to point blank drop-out points. Since none of these tests was statistically significant, we concluded attrition rates between conditions to be reasonably similar. Figures S2-S5 are also supportive of this conclusion: Given the construct of interest, it makes sense that there is a greater prevalence of tailmiss values falling between 0 and 8 in the covert conditions relative to the point-blank. Beyond 8 tailmisses, however, it becomes increasingly implausible that the tailmiss value represents true admissions or denials. Consistent with this notion, most covert condition tailmiss values greater than 8 (i.e. 72% or 32/44 in the commission condition, and 83% or 35/42 in the omission condition) correspond to page breaks (Figures S2 & S5).

*Step 2: Adjust covert results by point blank non-response rates*

Because attrition rates between conditions appeared to be similar, we adjusted the commission and omission results by the point blank non-response rates. Specifically, we calculated the point blank non-response rate for each of the 34 questions. Point-blank non-response rates can be assessed with certainty, since in this condition leaving a question blank is independent from the dependent variable.

Specifically, for the commission condition, where true non-response (i.e. leaving the item blank irrespective of whether one has done the behavior) is misinterpreted as *denial*, we adjusted the denial rate for the given question downward by the non-response rate in point-blank. Likewise, for the omission condition, where true non-response is disguised as *admission*, we adjusted the admission rate for the given question downward by the non-response rate in point-blank.

**Appendix S7.** Experiment 4: Question layout by condition

Overt inquiry:

<b>Carnegie Mellon</b>		19%
[REDACTED]		
<p><b>The behavior:</b> Letting a friend drive after thinking he or she had had too much to drink.</p> <p><b>Have you done this behavior?</b></p> <p><input type="radio"/> Yes</p> <p><input type="radio"/> No</p> <p><b>Please rate how unethical you think this behavior is:</b></p> <p><input type="radio"/> Not at all unethical</p> <p><input type="radio"/> Somewhat unethical</p> <p><input type="radio"/> Quite unethical</p> <p><input type="radio"/> Extremely unethical</p> <p><input type="radio"/> It depends</p> <p><input type="radio"/> Nothing to do with ethics</p>		

Covert inquiry:

<b>Carnegie Mellon</b>		19%
[REDACTED]		
<p><b>The behavior:</b> Letting a friend drive after thinking he or she had had too much to drink.</p> <p><b>A) If you have <i>EVER</i> done this, how unethical do you think it was?</b></p> <p><input type="radio"/> Not at all unethical</p> <p><input type="radio"/> Somewhat unethical</p> <p><input type="radio"/> Quite unethical</p> <p><input type="radio"/> Extremely unethical</p> <p><input type="radio"/> It depends</p> <p><input type="radio"/> Nothing to do with ethics</p> <p><b>B) If you have <i>NEVER</i> done this, how unethical do you think it would be, if you were to do it?</b></p> <p><input type="radio"/> Not at all unethical</p> <p><input type="radio"/> Somewhat unethical</p> <p><input type="radio"/> Quite unethical</p> <p><input type="radio"/> Extremely unethical</p> <p><input type="radio"/> It depends</p> <p><input type="radio"/> Nothing to do with ethics</p>		

**Table S1.** Experiment 1: Academic integrity question admission rates of participants who received a consent warning, by assurance conditions (listed in order of presentation)

Item	Percent affirmative admissions	
	No assurance	Assurance
3. Have you ever plagiarized text for any kind of assignment?*	28.0%	10.0%
4. Have you ever let a classmate copy from you during an exam?	28.6%	16.0%
7. Have you ever copied a classmate's homework?*	60.0%	38.0%
9. Have you ever cheated on an exam?*	22.0%	8.0%
12. Have you ever lied to a teacher in order to avoid taking an exam or handing in a term paper on time?	18.0%	20.0%
13. Have you ever lied about your grade point average?	10.0%	16.0%

\*chi square test significant at  $p \leq 0.05$  (2-sided)

**Table S2.** Experiment 2: Affirmative admission rates by question and condition (listed in order of presentation)

Item	Percent affirmative admissions	
	Baseline	Frivolous
1. Have you ever made up a serious excuse, such as grave illness or death in the family, to get out of doing something?	15.7%	27.3%
2. Have you ever looked at pornographic material?	71.4%	80.3%
3. Have you ever used sex toys?	8.6%	16.7%
4. Have you ever smoked marijuana (i.e. pot, weed)?	35.7%	43.9%
5. Have you ever "cheated" while in a relationship?	11.4%	21.5%
6. Have you ever driven when you were pretty sure you were over the legal blood alcohol level?	21.4%	24.2%
7. Have you ever taken nude pictures of yourself or a partner?*	15.7%	31.8%
8. Have you ever encouraged someone to drink when you were trying to seduce them?	15.7%	15.2%
9. Have you ever tried to peek at someone else's (e.g., a classmate's, boyfriend's, girlfriend's) email account without them knowing?	35.7%	45.5%
10. Have you ever fantasized about having violent non-consensual sex?	15.7%	22.7%
11. Have you ever tried cocaine?	2.9%	3.0%
12. Have you ever had sexual thoughts about a member of your same sex?***	14.3%	33.3%
13. Have you ever sold marijuana (i.e. pot, weed) to someone?	7.2%	7.6%
14. Have you ever watched someone while they undressed, without their knowledge?	4.3%	13.6%
15. Have you ever had anal sex?	4.3%	12.1%

