

Misplaced Confidences: Privacy and the Control Paradox

Laura Brandimarte
Alessandro Acquisti
George Loewenstein

Under review – Please contact the authors for the most recent version.

A previous version of this manuscript was selected by the Future of Privacy Forum in its “Privacy Papers for Policy Makers” competition, and was presented at the Ninth Annual Workshop on the Economics of Information Security (WEIS).

Abstract

We introduce and test the hypothesis that mere control over *publication* of private information affects individuals’ privacy concerns and their propensity to disclose sensitive information even when the *objective* risks associated with such disclosures do not change or, in fact, worsen. We designed three experiments in the form of online surveys administered to students at a North-American University. In all experiments we manipulated the participants’ control over information publication, but not their control over the actual access to and usage by others of the published information. Our findings suggest, paradoxically, that more control over the *publication* of their private information decreases individuals’ privacy concerns and increases their willingness to publish sensitive information, even when the probability that strangers will access and use that information stays the same or, in fact, increases. On the other hand, less control over the *publication* of personal information increases individuals’ privacy concerns and decreases their willingness to publish sensitive information, even when the probability that strangers will access and use that information actually decreases. Our findings have both behavioral and policy implications, as they highlight how technologies that make individuals feel more in control over the publication of personal information may have the paradoxical and unintended consequence of eliciting their disclosure of more sensitive information.

Keywords: privacy, control, paradox, Web 2.0 applications, online social networks, experimental design

We gratefully acknowledge research support from the U.S. Army Research Office under Contract DAAD190210389 through Carnegie Mellon CyLab, and TRUST (Team for Research in Ubiquitous Secure Technology), which receives support from the National Science Foundation (NSF award number CCF-0424422) and the following organizations: AFOSR (#FA9550-06-1-0244), BT, Cisco, DoCoMo USA Labs, EADS, ESCHER, HP, IBM, iCAST, Intel, Microsoft, ORNL, Pirelli, Qualcomm, Sun, Symantec, TCS, Telecom Italia, and United Technologies.

1. Introduction

Modern information technologies grant us great power to broadcast our personal information to the world, but afford us much less control on how that information, once disseminated, will be used. Poorly thought emails, sent to an acquaintance in the heat of the moment, are forwarded to others causing embarrassment and regret; raunchy photos uploaded to a supposedly private network are made public; sensitive data privately revealed to companies are breached and stolen. And yet, most of us, lost in the immediately gratifying control over tweets, blogs, and status updates, seem to spend little attention to the relative lack of control upon the longer term, yet more tangible consequences of our public disclosures. We even react in seemingly contradictory manners, when we feel that our right to decide what information should be disseminated, and how, has been violated. We do so, regardless of the objective privacy (or lack thereof) of that information: Many online social networks users compulsively compile and update detailed personal profiles, revealing plenty of private and even sometimes embarrassing information; however, how would they react if the *same* information they willingly disclosed had been published by *another* party?

The contrast between control over the *publication*, and control over the *access* and *usage* of personal information, is at the center of our enquiry. Users of the Facebook busted in outcry when, in 2006, the social network launched a feature called News Feed (Solove, 2007; Hoadley et al., 2010). The News Feed posts on the users' front page their last activities and those of their Facebook "friends:" whether they have changed their relationship status from "single" to "engaged," whether they have posted new pictures or new videos, whether they have become "friends" with somebody, and so on. Notably, all this information was accessible even before the introduction of News Feed – although one would only see it exploring each of his friends' profiles. However, users reacted against the introduction of this feature, as if their privacy concerns strongly depended on the type of information revelation the system enacted: a *pull* model, where one has to go and search for information if he is interested, may be less privacy-invasive than a *push* model, where the same amount of information is provided by default, without one asking for it or putting any effort in searching for it (Malhotra et al. 1997).

Facebook users' reaction can be justified in rational terms: after all, a pull model implies higher transaction costs to retrieve information, which may in turn lower the likelihood that the information will be, in fact, retrieved and then abused by others. However, the reaction may also be affected by a less obvious and more ethereal discomfort with the inherent lack of control on access and usage that the News Feed suddenly made explicit. Most social networks users certainly understand that, in principle, the information they reveal online may be accessed by strangers. Indeed,

users display specific concerns about the way in which their private information may be accessed and used by strangers (Consumer Reports Poll, 2008; Norberg et al., 2007). However, at the moment of deciding whether to reveal personal information, the consideration that the information provided may later become available to individuals other than the intended recipients may remain as “latent” as such later access appears distant. The individual, therefore, may not feel concerned about the relative lack of control on future access to and usage of their data; such concerns may be trumped by the satisfactory perception of having control on the very act of revealing, and publishing, the information. Control over publication makes the lack of control over usage less salient. This paper tests such conjecture: we hypothesize that one of the psychological mechanisms that leads people to expose themselves to such a large extent is a control paradox on the information they reveal: Since we have control over the publication of our private information, we give less importance to control (or lack thereof) over the accessibility and use of that information by others.

In order to investigate issues of control in privacy decision making, we designed three survey-based experiments and administered them to students at a North-American University. Across all experiments, we manipulated the subjects’ feeling of control over information *publication*, without altering the actual conditions of *access* to and *usage* of the information they were asked to reveal. In two experiments, we decreased subjects’ feeling of or actual control, relative to a baseline condition. In one experiment, we increased the feeling of or actual control. We measured subjects’ propensity to reveal sensitive information about themselves as function of the amount of control they felt over the publication of their responses. Our results indicate that people may suffer from what we call a control paradox on personal information: more control over the publication of private information makes control over information access and use by others appear less salient, which consequently decreases individuals’ privacy concerns, and increases their willingness to publish sensitive information about themselves. Vice versa, individuals with less control over the publication of their private information may face increased privacy concerns, exhibiting lower willingness to publish sensitive information.

These results are significant on two levels. On a theoretical level, they challenge the traditional scholarly construct of privacy as “control” of personal information flows (see Section 2). Normatively, we have no doubt that granting individuals control on how their personal information is disseminated and used is an important (albeit not necessarily sufficient) condition for “privacy” protection. Positively, however, our results indicate that actually granting users control over their data is not guaranteed to make it easier for them to achieve some desired abstract balance between information revelation and information protection; if anything, the ultimate effect seems to be, paradoxically, to induce them to reveal more personal information, even when this may expose them to larger risks. On a practical level, our results challenge the view that Internet operators can soothe privacy concerns by simply affording users more control on their data. The Internet in general, and

Web 2.0 applications in particular, have granted individuals vast powers to disseminate personal thoughts and information to others. One of the Web 2.0 entrepreneurs' responses to the privacy concerns raised by their technologies has been the observation that such technologies also grant great user control in terms of to whom, when, and how to present one's online persona. For instance, in a 2004 interview, then Tribe.net CEO Mark Pincus claimed that "[s]ocial networking has the potential to create an intelligent order in the current chaos by letting you manage how public you make yourself and why and who can contact you" (Black, 2004). Similarly, in announcing "more privacy options" and settings that users could control, Facebook's official blog stated: "Today, we are introducing privacy changes that work towards our goal of giving you the control you need in order to share information comfortably on Facebook."¹ Our results, however suggest that affording more control to users may not necessarily help them to better protect their privacy, but rather it may induce them to reveal more sensitive information.

The remaining of this paper is structured as follows: Section 2 provides an overview of the existing literature about privacy, control and different explanations of people's contradictory privacy-related decisions. Section 3 formally presents our hypotheses of control paradox. Section 4 describes the empirical methodology we used in order to test those hypotheses, and shows and explains the results of our first experiment. Sections 5 and 6 describe design and results of our second and third experiment respectively. Section 7 concludes with final remarks and an agenda for future work.

2. Related literature

The concept of privacy of information has often been linked with the concept of control (Culnan, 1993; Smith et al., 1996). Quoting Alan Westin (1967), "privacy is the claim of individuals, groups, or institutions to determine for themselves when, how and to what extent information about them is communicated to others". Along these lines, Miller (1971) says that "the basic attribute of an effective right of privacy is the individual's ability to control the circulation of information relating to him." Fried (1984) defines privacy as "not simply an absence of information about us in the minds of others, rather it is the control we have over information about ourselves." And Elgesem (1996) believes that "to have personal privacy is to have the ability to consent to the dissemination of personal information." According to Lessig (2002), privacy, similarly to copyright, is a way of controlling information: "[j]ust as the individual concerned about privacy wants to control who gets access to what and when, the copyright holder wants to control who gets access to what and when."

¹" More Privacy Options," available at <http://blog.facebook.com/blog.php?post=11519877130>. Retrieved on April 7, 2010.

The digital age and the Internet have both increased but also reduced our ability to control the flow of information about ourselves. Solove (2007) notices that “with blogs and social network sites [...] there will be more instances when information we want to keep on a short leash will escape from our control.” The paradox of the Internet consists in the fact that, on the one hand, it gives people an extraordinary freedom of expression and communication; on the other hand, the Internet also constrains people, because it makes their private information more likely to be diffused in ways that can thwart future opportunities: “[a]s people use the freedom-enhancing dimensions of the Internet, as they express themselves and engage in self-development, they may be constraining the freedom and self-development of others – and even of themselves” (Solove, 2007).

Does this reduced control necessarily imply reduced privacy? In contrast with the literature referenced above, which defines privacy in terms of control, Laufer and Wolfe (1977) argue that “[a] situation is not necessarily a privacy situation simply because the individual perceives, experiences or exercises control.” Control should be one of the factors that determine privacy state (Dinev and Hart, 2004; Xu, 2007), but not necessarily should it be identified with privacy. Distinguishing privacy from control helps to understand how having control doesn’t necessarily mean having privacy and vice versa. In fact, Tavani and Moor (2001) state that “the concept of privacy itself is best defined in terms of restricted access, not control. Privacy is fundamentally about protection from intrusion and information gathering by others. Individual control of personal information, on the other hand, is part of the justification of privacy”. We don’t have control when we have to reveal our private information to the IRS, but we certainly expect the government to have high regard for our right to privacy. On the other hand, we have control when we create our profile on Facebook but we accept the idea that the information we post is no more private. However, also Tavani and Moor would agree that “[i]ndividual control plays a central role [...] in the management of privacy” (*op cit*, p. 8). Tavani and Moor acknowledge that such individual control has limits, but focus on the issue that “[t]he management of privacy requires controls beyond individual control that will ensure restrictions in access” (*op cit*, p. 8). Our experiments, instead, investigate how certain types of individual control (or feeling of control) on personal information may paradoxically lead individuals to overexpose themselves.

The relationship between privacy and control has been empirically addressed in few recent studies. Xu’s (2007) experiment tests the hypothesis that the negative relationship between privacy assurance mechanisms and privacy concerns is mediated by perceived control and self-construal construct, which can be instilled in consumers through privacy-enhancing technologies (PETs), industry self-regulation or government legislation. Her results suggest that the appropriate privacy protection device may change across types of people: those who evaluate themselves as independent-selves prefer PETs as tools to maintain control of personal information; those who evaluate

themselves to be less independent-selves and believe others to be powerful, influential, and responsible for events in others' lives prefer industry self-regulation and government legislation. Hoadley et al. (2010) analyze the psychological and behavioral reasons for Facebook users' protest after the introduction of the News Feed application. They find that, even though in general privacy concerns are enhanced simply because of the media coverage that privacy issues currently receive (see also Xu et al., 2008), Facebook users' discontent with the News Feed application was only partly explained by media coverage. Their survey-based study shows the effect of *perceived* control and easy access to information on privacy concerns: the easier the accessibility to some information, the lower the perceived control on that information, and therefore, the higher users' privacy concerns. Furthermore, a study published in September 2008 by the Consumer Reports National Research Center showed that most Americans are very concerned about what is being done with their personal information online and want more "control" over how their information is collected and used.² Norberg et al. (2007)' empirical tests of the so-called "privacy paradox" (the dichotomy between individuals' intentions to disclose private information and their actual behaviors; see also Spiekermann et al., 2001) is also related to control, in that consumers often voice concerns that their rights and ability to control their personal information in the marketplace are being violated. However, despite those complaints, it appears that consumers freely provide personal data under many conditions. Finally, Acquisti and Gross (2006) found that Facebook users who were *not* concerned about the privacy of the information they posted online lacked such concerns because they felt in control of the information provided.

