Carnegie Mellon

Misplaced Confidences:
Privacy and the Control Paradox

Laura Brandimarte

Alessandro Acquisti

George Loewenstein

Why are we willing to disclose private information to strangers if, at the same time, we claim to be concerned about privacy?

Why are we willling to reveal sensitive information – but get concerned if the same *information was revealed by others?*

Example: Facebook users' outcry against News Feed

We introduce and test the hypothesis that control over publication of private information may influence individuals' privacy concerns and affect their propensity to disclose sensitive information, even when the objective risks associated with such disclosures do not change

The Control Paradox in Privacy Decision Making

- Consider: Distinction between control over *publication* and control over *access to/use of* personal information (Jiang et al., 2002; Noam, 1997)

- However: Because the act of publishing information (immediate and certain) is more <u>*salient*</u> (Slovic, 1975; Klein, 1998) than its access/use by others (future and uncertain), people may give more importance to control over publication of private information, and less importance to control (or lack thereof) over access/use by others

  - Even though arguably threats to privacy derive from access to/use of available information by others!

I have control over publication… therefore my privacy is protected! Let's share this information

I don't have control over publication… therefore my privacy is not protected! Let me keep this information to myself

Hypotheses

- *H1:* Lower perceived control on publication will trigger lower willingness to reveal, even when the objective risks associated with accessibility/usage do not change, or in fact decrease

- *H2:* Higher perceived control on publication will trigger higher willingness to reveal, even when the objective risks associated with accessibility/usage do not change

Our approach

- We test whether people respond to manipulations of control over *publication* of information, even if *accessibility/usage* don't change – or indeed decrease

- If our hypothesis is correct, we should observe higher (lower) willingness to reveal when subjects have more (less) control over publication

- Implication: Giving *more privacy controls to users* over information disclosure and publication may lead to over-confidence, and *paradoxically cause users to disclose more sensitive information*

Three studies

- Study 1: survey-based randomized experiment, manipulating control over mediated or unmediated publication of personal information

- Study 2: survey-based randomized experiment, manipulating control over probability of publication of personal information

- Study 3: survey-based randomized experiment, manipulating level of control over publication of personal information

Study 2

- Design (same as Study 1)

  – Subjects: CMU students recruited on campus, March 2009

  – Completed online survey

  – Justification for the survey: creation of CMU networking website

  – Questions focused on students' life on and off campus
    - Multiple choice, Yes/No, Rating and open-end questions
    - Included quasi-identifiers + privacy intrusive and non-intrusive questions
      – As rated by 31 subjects independently in a pre-study

Study 2

- **Examples of highly intrusive questions**
  - Email address
  - Home address
  - Have you ever cheated for homework/projects/exams (e.g. copy, plagiarize)?
  - Have you ever had a sexual relationship with somebody other than your partner without their knowledge or consent?

- **Examples of moderately intrusive questions**
  - Date of birth
  - Do you have a girlfriend/boyfriend?
  - Have you ever had troubles with your roommates?

- **Examples of non intrusive questions**
  - Do you do any sport on campus?
  - Which courses are you taking at the moment?
  - How would you rate the quality of the education you are receiving?

Study 2

- Manipulation: Profile automatically published vs. profile published with 50% probability

– Control group

*"The information you provide will appear on a profile that will be automatically created for you. **The profile will be published on a new CMU networking website**, which will only be accessible by members of the CMU community, starting at the end of this semester. The data will not be used in any other way. NO QUESTION/FIELD REQUIRES AN ANSWER."*

– Treatment group

*"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 CMU networking website**, which will only be accessible by members of the CMU community, starting at the end of this semester. The data will not be used in any other way. NO QUESTION/FIELD REQUIRES AN ANSWER."*

Study 2

- Hypothesis: Loss of control over publication should decrease willingness to disclose private information even though, statistically, the probability of third parties accessing/using it is halved in the treatment condition

- Confounding factors

Study 2

- Dependent variable
  - Response rate (whether subject answered or not)

- Explanatory variables
  - Treatment
  - Intrusiveness
  - Demographics (age, gender)

Study 2

- Participants: 67 subjects in control condition, 65 subjects in treatment condition

  - 62 females (34 in control group) and 62 males (29 in control group), 8 missing, evenly split