\*chi square test significant at  $p \leq 0.05$  (2-sided)

\*\*\*chi square test significant at  $p \leq 0.01$  (2-sided)

**Table S3.** Experiment 3: Questions (listed in order of presentation)

Item number	Item	Item type T = Tame M = Moderate I = Intrusive
1	Have you ever stolen anything worth more than \$100?	M
2	Have you ever given preferential treatment (e.g., a favor, discount, etc.) to a person just because of how they look?	T
3	Have you ever claimed to have education that you didn't actually have?	M
4	Have you ever had sex with someone who was too drunk to know what they were doing?	I
5	Have you ever pretended not to see a beggar to avoid being seen as stingy?	T
6	Have you ever fantasized about having violent non consensual sex with someone?	I
7	Have you ever masturbated in a public place such as a public rest room?	I
8	While an adult, have you ever had sexual desires for a minor?	I
9	Have you ever gone on a date only to make somebody else jealous?	M
10	Have you ever watched someone through the window of their home without their knowledge?	M
11	Have you ever failed to do chores in a shared house or apartment?	T
12	Have you ever taken credit for someone else's work?	M
13	Have you ever tried to gain access to someone else's (e.g., a partner, friend, or colleague's) email account?	I
14	Have you ever neglected to tell a partner about a sexually transmitted disease from which you were currently suffering?	I
15	Have you ever let a friend drive after you thought he or she had had too much to drink?	M
16	In the last year, have you eaten meat, poultry, or fish?	T
17	While in a relationship, have you ever flirted with somebody other than your partner?	M
18	Have you ever downloaded pirated songs from the Internet?	T
19	Have you ever drunk so much that you got a hangover?	T
20	Have you ever lied about your income to someone?	M
21	Have you ever had anal sex?	I
22	Have you ever known about or witnessed a serious crime and failed to report it or stop it?	M
23	Have you ever visited an internet dating website, even just to check out what types of people might be available?	T
24	Have you ever littered in a public space?	T

25	Have you ever had a fantasy of doing something terrible (e.g., torturing) to someone?	I
26	Have you ever called in sick when you were not sick?	T
27	Have you ever viewed pornography when you were unsure whether the subjects were underage?	I
28	Have you ever made a false insurance claim?	I
29	Have you ever cheated on your tax return?	M
30	Have you ever failed to tip a waiter in a country in which tipping is customary?	T
31	Have you ever failed to turn the lights out at home or work, just because you were feeling lazy?	T
32	Have you ever looked at pornographic material?	M
33	Have you ever made up a serious excuse, such as grave illness or death in the family, to get out of doing something?	M
34	Have you ever had sex with the current husband, wife, or partner of a friend?	I



**Table S4.** Experiment 3: Item by item analysis of intrusive items, using nonresponse adjusted admission rates (listed in order of presentation)

Item	Percent affirmative admissions		
	Point blank	Commission	Omission
4. Having sex with someone who is too drunk to know what they are doing.**	4.3%	16.5%	10.9%
6. Fantasizing about having violent non-consensual sex with someone.**	20.1%	33.0%	20.9%
7. Masturbating in a public place such as a public rest room.**	32.3%	26.0%	22.5%
8. While an adult, having sexual desires for a minor.	15.8%	21.4%	20.9%
13. Trying to gain access to someone else's (e.g. a partner, friend, or colleague's) email account.	30.4%	27.0%	30.6%
14. Neglecting to tell a partner about a sexually transmitted disease from which one is currently suffering.**	1.9%	7.7%	13.6%
21. Having anal sex.	34.6%	33.7%	36.8%
25. Having a fantasy of doing something terrible (e.g. torturing) to someone.**	28.9%	44.6%	40.7%
27. Viewing pornography when unsure whether the subjects are underage.**	21.0%	26.0%	34.1%
28. Making a false insurance claim.**	5.2%	9.8%	20.2%
34. Having sex with the current husband, wife, or partner of a friend.**	9.1%	14.7%	22.9%

\*chi square test significant at  $p \leq 0.05$  (2-sided)

\*\*chi square test significant at  $p \leq 0.01$  (2-sided)

**Table S5.** Experiment 4: Questions

<b>Item number</b>	<b>Item</b>	<b>Item type</b> T = Tame I = Intrusive
1	Going on a date only to make somebody else jealous.	T
2	Letting a friend drive after thinking he or she had had too much to drink	I
3	Eating meat, fish, or dairy products.**	T
4	Neglecting to tell a partner about a sexually transmitted disease from which one is currently suffering.	I
5	Lying about one's income or that of one's family to someone.**	I
6	Drinking so much as to get a hangover.	T
7	Having sex with the current husband, wife, or partner of a friend.	I
8	Cheating on one's tax return.**	I
9	Visiting an Internet dating website, even just to check out what types of people might be available.	T
10	Having a fantasy of doing something terrible (e.g. torturing) to someone.	I
11	Calling in sick when one is not sick.	T
12	Viewing pornography when unsure whether the subjects are underage.*	I
13	Knowing about or witnessing a serious crime and failing to report it or stop it.*	I
14	Downloading pirated songs from the Internet.	T
15	Making a false insurance claim.*	I
16	Failing to send a 'thank you' note after attending an event to which one had been invited.**	T