3. Hypotheses

The privacy literature distinguishes between separate concerns associated with the *dissemination*, the *collection*, and the *access* or *usage* of personal information. Control on dissemination, is noted, does not necessarily imply control on who can access and use the information disseminated, and vice versa (Jiang et al., 2002; Noam, 1997). Hence, such different actions associated with personal data may elicit diverse concerns and reactions. While these distinctions are, in theory, quite clear and readily understood, in everyday practice people may inadvertently conflate concerns over different actions associated with their personal information. Consider, for example, the case of online social networks. Facebook provides its members with a strong feeling of control, because members can change every little detail of their default privacy settings, including what type of information will be available to whom. Not only can a Facebook member decide whether or not to post a particular piece of information on his or her profile, such as a mildly embarrassing photo, but she can also restrict the visibility of that photo to a limited list of "friends". However, the member will have no control on the

² http://www.consumersunion.org/pub/core_telecom_and_utilities/006189.html, accessed on February 16, 2010.

way in which that photo will be used by those “friends”, who could, for instance, post it somewhere else, or share it with other people, making it accessible to unintended third parties. Therefore, on the one hand, changing such detailed privacy settings can be time consuming and costly; on the other hand, the high level of control over information publication may induce users to feel an overall sense of confidence and could make control over publication more salient than control over access and use by others – and encourage users to publish private information. The distinction between control over publication versus control over access or use of information is crucial for our hypothesis of a paradox of control in privacy decision making: individuals reveal more when they feel in control over information dissemination, regardless of the actual level of control over access and usage.

We derived such hypothesis by combining observations from the above mentioned privacy literature, which distinguishes between information revelation, access, and usage, and the psychological and behavioral decision research literature, which has investigated the impact of saliency (Slovic, 1975; Klein, 1998) on individual choice. Consider an individual who is debating whether to post a suggestive personal photo on a social network. Imagine that the individual harbors some latent concerns about who will access that information and how they will use it: out of context and with the wrong audience, the photo may be embarrassing (consumers’ concerns about *access* and *usage* are some of the most pressing: Consumer Reports Poll, 2008 and Norberg et al., 2007). Rationally, individuals may realize that, although they have complete control on the decision of whether to post or not to post that photo, once the information is posted, they will not be able to maintain control on who will access it and how it will be used, no matter how granular the privacy settings provided by the system.³ Furthermore, one may reasonably argue that the costs associated with the mere dissemination of personal information are purely psychological (if a picture is uploaded to the Facebook and nobody is around to see it, does it make a sound?); on the other hand, it is how one’s personal information is accessed and used that will determine tangible consequences for the subject (for instance, a salacious photo will be taken out of context, leading to embarrassment, social stigma, denied job applications, or even blackmailing). However: access and usage of information by others are not certain events, and are distant in time. This reduces the saliency of those concerns, and therefore, we conjecture, may reduce their influence on the individual’s decision process (Slovic, 1975; Klein, 1998). On the other hand, the decision to publish is a factual and immediate choice, an act of free will. This, we argue, may provide a sense of empowerment, likely to affect the decision. The result would be that control over publication trumps over the relative lack of control over access and usage, even though the latter was the actual source of concern. In a sense, the two forms of control get conflated: the more control the individual feels over publication, the more concerns over access and usage become muted, and therefore the more likely the individual will be to reveal

³ Acquisti and Gross (2006) found that a significant portion of Facebook users speculated that their information could be accessed by strangers.

personal, sensitive information. On the other hand, the less control the individual feels over publication, the more concerns over access and usage become relevant, and therefore the less likely the individual will be to reveal personal, sensitive information. Just like illusion of control (Langer, 1975) makes gamblers feel overconfident and, thus, bet more (Goodie, 2005), control over information publication makes people feel empowered and overconfident about the effects of their “confidences” and induces them to accept higher risks (Campbell et al., 2004) and reveal more.

We call this the paradox of control on personal information. We detect it, by inference, manipulating (namely: increasing, in one experiment; and decreasing, in two other experiments) the level or feeling of control an individual has over the publication of personal information, without factually altering her control (or lack thereof) on access and usage (for instance, the objective likelihood that information will be later accessed or used). If we find that people respond to manipulations of control over information publication even when the objective conditions, terms, and likelihood of information access and usage are not changing, we can thus infer that control over publication affects the individual’s decision to reveal with larger weight. This leads us to formally state the following symmetric hypotheses:

H1: If people suffer from the paradox of control over private information, they will be willing to reveal more [less] if they have more [less] control over information publication, even if their control over access and use of that information by others remains unaltered.

The paradox of control will not only affect the amount of information that people are willing to publish online, but also the type of information that people share. If people feel an overall sense of control over the information they publish, they should be less concerned about revealing more sensitive information if they have control over publication, even if, again, control over information accessibility and use by others doesn’t change. In Section 4 we describe in detail how we established which information is more or less sensitive. This leads us to our second hypothesis:

H2: If people suffer from the paradox of control over private information, they will be willing to reveal more sensitive information if they have control over information publication, even if their control over access and use of that information by others remains unaltered.

The paradox of control hypothesis does not depend on the likelihood that published information will in fact be accessed or used, or even the rational awareness that an individual may not be able to control the usage and access to her data. The simple fact of being personally responsible for the publication of some information, and voluntarily and consciously deciding to disclose it, makes him feel endowed with the power of controlling it. This construct is therefore particularly relevant in

situations where information disclosure takes place through information technologies, because the intermediation of IT makes the act of sharing and disseminating information easier, almost effortless. This perception of control will make the individual feel comfortable with revealing private information - possibly more information than he would reveal if he wasn't subject to this paradox. Since this is a psychological more than a technological problem, the paradoxical implication is that Web 2.0 applications, by giving greater freedom and power to reveal and publish personal information, may trump on the concerns that people have over the control over access and usage of that information, independently of the amount of granular controls (even on access and usage) provided to users.

Naturally, the paradox of control is only one of the mechanisms that contribute to explain apparently inconsistent privacy-related decisions. Other concurring explanations for people's willingness to reveal private information can be derived from the literature and include wrong estimation of probabilities (prospect theory: Kahneman and Tversky, 1979), trust (Culnan and Armstrong, 1999), hyperbolic time discounting and immediate gratification (Acquisti and Grossklags, 2003; Acquisti, 2004), and perfectly rational models of decision making (Posner, 1978; Stigler, 1980). Our contribution attempts to motivate and understand privacy related decisions borrowing concepts from the psychology and behavioral economics literature.

4. Study 1

In order to test our hypotheses we designed three randomized experiments (Study 1, discussed in this section; Study 2, discussed in Section 5; and Study 3, discussed in Section 6). All three experiments were survey-based and subjects were recruited among students at a North-American University. The design of the first two experiments was essentially the same – and we present it in this section. The third experiment had a different design, which we describe in Section 6.

Experiments 1 and 2 manipulated subjects' sense of control in order to make them feel *less* in control over information publication, relative to a baseline condition of direct control. Experiment 3, instead, manipulated subjects' sense of control in order to make them feel *more* in control over information publication, relative to a baseline condition.

For the first two experiments, the questions contained in the survey were the same, although we slightly altered the wording for two questions. The surveys focused on students' life in the city around the university and on campus (the list of questions is provided in the Appendix). The justification for the survey was the creation of a new university networking website that would be launched at the end of the ongoing semester. Students were invited to become members of the

network. While the online social networks setup was chosen to be familiar and credible for our sample of subjects, our results are not confined to this particular social medium but apply, more universally, to other instances of information revelation, especially on the Internet.

The questions varied in terms of level of perceived privacy intrusiveness. In order to establish an objective level of intrusiveness of the questions, we recruited from the same population a separate sample of 31 subjects (who did not participate in the actual experiment), and asked them to rate all the questions according to their perceived privacy intrusiveness. Subjects were presented with a list of all the questions reported in the surveys and asked to rate them as either not at all intrusive, or moderately intrusive, or very intrusive. A table with the results of this pre-study can be found in the Appendix.

In the actual experiments, when subjects clicked on the link provided for the survey, they were randomly redirected to one of two conditions. The manipulation across the two conditions in each experiment consisted of the amount of control subjects had on the publication of the information they provided.

4.1 Design

The first page of the questionnaire in Study 1 contained three lines of instructions (reported in the Appendix), explaining that none of the questions required an answer, but that all the answers provided would be part of a profile that would appear on a new university networking website under construction, accessible to the university community only (students, professors, staff). In one condition (Condition 1) subjects were told that a profile would be automatically created for them, containing the information they provided, and that this profile would be published online once the website was completed, without any intervention by the researcher. In the other condition (Condition 2) subjects were told that a researcher would have collected the data, created a profile for them and published it on the network. The manipulation focused on how much control subjects had on the publication of their information. In Condition 1, subjects were given more control over the *publication* of their information: they decided exactly what to publish, if they wanted to publish anything at all. In Condition 2, on the other hand, an unknown “researcher” was responsible for the publication of their information: if subjects decided to disclose, they may have been somewhat less sure about what would happen to their information, because it ended up in possession of a researcher. This manipulation allows us to separate subjects’ feeling of control over revelation from the feeling of control over publication of private information. This manipulation was designed to emulate the condition of online social network users, who can obviously decide if and what to reveal. This type of

control is more salient than the lower level of control over how third parties are going to access and use this revealed information, which is neither a certain nor an immediate outcome.

Except for the control manipulation, the surveys were otherwise identical: 40 questions requiring approximately 10 minutes for completion. Since the website didn't really exist, in order to make the setup credible, some questions regarded everyday life of students at the university: what program they were enrolled in, what courses they were taking, how satisfied they were with their program, whether they practiced any sport on campus, and so forth. The survey contained 40 questions: seven highly intrusive questions, seven moderately intrusive questions and 24 non-intrusive questions which we used as additional controls for our hypothesis (details can be found in the Appendix). Sixteen questions had open-ended responses, 14 questions had binary – yes/no – responses, five were multiple choice questions and three were rating questions.

Our dependent variable was whether a subject answered to a certain question: to test hypothesis H1, we considered whether the subject decided to *answer* or not a question; to test hypothesis H2, we considered whether the subject decided to answer the *more privacy intrusive* questions. If, indeed, our subjects are affected by the paradox of control, they would be willing to answer more questions - and, specifically, more sensitive questions - in Condition 1 (where they felt personally responsible for the publication of that information) than in Condition 2 (where a researcher stood between them and the online publication). Note that subjects' objective privacy risks (which should also affect their propensity to answer sensitive questions) are associated with actual risks of disclosure; but the latter do not depend on who *publishes* the information (once the information is out there, it doesn't really matter who was responsible for its publication) but on who *accesses* that information and how they *use* it (Consumer Reports Poll, 2008 and Norberg et al., 2007). Our manipulation left control on accessibility and usage unchanged: more precisely, subjects had no control on that in either condition. One should therefore expect no difference in the propensity to answer (sensitive) questions across the two conditions, unless it is not the publication of their private information per se that disturbs the subjects, but the fact that someone else will publish it for them. In other words: lower willingness to reveal in Condition 2 would suggest that the saliency of control on information publication trumps over the lack of control over access and use by others.