  - Average age = 21.4 in control group; 21.6 in treatment group (difference not significant)
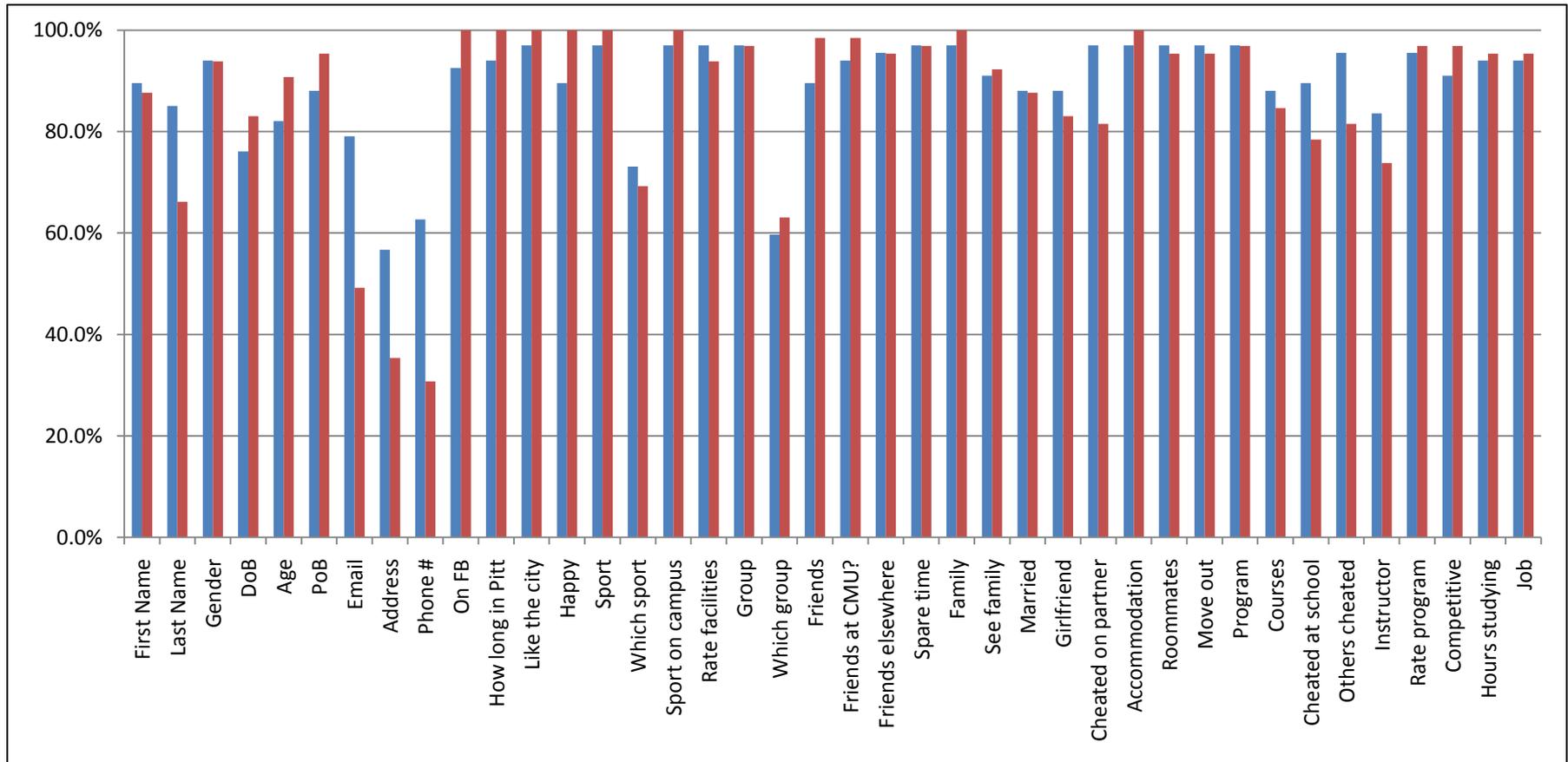
Study 2

Figure 2: Percentage of subjects answering each question in control
and treatment condition

Study 2

Table 2.