\*chi square test significant at  $p \leq 0.05$  (2-sided)

\*\*chi square test significant at  $p \leq 0.01$  (2-sided)

**Table S6.** Experiment 4: Affirmative admissions to intrusive items, by inquiry condition (listed in order of presentation)

Item	Percent affirmative admissions	
	Overt	Covert
2. Letting a friend drive after thinking he or she had had too much to drink	43.6	49.7
4. Neglecting to tell a partner about a sexually transmitted disease from which one is currently suffering.	3.5	5.6
5. Lying about one's income or that of one's family to someone.**	29.1	41.3
7. Having sex with the current husband, wife, or partner of a friend.	15.5	11.4
8. Cheating on one's tax return.**	12.9	21.6
10. Having a fantasy of doing something terrible (e.g. torturing) to someone.	49.3	55.2
12. Viewing pornography when unsure whether the subjects are underage.*	11.1	16.3
13. Knowing about or witnessing a serious crime and failing to report it or stop it.*	4.7	8.3
15. Making a false insurance claim.*	2.6	6.2

\*chi square test significant at  $p \leq 0.05$  (2-sided)

\*\*chi square test significant at  $p \leq 0.01$  (2-sided)

- Acquisti, A. "Privacy in Electronic Commerce and the Economics of Immediate Gratification," Proceedings of the ACM Conference on Electronic Commerce (EC '04), 2004, pp. 21-29.
- Danezis, G., Lewis, S., and Anderson, R. "How much is location privacy worth?," Fourth Workshop on the Economics of Information Security, Cambridge, MA, 2005.
- Hann, I.-H., Hui, K.-L., Lee, S.-Y.T., and Png, I.P.L. "Online information privacy: measuring the cost-benefit trade-off," 23rd International Conference on Information Systems, 2002.
- Himmelstein, P., and Kimbrough, W.W. "A study of self-disclosure in the classroom," *Journal of Psychology* (55) 1963, pp 437-440.

- Hsee, C., Loewenstein, G., Blount, S., and Bazerman, M. "Preference reversals between joint and separate evaluations: A review and theoretical analysis," *Psychological Bulletin* (125:5) 1999, pp 576-590.
- Jourard, S.N., and Lasakow, P. "Some factors in self-disclosure," *Journal of Abnormal and Social Psychology* (56) 1958, pp 91-98.
- Kahneman, D., Ritov, I., Jacowitz, K.E., and Grant, P. "Stated willingness to pay for public goods," *Psychological Science* (4:5) 1993, pp 310-315.
- Norberg, P.A., Horne, D.R., and Horne, D.A. "The privacy paradox: Personal information disclosure intentions versus behaviors," *The journal of consumer affairs* (41:1) 2007, pp 100-126.
- Posner, R.A. "The Economics of Privacy," *American Economic Review* (71:2) 1981, pp 405-409.
- Simonson, I., and Tversky, A. "Choice in context: tradeoff contrast and extremeness aversion," *Journal of Marketing Research* (29:3) 1992, pp 281-295.
- Singer, E., Hippler, H.-J., and Schwarz, N. "Confidentiality assurances in surveys: Reassurance or threat?," *International Journal of Public Opinion Research* (4) 1992, pp 256-268.
- Slovic, P. "The construction of preference," *American Psychologist* (50:5) 1995, pp 364-371.
- Stigler, G.J. "An Introduction to Privacy in Economics and Politics," *Journal of Legal Studies* (9) 1980, pp 623-644.
- Tourangeau, R., and Yan, T. "Sensitive questions in surveys," *Psychological Bulletin* (133:5) 2007, pp 859-883.
- Tversky, A., and Kahneman, D. "The framing of decisions and the psychology of choice," *Science* (211:4481) 1974, pp 453-458.
- Tversky, A., Sattath, S., and Slovic, P. "Contingent weighting in judgment and choice," in: *Choices, Values, and Frames*, D. Kahneman and A. Tversky (eds.), Cambridge University press, Cambridge, 1995.
- Tversky, A., Slovic, P., and Kahneman, D. "The causes of preference reversal," *American Economic Review* (80:1) 1990, pp 204-217.
- Weber, E.U., and Johnson, E.J. "Constructing preferences from memories," in: *The construction of preference*, S. Lichtenstein and P. Slovic (eds.), Cambridge University Press, New York, 2006, pp. 397-410.
- Westin, A.F. *Harris-Equifax Consumer Privacy Survey 1991* Equifax Inc., Atlanta, GA, 1991.