Note that one possible confounding factor in this design is that subjects may be less willing to reveal private information in Condition 2 if they do not trust the unspecified "researcher." the researcher may not report the information correctly, may not securely store it, or may use it maliciously. All this is a consequence of lack of control, of course, but we can't exclude that lack of trust has a direct effect on willingness to reveal information, not mediated by control. We account for such potential confounds in the design of our second experiment.

4.2 Results

Twenty-nine subjects took the survey in Condition 1 (nine females and 18 males, two missing answers) and 32 took the survey in Condition 2 (ten females and 16 males, six missing answers). The modal age was 22 and 20 in Conditions 1 and 2 respectively, the average was 22.7 and 21.3 (sd: 5.4, 3; p-value = .18). Most of our subjects were born in the US and the percentage of Americans did not differ significantly across conditions (61.5% in Condition 1 and 60.7% in Condition 2, p-value = 0.87).

Figure 1 shows the percentage of subjects answering each of the questions in the two conditions. The average response rate (percentage of questions answered, averaged across subjects) was 84% in Condition 1 and 71% Condition 2. This might seem an unexpectedly high response rate, given the little amount of information that subjects received about the study. However, it is explained by the fact that most questions were not particularly intrusive. A cursory look at the histogram suggests that blue bars seem generally higher than red bars: this means that the response rate in Condition 1 was in general higher than in Condition 2, which would be consistent with our hypothesis H1. We found no significant difference in willingness to reveal across genders: in Condition 1, the average response rate for female was 89%, for male it was 85%. In Condition 2, the average response rate was approximately 83% for female and 81% for male.

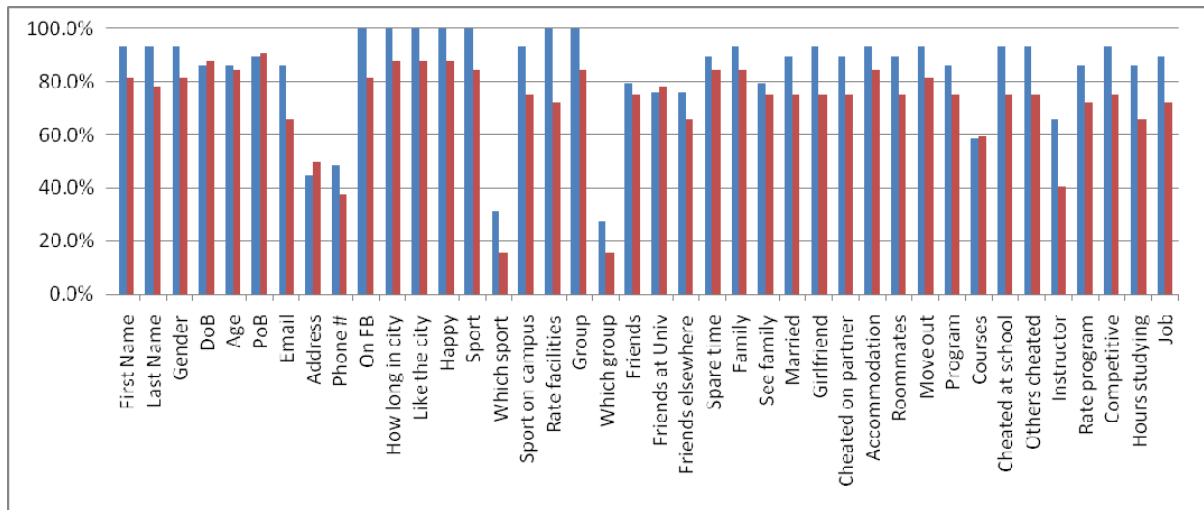


Figure 1: Percentage of subjects answering each question in the control condition (blue, 29 subjects total) and in the treatment condition (red, 32 subjects total).

To formally test our hypotheses we estimated a random effects panel Probit model. This methodology allows us to estimate the effect of the treatment (lack of control over information

publication) on the probability of a question being answered, taking into account the fact that the answers provided by one subject are not independent of each other. The unit of observation is the subject, while each of his/her answers constitutes a data point. This implies that answers by the same subject are very likely to be correlated with each other. We allow for that correlation when we estimate the variance-covariance matrix of the coefficients, assuming constant correlation between any two answers within a subject (exchangeable correlation structure: Liang and Zeger, 1986).

This is the equation of interest:

$$q_{ij} = \beta_0 + \beta_1 * Treatment_i + \beta_2 * Intrusive_j + \beta_3 * Treat_Int_{ij} + \beta_4 * Age_i + \beta_5 * Male_i + v_{ij},$$

$$v_{ij} = \alpha_i + u_{ij}$$

where $i=\{1,...,29\}$ for Condition 1 and $i=\{30,...,61\}$ for Condition 2 indexes the subject, and $j=\{1,...,40\}$ indexes the question. *Treatment* is a dummy variable for the experimental condition: it is equal to 1 if the subject was in Condition 2 and zero otherwise. *Intrusive* is a dummy variable that denotes the questions classified as most intrusive during the pre-study. *Treat_Int* represents the interaction between the treatment and the intrusive questions. *Male* is also a dummy, and *Age* is a discrete variable.⁴ Finally, q , our dependent variable, is a dummy set to 1 if a given question was answered and zero otherwise.

Our key variables of interest are *Treatment* and *Treat_Int*. The classical interpretation of an unobserved “latent” variable and a corresponding observed indicator variable applies: we don’t observe to what extent people suffer from illusion of control (call this unobserved continuous variable q^*), but we observe whether or not subjects provided with more control over publication (Condition 1) are more or less willing to answer the questions. Therefore, q_{ij} will be equal to 1 if subjects suffer from illusion of control ($q^*>0$) and zero otherwise. Assuming that the error term in the equation for the unobserved variable is normally distributed, we will obtain a standard Probit model. A negative coefficient on *Treatment* would then suggest that people are less willing to reveal private information if they are not personally responsible for its publication: so, it would support H1. A negative coefficient on *Treat_Int* - the interaction between treatment and level of intrusiveness - would suggest that this effect is larger for very sensitive information: so it would support H2; the interaction between *Treatment* and *Intrusive* tells us whether the treatment has a different effect on very intrusive questions versus non intrusive or moderately intrusive questions.

⁴ Since subjects were randomly assigned to one of the two conditions, we didn’t need to control for any effect other than the treatment, in order to obtain an unbiased estimate of the lack of control; but for descriptive analysis, and also as a randomization check, we included some demographic characteristics (age and gender).

The coefficients estimated with this model do not represent the marginal effects of the explanatory variables on the probability of the question being answered, but they are proportional to them, and the sign of the estimated coefficient will be the same as the marginal effect.⁵ For example, if the coefficient on the treatment variable is estimated to be -.4, this doesn't mean that the probability of answering a question is lower by 40% in Condition 2 as compared to Condition 1. It does mean that the effect of the treatment is negative and, in words, that lack of control over information publication decreases the probability of subjects answering a question.

Table 1.

	<i>Coeff</i>	<i>P-value</i>
Treatment	-.37*	.08
Intrusive	-.43**	.00
Treat_Int	-.03	.19
Age	.00	.98
Male	.08	.32
N= 61		Prob > χ^2 = .000

Random effects Probit coefficients of regression of response rate on treatment, with dummy for most intrusive questions, interaction and demographics.

** indicates significance at 10% level, ** indicates significance at 5% level.*

Table 1 reports the results of our estimation. The negative and significant coefficient on the dummy variable for very intrusive questions suggests that the sensitivity of the questions has a negative impact on the willingness to reveal: not surprisingly, subjects in Condition 1 were less willing to answer to very intrusive questions than to other questions. Mildly supporting our hypothesis H1, the coefficient on the treatment is negative and significant at the 10% level (as a randomization check, we noticed that removing demographics from the list of covariates doesn't change our main results). This indicates that when people do not feel personally responsible for the publication of their private information, they tend to answer fewer questions about themselves. The coefficient on the interaction between the treatment and the level of intrusiveness of the questions is negative, as predicted. However, it is not significant, suggesting that lack of control has the same negative effect on willingness to reveal very privacy intrusive information as on willingness to reveal non intrusive or moderately intrusive information (H2 is not supported).

⁵ In order to obtain the magnitude of the marginal effect, one has to evaluate the standard normal cumulative distribution function (CDF) at the estimated coefficients, and adjust for the correlation coefficient. This adjustment is due to the fact that in order to estimate the Random Effects, we can only estimate the β parameters up to a constant (see Arulampalam, 1998 for technical details).

Even though the coefficient on the interaction term is not significant, if we look at individual questions (details on regressions for individual questions can be found in Table 8 in the Appendix), we notice that the difference across conditions in the probability of answering a question is larger for many of the questions that were rated as very intrusive. For example, 65% of the subjects in Condition 1 answered to the question that asked about reporting a cheating student to the instructor, while only 40% of the subjects in Condition 2 answered that question. Similarly, 86% of the subjects in Condition 1 provided their email address, while only 65% did in Condition 2. For other very intrusive questions, though, the difference across condition was not so pronounced: for instance, 48% of the subjects in Condition 1 provided a phone number, compared to 37% in Condition 2.

5. Study 2

5.1 Design

Our first experiment provided mild evidence in support of a paradox of control hypothesis, but suffered from a potential confounding factor: the lower willingness to provide information, which is observed when there is lack of control over information publication, may be partly due to a lack of trust toward the researcher who is supposed to collect the information and compile the online profiles. Our second experiment resolved this confound. Study 2 mimicked Study 1's design. It consisted of the same survey we used in Study 1 and focused on the same population of subjects (although no subject who took part in Study 1 was allowed to participate in Study 2). Naturally, we changed the control manipulation. Condition 1 remained unaltered with respect to Study 1, while in Condition 2 participants were told that a 50% subset of the profiles created would have been randomly picked and published on the new university networking website.

Similarly to Study 1, subjects in either condition had no control over information accessibility and usage by others in either condition; however, subjects in Condition 2 had also no control over information publication. Due to the different manipulation, though, a rational decision making model would no longer predict that willingness to reveal should remain unaltered across conditions: one would expect that subjects would be less concerned about answering intrusive questions in Condition 2, because the probability of private information being published – and therefore available for access and use to others – is halved. If, on the other hand, our hypothesis H1 was correct and the revelation decision was based on the level of control that participants have over the information published on the network, then participants should be willing to reveal more in Condition 1: the feeling of overall control provided by a specific type of control – namely, over publication – decreases privacy concerns, which leads to larger willingness to reveal (Hypothesis 1), and especially so for sensitive information (Hypothesis 2).

Note, however, that a lower response rates in Condition 2 could, in principle, be attributable to two different factors: a) paradox of control; or b) less overall *interest* of the subjects in the creation of an online profile for the university network. This last factor could be playing a role because, if the probability of a profile being published on the network is only 50%, whatever benefit or risk the subject may incur by creating such profile will be halved, relative to the benefit or risk that subjects assigned to Condition 1 incur. This, in turn, may cause subjects in Condition 2 to simply care less about the study in the first place, and may push them to skip more questions. We designed our study so to be able to test which of the two factors above is supported by the data: if less interest is the main driver of lower willingness to reveal, then we should observe lower response rates specifically in the questions that required more time and effort for completion. Therefore, we added open ended questions such as “What program are you in?” or “Which courses are you taking?” In order to answer these questions, subjects had to reflect and then provide an (often long) list of items. If they cared about the study and the publication less in Condition 2 than in Condition 1, they would also be more likely to skip, or answer with fewer details, those questions. If, on the other hand, the paradox of control was the major factor in determining a lower willingness to reveal information, then we shouldn’t detect a lower response rate specifically for more time consuming questions.

We summarize below a hypothesis which will then be alternative to H1:

H1b: If people care less about the study overall, they will be willing to reveal more if they have control over information publication, and particularly so for time-consuming questions.

5.2 Results

One hundred and thirty-two subjects took part in our second experiment. Of those subjects, 67 were randomly assigned to condition 1 (34 females and 29 males, 4 missing answers) and 65 took the survey in condition 2 (28 females and 33 males 4 missing answers). The age distribution of subjects in the two conditions was described by the following statistics: the mode was 19 and 20 in Conditions 1 and 2 respectively, the average was 21.4 and 21.6 (not statistically different: $p\text{-value} = .96$), the standard deviation was 2.85 and 2.86. Similarly to Study 1, the majority of subjects were born in the US and the percentage of Americans did not differ significantly across conditions (57.6% in Condition 1 versus 51.6% in Condition 2, $p\text{-value} = .51$).

The average response rate was 89% and 87% in Conditions 1 and 2 respectively. We found no significant difference in willingness to reveal across genders: in Condition 1, the average response rate for female was 90%, while for male it was 95%. In Condition 2, the average response rate was approximately 87% for both sexes. Figure 2 shows the percentage of subjects answering each of the

questions in the two conditions. Looking at the histogram, we notice that the difference in response rate across conditions seems substantial for specific questions, but negligible for all other questions. In order to formally test our hypotheses we again estimated a random effects Probit model.

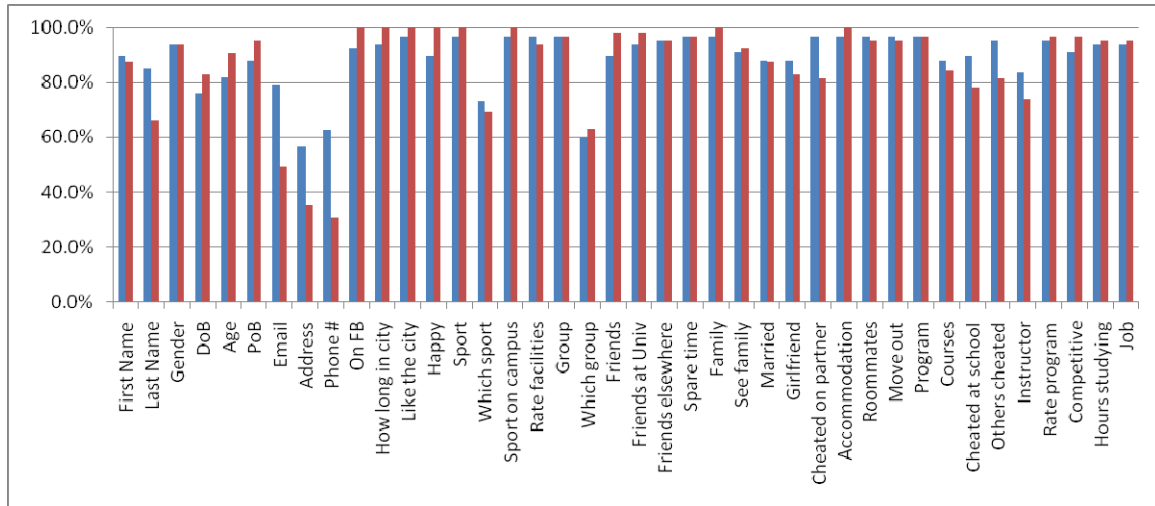


Figure 2: Percentage of subjects answering each question in the control condition (blue, 67 subjects total) and in the treatment condition (red, 65 subjects total).

Table 2 shows the results of our panel random effects Probit regression: the treatment has a negative effect on willingness to reveal private information, thus supporting H1. Furthermore, supporting hypothesis H2, the coefficient on the interaction of the treatment with the dummy for highly intrusive questions is negative and significant. Similarly to what we found in Study 1, also the coefficients on the treatment variable and on the dummy for privacy intrusive questions are negative and significant.

Table 2.

	<i>Coeff</i>	<i>P-value</i>
Treatment	-.25**	.05
Intrusive	-.64**	.00
Treat_Int	-.67**	.00
Age	-.02	.28
Male	.20*	.10
N= 132		Prob > χ^2 = .000

Random effects Probit coefficients of regression of response rate on treatment, with dummy for most intrusive questions, interaction and demographics.

** indicates significance at 10% level, ** indicates significance at 5% level.*

In order to test whether H1 or H1b is supported, we also estimated the model including two more variables: a dummy for time consuming questions (equal to 1 for questions 30 and 31, zero otherwise) and its interaction with *Treatment*. If the coefficient on this interaction is significant, then H1b would be supported, otherwise the data would suggest that indeed illusion of control is the reason for lower willingness to reveal in Condition 2. As we can see from Table 3, our estimates of the other coefficients remain similar to those estimated in Table 2, and the coefficient on the interaction between the treatment and the dummy for time-consuming questions is not at all significant.

Table 3.

	<i>Coeff</i>	<i>P-value</i>
Treatment	-.24*	.06
Intrusive	-.61**	.00
Treat_Int	-.69**	.00
Age	-.02	.28
Male	.19*	.10
Time	5.1	.99
Treat_Time	-4.6	.99
N= 132		Prob > χ^2 = .000

Random effects Probit coefficients of regression of response rate on treatment, with dummy for most intrusive questions and time consuming questions, interactions and demographics.

** indicates significance at 10% level, ** indicates significance at 5% level.*

Moreover, we found no significant difference in the response rates across conditions in the probit regression for individual more time consuming questions (which were questions where we asked which program and courses the subjects were enrolled in) . Therefore, the data support hypothesis H1 (paradox of control).

Looking at individual questions (details on regressions for individual questions can be found in table 9 in the Appendix), we notice that the difference in response rates across conditions is larger for very intrusive questions: for example, 63% of the subjects in Condition 1 provided a phone number, while only 31% did in Condition 2. Similarly, 79% of the subjects provided an email address in Condition 1, while only 49% did in Condition 2. The difference across conditions was smaller for other moderately or not intrusive questions. For instance, 88% of the subjects in Condition 1 answered

to the question on whether they had a girlfriend or a boyfriend, and the response rate in Condition 2 was 83%.

A possible confounding factor that deserves careful evaluation for both experiments is the possibility that some subjects may be unaware of the consequences of disclosing certain type of private information (for the relevance of awareness in privacy decision making see, for example, Malhotra et al. 2004 and Culnan, 1995), particularly privacy intrusive information. Let us assume, conservatively, that some online social networks users do not know about the consequences of sharing private information with their “friends.” Still one would have to consider that awareness doesn’t always explain people’s behavior: tobacco smoking or drug use have been scientifically proved to damage people’s health and cause addiction, and yet many people consume them. Similarly, everybody reckons that waste separation and recycling is a friendly attitude towards the environment, but not everybody does it. Knowing about the risks of disclosing private information doesn’t prevent everybody from doing it. Moreover, it seems that the level of awareness among users of online social networks has significantly increased over time: Acquisti and Gross (2006) found that less than 1% of Facebook users ever changed their privacy settings (based on 2005 data). According to Facebook’s Chief Privacy Officer, that percentage was at least 20% in March 2009 (as reported by Randall Stross in his New York Times article, March 7th, 2009, <http://www.nytimes.com/2009/03/08/business/08digi.html>), suggesting increased awareness and changes in behavior following the significant media interest towards privacy issues on the Facebook. These findings make lack of awareness harder to justify and lower any weight that this confounding factor may have in explaining the results of our experiments.

6. Study 3

6.1 Design

The design of our third, survey-based study differs substantially from the design of the previous two studies, but still aims at manipulating subjects’ feeling of control over the publication of their private information. Unlike the two previous experiments, which had subjects feel *less* in control over information publication, relative to a baseline condition of direct control, our third experiment included a baseline condition with “baseline” control and four treatment conditions with augmented (whether it be perceived or real) control over information publication.

The alleged motivation for the survey was that we were interested in studying “ethical behaviors” and that we would ask a series of questions related to this topic. The survey consisted of ten yes/no questions regarding more or less sensitive and moot behaviors, such as stealing, lying or

consuming drugs (the complete list, together with the exact wording used in each condition, can be found in the Appendix). Subjects were informed that none of the questions required an answer. Subjects were also told that the researchers were meaning to publish the answers provided by the participants in a Research Bulletin among the results of the study, but no detail was given as to whether this Bulletin would have been printed or published online, nor as to whom this Bulletin would have been visible/available to. What is relevant is that, similarly to the first two studies, subjects had no control over the access to, or the usage of, their information by others.

Besides the ten questions on ethical behaviors, subjects were asked to provide some demographic information (age, gender, country of birth and email address; email addresses were not saved in our results file to protect subjects' anonymity; however, we took note of whether a subjects had inserted or not a meaningful email address in the appropriate field) and some final questions needed as manipulation checks.

Subjects were randomly assigned to one of five conditions. What varied was the control subjects had over the publication of the answers they would provide.

In Condition 1 (our baseline condition, the one to which all the others will be compared), subjects were told that by answering a question, they would implicitly give the researchers permission to publish the answer provided. This therefore can be considered a condition of “baseline” control over information dissemination and publication, in that subjects could still decide not to answer any question (and therefore deny the researchers the ability to publish their answers to that question). In the other conditions, however, the control on the publication was made increasingly explicit and, in a sense, stronger.

In Condition 2, for each individual question, subjects were also explicitly asked to check a box, next to the question, to signal that they were willing to grant publication permission. However, the default option was that the answers wouldn't be published. This condition emulates several Web 2.0 applications that provide users with granular control on what to publish online. Hypothesis H1 would be supported if we observed that the willingness to answer *and* publish in Condition 2 is larger than the willingness to answer in Condition 1 (if this is especially true for the most privacy intrusive questions⁶, also hypothesis H2 would be supported). If a rational agent was concerned about the publication of some answers, she would refrain from answering if those answers were to be

⁶ In order to establish an objective level of privacy intrusiveness for the 10 questions on ethical behaviors, we followed the same procedure we used in Studies 1 and 2: we asked a separate set of 49 subjects, who did not participate in Study 3 afterwards, to rate the questions as either not intrusive at all, or moderately intrusive or very intrusive. Results of this pre-study can be found in the Appendix.

automatically published (condition 1), and she would decide not to publish them if she was explicitly asked to (Condition 2). Choosing not to answer in Condition 1 but, on the other hand, to provide an answer *and* allow its publication in Condition 2 would suggest that the simple fact of having control over information publication decreases privacy concerns and, paradoxically, leads to publication of a larger amount of information – which would then be accessible and available for use by others. (Note that participants in Condition 2, as well as those in the additional conditions presented below, could have chosen to answer a question but not to permit its publication; this allows us to separately study the effect of our manipulation on the propensity to answer – revealing the answers just to the researchers– and to publish – revealing the answers to all the readers of the purported Research Bulletin.)

One possible confounding factor is that subjects may incur too much cost (in terms of time spent participating) in checking every single publication permission box. Consequently, if we observed that they didn't allow the publication of their answers, it may have been because indeed they didn't want their answers to be published, or because they didn't want to spend so much effort in checking all boxes. This confound was removed in Conditions 3 and 4. In Condition 3, subjects were explicitly asked for publication permission, but not for each individual question: before answering the 10 questions on ethical behaviors, subjects were asked to check a box if they agreed to give the researchers permission to publish all their answers among the results of the study. Even though Condition 3 may have implied less granular control than Condition 2, it still represented an increased feeling of control relative to the baseline condition. Hence, similarly to what we explained above, allowing the publication (of all answers) and providing an answer in condition 3, while not responding in condition 1, would paradoxically indicate that the simple fact of being explicitly asked for publication permission, and therefore of feeling more control over information publication, increases willingness to reveal relative to the case where the publication permission is not explicitly requested. In other words, such results would suggest that the publication of private information only becomes an issue if it is not under the subjects' explicit control.

Condition 4 was identical to condition 2, but the default was that each and every answer provided would have been published. If the subject didn't want to grant publication permission, he had to check the box corresponding to the answer he didn't want to publish. We included this condition in order to account for default effects and *status quo* bias (tendency of people to choose the default option, regardless of what their true preferences are; see for instance Samuelson and Zeckhauser, 1998; Johnson et al., 2002; and Park et al., 2000). Since both Conditions 2 and 4 provide complete control over information publication, we would expect similar results in terms of willingness to answer a question. As regards willingness to publish, if the default effect is strong enough, we would observe on average a difference across the two conditions.

In all four conditions described above, subjects were informed that the demographic information provided by the subjects wouldn't have been published. The last condition – Condition 5 – was identical to Condition 3, but also explicitly asked for permission to publish demographic information. Subjects could click on a publication permission box for each and every of the following items: gender, age, and country of birth. This condition was included to test whether providing subjects with explicit control on the publication of *more* information would induce them in fact to decide to publish it. Notice that it doesn't really matter whether this information is perceived as privacy intrusive or not: what matters is that this information is *additional* to the one already provided about what we called “ethical behaviors”. Furthermore, providing certain demographic information increases the objective risks associated with answering the other questions in the survey, by lowering their probability of remaining anonymous: researchers have proved now and again that individuals can be uniquely re-identified based on combinations of otherwise non-uniquely identifiable personal information (Sweeney 1997).⁷ Observing similar willingness to answer as compared to Condition 3, associated with willingness to publish demographics, would support our hypothesis that providing people with explicit control over information publication decreases their privacy concerns and pushes them to answer and publish more, regardless of the control over information accessibility and availability for others.

Finally, we note that the literature on *status quo* and default settings biases would suggest that our subjects would be unlikely to click the “publication permission” boxes in Conditions 2, 3 and 5; or unclick the “no publication permission” box in Condition 4. In other words, to show a detectable effect, the paradox of control should be stronger than the tendency to choose the default option.

6.2 Results

Before presenting the estimations of our panel Probit model, we provide in Table 4 some descriptive statistics and qualitative results. Note that our main variable of interest was not, per se, the permission to publish, but whether subjects would answer the actual questions, as function of the permission to publish. First of all, we notice that the average response rate is much lower in condition 1 than in any other condition – which seems to suggest that hypothesis H1 is supported. Also, subjects seemed on average not very concerned about providing their email address.⁸

⁷ Sweeney (1997) showed that most US residents are likely to be uniquely identifiable based on gender, residence ZIP code, and date of birth alone. The literature on statistical re-identification (see Acquisti and Gross 2009) has since found several other examples of non-unique pieces of data, once combined, allowing the identification of individuals.

⁸ We didn't store email addresses, but let subjects believe we would. Subjects, however, were informed that their email address would not be published among the results of the study. In order to make sure that this was

Table 4.

<i>Experimental condition</i>	<i>Number of subjects</i>	<i>Average age</i>	<i>% Male</i>	<i>Average response rate (%)</i>	<i>Subjects providing email (%)</i>	<i>Subjects answering all questions</i>	<i>Subjects publishing all questions</i>	<i>Subjects publishing no question</i>
1	33	22.03	45.4	60.6	78.8	5 (15.1%)	-	-
2	36	22.11	50.0	96.1	80.5	28 (75.0%)	10 (27.8%)	10 (27.8%)
3	32	21.87	46.9	84.4	81.2	12 (37.5%)	32 (100%)	-
4	35	21.80	48.6	96.0	80.0	26 (74.3%)	19 (54.3%)	0 (0%)
5	33	22.09	54.5	83.3	87.9	13 (39.4%)	33 (100%)	-
Total	169	21.98	49.1	86.0	81.6	83 (49.1%)	94 (69.1%)	

Some descriptive statistics and qualitative results from Study 3.

Looking at the second to last column, we notice that all subjects in Conditions 3 and 5 clicked the unique publication permission box, and that all subjects in Condition 5 granted permission to publish all three demographic items. The pattern that seems to emerge from this striking result is that, as long as subjects are endowed with control over information publication, and this control is not too costly in terms of time/effort spent exercising it, subjects do not show concern in publishing private information. Note, again, that our focus was not on whether, per se, subjects clicked on the “control” boxes provided to them: for our analysis, the motives why our subjects chose or did not choose to allow publication are irrelevant; what we focus upon is whether, *once such control was given to the subjects*, the subjects were more or less likely to answer the actual survey questions. In analyzing such variable, we considered both overall response rates, as well as rates conditional on subjects’ having allowed their dissemination.

Moreover, and not surprisingly, looking at the last two columns in Table 4, we detect some effect of the default option. Nobody changed the default option (publication allowed) for all questions in Condition 4, and only 10 subjects changed the default state (publication not allowed) for all questions in Condition 2. This suggests that some subjects didn’t want to spend time/effort in checking all boxes for publication permission, and therefore their revealed preferences about information publication may be different from their true preferences. We tried to limit this (expected) effect, by asking a relatively short list of questions, so that the overall survey wouldn’t require a long time for completion, and asking for publication permission on top of the page requesting the actual yes/no answer (Figure 3 in the Appendix shows what the question matrix looked like in each condition).

clear to all subjects, we explicitly asked, as a manipulation check at the end of the questionnaire, whether they thought their email address would be published, and all subjects answered negatively.

In order to formally test our hypotheses, we used the same regression approach as the one employed to analyze the results of Studies 1 and 2. Table 5 summarizes our results.

Table 5.

	<i>Comparing conditions:</i>			
	<i>1 and 2</i>	<i>1 and 3</i>	<i>1 and 4</i>	<i>1 and 5</i>
Treatment	1.51** (.000)	1.92** (.000)	1.52** (.000)	.91** (.000)
Intrusive	-.85** (.000)	-.85** (.000)	-.85** (.000)	-.84** (.000)
Treat_Int	.59* (.071)	-1.21** (.002)	.44 (.177)	-.08 (.741)
Age	.01 (.753)	.03 (.521)	.003 (.942)	.05 (.158)
Male	.10 (.653)	-.11 (.593)	-.08 (.684)	-.03 (.861)
N	69	65	68	66
Prob > χ^2	.000	.000	.000	.000

Random effects Probit coefficients of regression of response rate on treatment, with dummy for most intrusive question, interaction and demographics. P-value in brackets.

** indicates significance at 10% level, ** indicates significance at 5% level.*

Supporting hypothesis H1, we find that the coefficient on *Treatment* is always positive and significant, suggesting that providing subjects with more control over information publication increases their willingness to answer a question. The coefficient on the dummy for very intrusive questions is, not surprisingly, negative, indicating that subjects across conditions were on average less likely to respond to a question that was rated as very privacy intrusive. The coefficient on the interaction between the treatment and the intrusiveness of the question is of the expected (positive) sign when comparing Condition 1 with Conditions 2 and 4, but it is only mildly significant in the first case. Hypothesis H2 is therefore only supported when comparing a condition of little control with a condition of higher control. When comparing Conditions 1 and 3, instead the interaction is negative and significant: the probability of answering a very intrusive question is lower if the subject has partial control over information, as compared to a condition of zero control. This results is likely due to the very nature of the treatment, which makes publication of very sensitive information more salient, but at the same time does not allow the prohibition of the publication of specific questions. Adding a dummy variable for the provision of an email address, which should have made subjects feel more identifiable, doesn't affect our results, probably due to the high overall willingness to provide an email address. In other words, subjects who provided an email address did not differ significantly from those who didn't in terms of willingness to answer or to publish their information.

Table 6.

	<i>Comparing conditions:</i>			
	<i>1 and 2</i>	<i>1 and 4</i>	<i>1 and 2</i>	<i>1 and 4</i>
	<i>Strategy A</i>		<i>Strategy B</i>	
Treatment	1.55** (.000)	1.51** (.000)	-.25 (.534)	.72** (.004)
Intrusive	-.88** (.000)	-.85** (.000)	-.93** (.000)	-.86** (.000)
Treat_Int	1.31** (.012)	.45 (.191)	.37 (.160)	.39 (.108)
Age	.02 (.720)	-.01 (.891)	-.02 (.788)	-.04 (.360)
Male	.21 (.491)	-.07 (.762)	.51 (.193)	-.16 (.428)
N	59	68	69	68
Prob > χ^2	.000	.000	.000	.000

Random effects Probit coefficients of regression of modified response rate on treatment, with dummy for most intrusive question, interaction and demographics. P-value in brackets.

** indicates significance at 10% level, ** indicates significance at 5% level.*

In order to properly compare Condition 1 with each of the other conditions, though, we have to make sure that subjects answered a question *and* allowed the publication of their answers, because for subjects assigned to Condition 1 answering a question automatically implied the publication of their answer. Since everybody in Conditions 3 and 5 allowed the publication, the results we reported in Table 5 for those conditions pass this test. For Conditions 2 and 4, however, we adopted two independent strategies to analyze the data. First, we restricted our sample of answers to the ones that the subjects allowed to publish (Strategy A; since 10 subjects in Condition 2 did not allow the publication of any answer, they disappeared from the dataset, which explains the lower number of observations relative to the previous regression); second, very conservatively, we considered the lack of publication permission as a lack of answer (Strategy B). The results of the regressions with our modified dataset are reported in Table 6. The results don't change much, due to the high overall willingness to publish information: in particular, we notice that the treatment effect in Condition 2 becomes negative in Strategy B – but it's not significant. This is not surprising, since Strategy B artificially decreases the probability of answering a question. The fact that the treatment remains significant for Condition 4 is particularly reassuring, and is due to the fact that, as observed while discussing qualitative results, most subjects in that condition decided to publish everything, and no subject prohibited the publication of all answers – so there were only few individuals and questions for which the variable *Answer* was actually modified from 1 to 0. Finally, the interaction term in Strategy A is still positive and significant for Condition 2, and positive but not significant for

Condition 4; it is also positive but not significant in both cases for Strategy B. Hence, again, hypothesis H2 is only supported if the treatment is complete control and the default effect is not too strong.

Table 7.

<i>Experimental condition</i>	<i>Control</i>	<i>Privacy concern</i>
1	1.39	5.70
2	3.61 [.000]	3.58 [.000]
3	2.62 [.000]	4.87 [.021]
4	3.57 [.000]	3.54 [.000]
5	2.97 [.000]	3.88 [.000]

Some tests from the manipulation checks in Study 3: sample averages. In square brackets: P-values for t-tests on the difference in means – each condition compared to condition 1.

We concluded our survey with a set of manipulation checks that allowed us to verify whether subjects understood how their information would have been used and how much control they had over the information provided. These final questions were the same across all conditions, except for the differences required by the nature of the treatment (a list of all questions can be found in the Appendix). In Table 7 we report the sample averages across conditions of two indicators of control and privacy concerns. The control question was the following: “Do you think you were given enough control on whether your answers would be published among the results of the study? (By control we refer to whether you felt you could decide what would be published or not)” – and the possible answers ranged from “No control at all” (coded as 1) to “Complete control” (coded as 4), with the “I don’t know” option being coded as 5. The privacy concern question was the following: “How concerned were you about your privacy as you answered the questions in this survey? If you didn’t feel neither one way or the other, please click on the middle choice.” – and the possible answers ranged from “Not at all” (coded as 1) to “Very much” (coded as 7). As we can see, subjects perceived a significantly lower level of control over information publication, and higher level of privacy concern, in Condition 1 as compared to any other condition. This suggests that our manipulations of control were successful, and that the results we obtained from our Probit estimations are most likely due to control (or lack thereof) over information publication, and their impact on privacy concerns.

7. Discussion and Limitations

Our between-subjects experiments provide empirical evidence in support of one possible reason why people tend to reveal private information, especially when using Web 2.0 applications. In three studies, we found that people respond to manipulations of control over information publication, while the control over information access and use by others remained unchanged. We can thus infer that control over publication receives a larger weight in people's decision to reveal. Even though people are likely to be aware that potential privacy threats derive from who *accesses* their information and how that information is *used*, they may neglect to fully consider, or even fail to realize, that control over information access and usage by others is what matters most for privacy protection, while control over information publication is less relevant: our subjects seemed to care more for control over publication of private information than for control over access and use of that information; when someone other than themselves was responsible for the publication, or when the publication itself was uncertain (which reduced the probability of access/use by others) our subjects were more likely to refrain from disclosing. This could be due to the fact that, since the publication of personal information is a certain and immediate event, it is also more salient than the risk of somebody accessing and using that information, an outcome which is uncertain and distant in time. This could also explain the backlash against initiatives that users perceive as detrimental to their privacy (such as Facebook News Feed), when similar or even larger amounts of personal information are willingly revealed by the same users. Arguably, the costs and benefits associated with the mere dissemination of personal information are psychological, while the trade-offs arising from other people's actual usage of our information are more tangible: social value, promotions, discrimination, and so forth. However, it would appear that individuals give more relevance to the former rather than to the latter trade-offs.

The paradoxical policy implication is that Web 2.0 applications, by giving greater freedom and power to reveal and publish personal information, may lower the concerns that people have regarding control over access and usage of that information. On the one hand, changing such granular controls on information publication is time-consuming for the user, and therefore costly; on the other hand, those controls may appear more salient than the control on information access and usage by others, therefore pushing users to publish even more information. Moreover, the strong feeling of control over information publication may increase the satisfaction in posting it, which, when compared to the less salient risk of privacy threats deriving from it, creates a self-control problem for the user (Loewenstein and Haisley, 2008): since those threats are uncertain and distant in time, users may prefer to take the risk but obtain immediate satisfaction, rather than refrain from revealing.

The overall high response rate we found in our studies reflects the general tendency of people to reveal a lot of private information, even though our subjects were given very little detail about the purpose of the study and the actual content of the survey they were about to take. Our results are based on three surveys conducted among students at a North-American University, an educated community that is likely to be familiar with the technology of online social networks and Web 2.0 applications and aware of the implications of joining them. Nonetheless, our subjects showed high willingness to reveal private information. More than lack of awareness, it seems that this is at least in part due to the particular sense of control that new technologies transmit to users, making them feel endowed with the power of managing the flow of information about them that stems from their voluntary willingness to reveal. It is the users' choice to join the network, it is their deliberate and conscious decision to make their profile meticulously detailed, it is again their choice to accept "friendship" or not and make the content of their profiles visible to "friends" or to everybody on the network.

On the one hand, the feeling of power conveyed by detailed controls in the privacy settings of several social media and the saliency of information publication generate confusion between control over publication of private information and control over accessibility and availability/usability of that information by others. On the other hand, the voluntary nature of the disclosure makes people perceive it as non privacy-threatening relative to the situation where disclosure is solicited or required, in which case reactive devaluation⁹ might instill suspicion in people and prevent them from revealing private information. This paper addressed the first issue, but the voluntary versus required nature of information disclosure deserves more direct and accurate analysis and provides the direction for future research.

Apart from the general problems related to survey-based studies, regarding truthful response from participants (Bradburn, 1983), which we try to minimize using a self-administered computerized data collection process rather than conventional interviewer administration (Schaeffer and Presser, 2003; Tourangeau and Smith, 1996), one limitation of this study is external validity: the Facebook was created as a community specifically dedicated to college or graduate students, and our sample could be at most considered representative of that population. But members of the Facebook today are not only university students: they include younger students, professionals, self-employed people who try to advertise their business, parents who want to control their sons' and daughters' activities on the Internet.¹⁰ a great variety of users that are very likely to have different sensitivity and concerns about privacy protection. Arguing that our results from the first two studies are generalizable to all these

⁹ Reactive devaluation is the phenomenon by which the simple fact that an offer or a concession is suggested by somebody other than the self diminishes the apparent value or attractiveness of the offer (Ross and Stilling, 1991).

¹⁰ Indeed, the statistics published by Facebook in July 2009 report that the fastest growing group of users is 55 year-olds and older.

types of users would be obviously inappropriate. In fact, it would be interesting to observe the intercultural differences in people's behavior in online social networks. Our sample was too over-representative of American students in order to allow us to address that question. Nonetheless, our results, tested not only in the context of online social networks (Study 3), may be generalized to similar IT environments such as blogs or forums: even in these cases, people are likely to give different weight to control over publication and control over access/use by others, resulting in revelation of "too much" private information.

References

- Acquisti, A., "Privacy in electronic commerce and the economics of immediate gratification", *Proceedings of the 5th ACM conference on Electronic commerce*, 2004.
- Acquisti, A. and Gross, R., "Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook", *Proceedings of the Sixth Workshop on Privacy Enhancing Technologies*, 2006.
- Acquisti, A. and Gross, R., "Predicting Social Security Numbers from Public Data," *Proceedings of the National Academy of Science*, 106(27), 10975-10980, 2009.
- Acquisti A., Grossklags J., "Losses, gains, and hyperbolic discounting: An experimental approach to information security attitudes and behaviors", *Proceedings of the 2nd Annual Workshop of on Economics and Information Security*, 2003.
- Arulampalam, W., 1998, "A Note on Estimated Coefficients in Random Effects Probit Models", *The Warwick Economics Research Paper Series (TWERPS)*, 520.
- Black, J., 2004. "The perils and promise of online schmoozing." *BusinessWeek Online*, February 20.
- Boyd, D., 2008, "Facebook's Privacy Trainwreck: Exposure, Invasion, and Social Convergence", *Convergence: The International Journal of Research into New Media Technologies*, 14(1):13-20.
- Bradburn, N. M., 1983, "Response Effects", in *Handbook of Survey Research*, ed. P. Rossi, J. Wright and A. Anderson, pp. 289-328 New York: Academic Press.
- Campbell, W. K., Goodie, A.S. and Foster, J.D., 2004, "Narcissism, Confidence, and Risk Attitude", *Journal of Behavioral Decision Making*, 17:297-311.
- Culnan, M. J., 1995, "Consumer Awareness of Name Removal Procedures: Implications for Direct Marketing", *Journal of Direct Marketing*, 9(2):10-19.
- Culnan, M. J ., 1993, "How Did They Get My Name? An Exploratory Investigation of Consumer Attitudes Toward Secondary Information Use," *MIS Quarterly*, 17(3): 341-363.
- Culnan, M. J. and Armstrong, P. K., 1999, "Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation", *Organization Science*, 10(1): 104-115.
- Dinev, T. and Hart, P., 2004, "Internet Privacy Concerns and their Antecedents: Measurement Validity and a Regression Model," *Behavior and Information Technology*, 23(6):413-23.
- Elgesem, D., 1996, "Privacy, Respect for Persons, and Risk", in *Philosophical Perspectives on Computer-Mediated Communication*, C. Ess ed., New York, State University of New York Press.
- Fenton-O'Creevy, M., Nicholson, N., Soane, E. and Paul, W., 2003, "Trading on illusions: Unrealistic perceptions of control and trading performance", *Journal of Occupational and Organizational Psychology*, 76(1):53-68.
- Fried, C., "Privacy", 1984, in *Philosophical Dimensions of Privacy*, F.D. Schoeman ed., New York, Cambridge University Press.
- Goodie, A. S., 2005, "The Role of Perceived Control and Overconfidence in Pathological Gambling", *Journal of Gambling Studies*, 21(4):481-502.

- Henslin, J. M., 1967, "Craps and Magic", *The American Journal of Sociology*, 73(3):316-330.
- Hoadley, C. M., Xu, H., Lee, J. J. and Rosson, M. B., 2010, "Privacy as Information Access and Illusory Control: The Case of the Facebook News Feed Privacy Outcry ", *Electronic Commerce Research and Applications* 9(1):50-60.
- Jiang, X., Hong, J.I. and Landay, J.A., "Approximate Information Flows: Socially-Based Modeling of Privacy in Ubiquitous Computing", *Proceedings of the 4th International Conference on Ubiquitous Computing* (UbiComp 02), Springer, 2002, pp. 176–193.
- Johnson, E.J., Bellman, S. and Lohse, G.L., 2002, "Defaults, Framing and Privacy: Why Opting In-Opting Out", *Marketing Letters*, 13(1):5-15.
- Kahneman, D. and Tversky A., 1979, "Prospect Theory: An Analysis of Decision under Risk", *Econometrica*, 47(2): 263-291.
- Klein, G.A., 1998, "Sources of Power: How People Make Decisions", MIT Press, Cambridge, Massachusetts.
- Langer, E. J., 1975, "The Illusion of Control", *Journal of Personality and Social Psychology*, 32(2):311-328.
- Laufer, R.S. and Wolfe, R., 1977, "Privacy as a Concept and a Social Issue: A Multidimensional Developmental Theory," *Journal of Social Issues*, 33:22-41.
- Lessig, L., 2002, "Privacy as Property", *Social Research*, 69(1).
- Liang, K. and Zeger S., 1986, "Longitudinal Data Analysis Using Generalized Linear Models", *Biometrika* 73:13–22.
- Loewenstein, G. and Haisley, E., 2008, "The Economist as Therapist: Methodological Ramifications of "Light Paternalism", in "The foundations of positive and normative economics: a handbook", Oxford University Press.
- Malhotra, N. K., Kim, S. S. and Agarwal, J., "Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale and a Causal Model", *Information System Research*, 15(4):336-355.
- Malhotra, A., Gosain, S. and Lee Z., 1997, "Push-Pull: The Information Tug of War – A framework for Information Delivery and Acquisition Systems Design", *Proceedings of the Third America's Conference on Information Systems*, Indianapolis.
- Noam, E.M., 1997, "Privacy and Self-Regulation: Markets for Electronic Privacy", in "Privacy and Self-Regulation in the Information Age", US Department of Commerce.
- Norberg, P. A., Horne, D. R. and Horne, D. A., 2007, "The privacy paradox: personal information disclosure intentions versus behaviors", *Journal of Consumer Affairs*, 41(1):100-126.
- Park, C. W., Jun, S. Y. and MacInnis, D.J., 2000, "Choosing What I Want versus Rejecting What I Do Not Want: An Application of Decision Framing to Product Option Choice Decisions", *Journal of Marketing Research*, 37(2):187-202.
- Pew Internet & American Life Project, "Teens, Privacy & Online Social Networks", April 2007, http://www.pewinternet.org/pdfs/PIP_Teens_Privacy_SNS_Report_Final.pdf.
- Posner, R. A., 1978, "Privacy, Secrecy, and Reputation", *Buffalo Law Review*, 28.

- Ross, L. and Stillinger, C., 1991, "Psychological barriers to conflict resolution", *Negotiation Journal*, 7(4):389-404.
- Sanderson, W. C., Rapee, R. M. and Barlow, D. H., 1989, "The Influence of an Illusion of Control on Panic Attacks Induced via Inhalation of 5.5% Carbon Dioxide-Enriched Air", *Archives of General Psychiatry*, 46(2):157-162.
- Schaeffer, N.C. and Presser, S., 2003, "The Science of Asking Questions", *Annual Reviews – Sociology*, 29:65-88.
- Slovic, P., 1975, "Choice between Equally Valued Alternatives", *Journal of Experimental Psychology: Human Perception and Performance*, 1:280-287.
- Smith, H. J., Milberg, S. J. and Burke, S. J., 1996, "Information Privacy: Measuring Individuals' Concerns about Organizational Practices," *MIS Quarterly*, 20(2):167-196.
- Solove, D. J., 2007, "The future of reputation – gossip, rumor, and privacy on the internet", Yale University Press, New Haven & London.
- Spiekerman, S., Grossklags, J. and Berendt, B., 2001, "E-privacy in 2nd Generation E-Commerce: Privacy Preferences versus actual Behavior", *Proceedings of the 3rd ACM Conference on Electronic Commerce*.
- Stigler, G. J., 1980, "An Introduction to Privacy in Economics and Politics", University of Chicago - Center for Study of Economy and State, 10.
- Sweeney, L., 1997, "Weaving Technology and Policy Together to Maintain Confidentiality," *Journal of Law, Medicine & Ethics*, 25, nos. 2&3: 98-110.
- Tavani, H. T. and Moor, J. H., March 2001, "Privacy Protection, Control of Information, and Privacy-Enhancing Technologies", *Computers and Society*.
- Tourangeau, R. and Smith, T. W., 1996, "Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Question Context", *The Public Opinion Quarterly*, 60(2): 275-304.
- Westin, A. R., 1967, "Privacy and Freedom", New York Atheneum.
- Xu, H. "The Effects of Self-Construal and Perceived Control on Privacy Concerns", *Proceedings of 28th Annual International Conference on Information Systems (ICIS 2007)*, Montréal, Canada, 2007.
- Xu, H., Dinev, T., Smith, H.J., and Hart, P., "Examining the formation of individual's information privacy concerns: toward an integrative view", *Proceedings of 29th Annual International Conference on Information Systems*, 2008

APPENDIX

Questions asked in the survey for studies 1 and 2

- Q1: First name, Middle name
- Q2: Last name
- Q3: Gender
- Q4: Date of birth
- Q5: Age (in years)
- Q6: Country of birth
- Q7: Email address
- Q8: Home address
- Q9: Phone number
- Q10: Do you have a Facebook profile?
- Q11: How long have you been in [city name]?
- Q12: On a scale from 1 (not at all) to 10 (very much), how do you like the city overall?
- Q13: How happy are you here?
- Q14: Do you do any sport?
- Q15: If so, which sport do you do?
- Q16: Do you do any sport on campus?
- Q17: How would rate the sport facilities offered on campus?
- Q18: Are you a member of any group/community/fraternity/sorority?
- Q19: If so, which group or groups are you a member of?
- Q20: How many of the people you know in [city name] do you consider close friends?
- Q21: How many of those are students at [university name]?
- Q22: How many are students at other universities in [city name]?
- Q23: Do you enjoy spending your spare time with your friends much more/ with your friends more/ alone more/ alone much more/ with your friends just as much as alone?
- Q24: Is your family in [city name]?
- Q25: How often do you see your family?
- Q26: Are you single or married?
- Q27: Do you have a girlfriend/boyfriend?
- Q28: Have you ever had a sexual relationship with somebody other than your partner without their knowledge or consent?
- Q29: Where do you live? (University housing, Private housing-alone, Private housing-shared)
- Q30: Have you ever had troubles with your roommates?
- Q31: Would you like to move somewhere else?
- Q32: What program are you in? (e.g.: Undergrad Psychology, Grad Math)
- Q33: Which courses are you taking at the moment?
- Q34: Have you ever cheated for homework/projects (e.g. copy, plagiarize) or on an exam?
- Q35: Have you ever seen someone else cheating?
- Q36: If so, did you inform the instructor?
- Q37: How would you rate the quality of the education you are receiving on a scale from 1 (very bad) to 5 (very good)?
- Q38: Do you think it will make you competitive on the job market?
- Q39: How many hours a day do you spend studying?
- Q40: Are you working at the same time?

Study 1

Instructions in condition 1

Please read these instructions carefully before you move on. No question/field is required. If you decide to answer, a profile will be automatically created for you, with no intervention by the researcher, and published on a new [university name] networking website, which will only be accessible by members of the [university name] community, starting from the end of April. The data will not be used in any other way.

Instructions in condition 2

Please read these instructions carefully before you move on. No question/field is required. If you decide to answer, your data will be collected by the researcher, who will create a profile for you and publish it on a new [university name] networking website, which will only be accessible by members of the [university name] community, starting from the end of April. The data will not be used in any other way.

Study 2

Instructions in condition 1

Please read these instructions carefully before you move on. This is not the usual yada-yada. The information you provide will appear on a profile that will be automatically created for you. The profile will be published on a new [university name] networking website, which will only be accessible by members of the [university name] community, starting at the end of this semester. The data will not be used in any other way. NO QUESTION/FIELD REQUIRES AN ANSWER. Did you understand these instructions? If so, click on Next.

Instructions in condition 2

Please read these instructions carefully before you move on. This is not the usual yada-yada. The information you provide will appear on a profile that will be automatically created for you. Half of the profiles created for the participants will be randomly picked to be published on a new [university name] networking website, which will only be accessible by members of the [university name] community, starting at the end of this semester. The data will not be used in any other way. NO QUESTION/FIELD REQUIRES AN ANSWER. Did you understand these instructions? If so, click on Next.

Results of our pre-study to establish the level of privacy intrusiveness of the questions – studies 1 and 2

Rating	Questions
Very intrusive	Q7: Email address Q8: Home address Q9: Phone number Q28: Have you ever had a sexual relationship with somebody other than your partner without their knowledge or consent? Q34: Have you ever cheated for homework/projects (e.g. copy, plagiarize) or on an exam? Q35: Have you ever seen someone else cheating? Q36: If so, did you inform the instructor?
Moderately intrusive	Q4: Date of birth Q5: Age (in years) Q19: If so, which group or groups are you a member of? Q20: How many of the people you know in [city name] do you consider close friends? Q27: Do you have a girlfriend/boyfriend? Q29: Where do you live? (University housing, Private housing-alone, Private housing-shared) Q30: Have you ever had troubles with your roommates? Q31: Would you like to move somewhere else?
Not at all intrusive	Q1: First name, Middle name Q2: Last name Q3: Gender Q6: Country of birth Q10: Do you have a Facebook profile? Q11: How long have you been in [city name]? Q12: On a scale from 1 (not at all) to 10 (very much), how do you like the city overall? Q13: How happy are you here? Q14: Do you do any sport? Q15: If so, which sport do you do? Q16: Do you do any sport on campus? Q17: How would rate the sport facilities offered on campus? Q18: Are you a member of any group/community/fraternity/sorority? Q21: How many of those are students at [university name]? Q22: How many are students at other universities in [city name]? Q23: Do you enjoy spending your spare time with your friends much more/ with your friends more/ alone more/ alone much more/ with your friends just as much as alone? Q24: Is your family in [city name]? Q25: How often do you see your family? Q26: Are you single or married? Q32: What program are you in? (e.g.: Undergrad Psychology, Grad Math) Q33: Which courses are you taking at the moment? Q37: How would you rate the quality of the education you are receiving on a scale from 1 (very bad) to 5 (very good)? Q38: Do you think it will make you competitive on the job market? Q39: How many hours a day do you spend studying? Q40: Are you working at the same time?

Results from probit regressions on individual questions, study 1.

For brevity, we only include questions for which the estimated coefficients are significant. For all other questions the coefficients were not significant.

Table 8.

	Q2	Q7	Q16	Q27	Q34	Q35	Q36	Q38	Q39	Q40
Treatment	-.71*	-.69*	-.81*	-.81*	-.81*	-.81*	-.64**	-.81*	-.68*	-.68*
(p-value)	(.10)	(.06)	(.06)	(.06)	(.06)	(.06)	(.05)	(.06)	(.06)	(.08)
Margfx	-.15*	-.21*	-.18**	-.18**	-.18**	-.18**	-.25**	-.18**	-.21*	-.18*
(p-value)	(.08)	(.05)	(.04)	(.04)	(.04)	(.04)	(.04)	(.04)	(.05)	(.07)
Chi-2 stat	2.7**	3.5*	3.6**	3.6**	3.6**	3.6**	3.8**	3.6**	3.5*	3.02*
(p-value)	(.09)	(.06)	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)	(.06)	(.08)

Probit coefficients and marginal effects of individual regressions of response rate on treatment – first experiment. * indicates significance at 10% level; ** indicates significance at 5% level; chi-squared statistic and corresponding p-value.

Results from probit regressions on individual questions, study 2.

For brevity, we only include questions for which the estimated coefficients are significant. For all other questions the coefficients were not significant.

Table 9.

	Q2	Q7	Q8	Q9	Q20	Q28	Q34	Q35
Treatment	-0.67**	-1.10**	-.75**	-1.22**	-.90**	-.98**	-.67**	-.59**
(p-value)	(.02)	(.00)	(.00)	(.00)	(.04)	(.00)	(.03)	(.03)
Margfx	-.16**	-.33**	-.28**	-.41**	-.09**	-.15**	-.14**	-.16**
(p-value)	(.02)	(.00)	(.00)	(.00)	(.02)	(.00)	(.03)	(.03)
Chi-2 stat	6.4**	12.8**	6.0**	13.5**	4.6**	8.3**	3.0*	13.5**
(p-value)	(.01)	(.00)	(.01)	(.00)	(.03)	(.00)	(.08)	(.00)

Probit coefficients and marginal effects of individual regressions of response rate on treatment – second experiment. * indicates significance at 10% level; ** indicates significance at 5% level; chi-squared statistic and corresponding p-value.

Questions asked in the survey for study 3

Demographics

Q1: Age

Q2: Gender

Q3: Country of birth

Q4: Email address

Questions on ethical behaviors

Q1: Are you married?

Q2: Have you ever been fired by your employer?

Q3: Have you ever stolen anything (e.g.: from a shop, a person)?

Q4: Have you ever used drugs of any kind (e.g.: weed, heroin, crack)?

Q5: Have you ever lied about your age?

Q6: Have you ever had cosmetic surgery?

Q7: Have you ever done any kind of voluntary service?

Q8: Have you ever had sex in a public venue (e.g.: restroom of a club, airplane)?

Q9: Have you ever made a donation to a non-profit organization?

Q10: Do you have any permanent tattoos?

Manipulation checks (Q4 differed across conditions)

Q1: Have you understood how your answers will be used? Please describe.

Q2: In answering the questions in the previous page, were you concerned about the publication of the information provided? Please briefly explain.

Q3: Do you think you were given enough control on whether your answers would be published among the results of the study? (By control we refer to whether you felt you could decide what would be published or not).

Q4

Condition 1: How did you feel about the fact that, for all the questions you actually answered, you could not control their publication? If you didn't feel neither one way or the other, please click on the middle choice. (Annoyed - Pleased, Powerless - Empowered, Frustrated - Calm, Controlled - Autonomous, Embarrassed - At ease)

Condition 2: How did you feel about the fact that, for all the questions you actually answered, you had to check a box to allow their publication? If you didn't feel neither one way or the other, please click on the middle choice. (Annoyed - Pleased, Powerless - Empowered, Frustrated - Calm, Controlled - Autonomous, Embarrassed - At ease)

Condition 3: How did you feel about the fact that, for each question you actually answered, you could not control its individual publication? If you didn't feel neither one way or the other, please click on the middle choice. (Annoyed - Pleased, Powerless - Empowered, Frustrated - Calm, Controlled - Autonomous, Embarrassed - At ease)

Condition 4: How did you feel about the fact that, for all the questions you actually answered, you had to check a box to avoid their publication? If you didn't feel neither one way or the other, please click on the middle choice. (Annoyed - Pleased, Powerless - Empowered, Frustrated - Calm, Controlled - Autonomous, Embarrassed - At ease)

Condition 5: How did you feel about the fact that, for each non demographic questions you actually answered, you could not control its individual publication? If you didn't feel neither one way or the other, please click on the middle choice. (Annoyed - Pleased, Powerless - Empowered, Frustrated - Calm, Controlled - Autonomous, Embarrassed - At ease)

Q5: How concerned were you about your privacy as you answered the questions in this survey?

If you didn't feel neither one way or the other, please click on the middle choice. (Not at all - Very Much) Please briefly explain why you felt that way.

Q6: Do you think that your email address and/or your demographic information will be published among the results of the study? Please briefly explain.

Study 3

Instructions in condition 1

Please read the following instructions carefully, as they are important.

In the following pages, you will be asked a number of questions related to ethical behavior.

The researchers would like to publish the answers you will provide in the following pages on their Research Bulletin, among the results of the study.

Please notice that the answers to the demographic questions that you provided in the previous page will NOT be published.

All answers are voluntary. By answering a question, you agree to give the researchers permission to publish your answer.

Instructions in condition 2

Please read the following instructions carefully, as they are important.

In the following pages, you will be asked a number of questions related to ethical behavior.

The researchers would like to publish the answers you will provide in the following pages on their Research Bulletin, among the results of the study.

Please notice that the answers to the demographic questions that you provided in the previous page will NOT be published.

All answers are voluntary. In order to give the researchers permission to publish your answer to a question, you will be asked to check the corresponding box in the following page.

Instructions in condition 3

Please read the following instructions carefully, as they are important.

In the following pages, you will be asked a number of questions related to ethical behavior.

The researchers would like to publish the answers you will provide in the following pages on their Research Bulletin, among the results of the study.

Please notice that the answers to the demographic questions that you provided in the previous page will NOT be published.

All answers are voluntary. In order to give the researchers permission to publish your answers to the questions, you will be asked to check a box in the following page.

Instructions in condition 4

Please read the following instructions carefully, as they are important.

In the following pages, you will be asked a number of questions related to ethical behavior.

The researchers would like to publish the answers you will provide in the following pages on their Research Bulletin, among the results of the study.

Please notice that the answers to the demographic questions that you provided in the previous page will NOT be published.

All answers are voluntary. In order to prevent the researchers from publishing your answer to a question, you will be asked to check the corresponding box in the following page.

Instructions in condition 5

Please read the following instructions carefully, as they are important.

In the following pages, you will be asked a number of questions related to ethical behavior.

The researchers would like to publish the answers you will provide in the following pages on their Research Bulletin, among the results of the study.

All answers are voluntary. In order to give the researchers permission to publish your answers to the questions, you will be asked to check a box in the following page.

Please notice that the answers to the demographic questions that you provided in the previous page will NOT be published without your explicit agreement: you will be asked permission to publish those answers separately.

Results of our pre-study to establish the level of privacy intrusiveness of the questions on ethical behaviors – Study 3

Rating	Questions
Very intrusive	Q2: Have you ever been fired by your employer? Q3: Have you ever stolen anything (e.g.: from a shop, a person)? Q4: Have you ever used drugs of any kind (e.g.: weed, heroin, crack)? Q6: Have you ever had cosmetic surgery? Q8: Have you ever had sex in a public venue (e.g.: restroom of a club, airplane)?
Moderately intrusive	Q10: Do you have any permanent tattoos?
Not at all intrusive	Q1: Are you married? Q5: Have you ever lied about your age? Q7: Have you ever done any kind of voluntary service? Q9: Have you ever made a donation to a non-profit organization?

Figure 3: Screenshots from survey for study 3

Condition 1

Qualtrics Survey Software - Mozilla Firefox

http://new.qualtrics.com/ControlPanel/PopUp.php?PopType=QuestionPreview&SID=SV_6tjTEIrNbAHFL3m&QID=QID26&WID=_WINDOW_PREVIEW_QID26&Q_Lang=EN

qualtrics.com

Study on Ethical Behavior

IMPORTANT: All answers are voluntary. By answering a question, you agree to give the researchers permission to publish your answer.

	Yes	No
1. Are you married?	☐	☐
2. Have you ever been fired by your employer?	☐	☐
3. Have you ever stolen anything (e.g.: from a shop, a person)?	☐	☐
4. Have you ever used drugs of any kind (e.g.: weed, heroin, crack)?	☐	☐
5. Have you ever lied about your age?	☐	☐
6. Have you ever had cosmetic surgery?	☐	☐
7. Have you ever done any kind of voluntary service?	☐	☐
8. Have you ever had sex in a public venue (e.g.: restroom of a club, airplane)?	☐	☐
9. Have you ever made a donation to a non-profit organization?	☐	☐
10. Do you have any permanent tattoos?	☐	☐

Close

Survey Powered By Qualtrics®

Condition 2

Qualtrics Survey Software - Mozilla Firefox

http://new.qualtrics.com/ControlPanel/PopUp.php?PopType=QuestionPreview&SID=SV_6tjTEIrNbAHFL3m&QID=QID56&WID=_WINDOW_PREVIEW_QID56&Q_lang=EN

qualtrics.com

Study on Ethical Behavior

IMPORTANT: All answers are voluntary. In order to give the researchers permission to publish your answer to a question, please check the corresponding box.

	Publication permission	Yes	No
1. Are you married?	☐	☐	☐
2. Have you ever been fired by your employer?	☐	☐	☐
3. Have you ever stolen anything (e.g.: from a shop, a person)?	☐	☐	☐
4. Have you ever used drugs of any kind (e.g.: weed, heroin, crack)?	☐	☐	☐
5. Have you ever lied about your age?	☐	☐	☐
6. Have you ever had cosmetic surgery?	☐	☐	☐
7. Have you ever done any kind of voluntary service?	☐	☐	☐
8. Have you ever had sex in a public venue (e.g.: restroom of a club, airplane)?	☐	☐	☐
9. Have you ever made a donation to a non-profit organization?	☐	☐	☐
10. Do you have any permanent tatoos?	☐	☐	☐

Close

Survey Powered By Qualtrics®

Condition 3

Qualtrics Survey Software - Mozilla Firefox

http://new.qualtrics.com/ControlPanel/PopUp.php?PopType=QuestionPreview&SID=SV_6tjTEIrNbAHFL3m&QID=QID112&WID=_WINDOW_PREVIEW_QID112&Q_lang=EN

Qualtrics.com

Study on Ethical Behavior

IMPORTANT: All answers are voluntary. In order to give the researchers permission to publish your answers to the questions, please check the box below.

☐ Publication permission

	Yes	No
1. Are you married?	☐	☐
2. Have you ever been fired by your employer?	☐	☐
3. Have you ever stolen anything (e.g.: from a shop, a person)?	☐	☐
4. Have you ever used drugs of any kind (e.g.: weed, heroin, crack)?	☐	☐
5. Have you ever lied about your age?	☐	☐
6. Have you ever had cosmetic surgery?	☐	☐
7. Have you ever done any kind of voluntary service?	☐	☐
8. Have you ever had sex in a public venue (e.g.: restroom of a club, airplane)?	☐	☐
9. Have you ever made a donation to a non-profit organization?	☐	☐
10. Do you have any permanent tatoos?	☐	☐

Close

Condition 4

Qualtrics Survey Software - Mozilla Firefox

http://new.qualtrics.com/ControlPanel/PopUp.php?PopType=QuestionPreview&SID=SV_6tjTEIrNbAHFL3m&QID=QID69&WID=_WINDOW_PREVIEW_QID69&Q_lang=EN

Qualtrics.com

Study on Ethical Behavior

IMPORTANT: All answers are voluntary. In order to prevent the researchers from publishing your answer to a question, please check the corresponding box.

	No Publication permission	Yes	No
1. Are you married?	☐	☐	☐
2. Have you ever been fired by your employer?	☐	☐	☐
3. Have you ever stolen anything (e.g.: from a shop, a person)?	☐	☐	☐
4. Have you ever used drugs of any kind (e.g.: weed, heroin, crack)?	☐	☐	☐
5. Have you ever lied about your age?	☐	☐	☐
6. Have you ever had cosmetic surgery?	☐	☐	☐
7. Have you ever done any kind of voluntary service?	☐	☐	☐
8. Have you ever had sex in a public venue (e.g.: restroom of a club, airplane)?	☐	☐	☐
9. Have you ever made a donation to a non-profit organization?	☐	☐	☐
10. Do you have any permanent tatoos?	☐	☐	☐

Close

Survey Powered By Qualtrics®

Condition 5

Qualtrics Survey Software - Mozilla Firefox

http://new.qualtrics.com/ControlPanel/PopUp.php?PopType=QuestionPreview&SID=SV_6tjTEIrNbAHFL3m&QID=QID141&WID=_WINDOW_PREVIEW_QID141&Q_lang=EN

Qualtrics.com

Study on Ethical Behavior

IMPORTANT: All answers are voluntary. In order to give the researchers permission to publish your answers to the questions on this page, please check the box below.

☐ Publication permission

Besides the answers to the questions you will provide on this page, the researchers would also like to ask you permission to publish your demographic information. Please check the box corresponding to the information that you allow the researchers to publish.

☐ Age

☐ Gender

☐ Country of birth

QUESTIONS ON ETHICAL BEHAVIOR

	Yes	No
1. Are you married?	☐	☐
2. Have you ever been fired by your employer?	☐	☐
3. Have you ever stolen anything (e.g.: from a shop, a person)?	☐	☐
4. Have you ever used drugs of any kind (e.g.: weed, heroin, crack)?	☐	☐
5. Have you ever lied about your age?	☐	☐
6. Have you ever had cosmetic surgery?	☐	☐
7. Have you ever done any kind of voluntary service?	☐	☐
8. Have you ever had sex in a public venue (e.g.: restroom of a club, airplane)?	☐	☐
9. Have you ever made a donation to a non-profit organization?	☐	☐
10. Do you have any permanent tatoos?	☐	☐

Close