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

RE Probit coefficients of panel 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;
*** indicates significance at 1% level

Study 2

- Possible confounding factors

  - Subjects may reveal less because they care less, since the probability of publication is lower
    - That is possible, but then we should observe an effect on those types of questions that required effort (program, courses). No such effect

  - Lack of awareness (Malhotra et al. 2004 and Culnan, 1995)
    - That is possible but consider the case of Harvard senior Lena Chen ("oversharing"), who turned from sex blogger in 2006 to organizer of conferences on virginity in 2010
    - More generally consider Stross, 2009 versus Acquisti & Gross, 2006

Study 3

- Design

  - Subjects: CMU students recruited on campus, March 2010

  - Completed online survey

  - Justification for the survey: study on ethical behaviors

  - Ten Yes/No questions that focused on sensitive behaviors (e.g. drug use, stealing)
    - Included demographics + privacy intrusive and non-intrusive questions
      - As rated by 49 subjects independently in a pre-study

Study 3

Questions asked and their level of intrusiveness

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?

Study 3

- Manipulation: Zero control / Some control / Complete control on publication

    – Condition 1 (zero control)

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

    – Condition 2 (complete control)

    "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."

    – Condition 3 (some control)

    "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."

    – Condition 4 (same as Condition 2, but the default is that answers will 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."

    – Condition 5 (some control + extra demographics)

    "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."

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?	○	○

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?	☐	○	○

Study 3

- Hypothesis: More control over publication should increase willingness to disclose (provide and publish) private information, and especially so for the most sensitive questions

  - It is not the publication of private information per se that disturbs people, but the fact that they don't get to decide what exactly will be published or not

Study 3

- **Dependent variable**
  - Response rate (whether subject answered or not)

- **Explanatory variables**
  - Treatment
  - Intrusiveness
  - Demographics (age, gender)

Study 3

Table 3.

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

RE Probit coefficients of panel 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

Study 3

- The coefficient on *Treatment* is always positive and significant: providing subjects with control over information publication increases their willingness to answer a question (results are similar if we only consider answers that subjects were willing to publish)

- The coefficient on the interaction is only significant when comparing condition 1 with condition 2

- The negative coefficient on the interaction in condition 3 may be due to the very nature of the treatment: makes publication of very sensitive information more salient, but does not allow the prohibition of the publication of specific questions

- Adding a dummy variable for the provision of an email address, which made subjects more easily identifiable, doesn't affect our results

Conclusions

Our results suggest the following:

- Control over publication leads to more revelation of private info

- This effect is stronger for privacy intrusive questions

Conclusions

- People seem to care more for control over publication of private information than for control over access and use of that information, possibly due to salience effects

- As noted, there is a paradoxical policy implication: Web 2.0 applications, by giving greater freedom and power to reveal and publish personal information, may lower concerns regarding control over access and usage of that information, thus leading to overexposure

Thank you!

Questions…

Understanding privacy decision making

- In recent years, growing research interest in understanding motivations and consequences of personal information disclosure and protection

  (Cranor et al., 1999; Moon, 2000; Phelps et al. 2000 Miyazaki & Krishnamurthy, 2002; Earp & Baumer, 2003; Milne et al., 2004)

Benefits and costs of personal information disclosure

- Benefits
  - Signaling (Posner, 1979; Stigler, 1980; Berman & Bruckman, 2001), self-representation (Donath and Boyd, 2005), socialization (Boyd, 2007), …

- Costs
  - Stalking (and cyber-stalking), identity theft (Gellman 2001), price discrimination (Odlyzko, 2003), lost employment opportunities,…
  - The case of Harvard undergrad student, Lena Chen

Concurrent explanations for privacy behavior and privacy dichotomy

- Rational cost-benefit analysis (Posner, 1978; Stigler, 1980)
- Awareness (Samuelson, 2001)
- Trust (Culnan & Armstrong, 1999)
- "Behavioral" explanations: prospect theory (Kahneman & Tversky, 1979), optimism bias, hyperbolic time discounting and immediate gratification (Acquisti, 2004)

Study 2

- Treatment has hypothesized effect on
  - 6 of the questions that were rated as most privacy intrusive (email, address, phone number, cheating on partner, cheating at school, others cheated)
  - 1 moderately intrusive question (friends)

- Treatment did not push subjects to admit more
  - The percentage of subjects answering "No" to questions about sensitive behaviors didn't change significantly (10% level) between the control and the treatment conditions

Study 3

Some descriptive statistics and qualitative results

Experimental condition	Number of subjects	Average age	% Male	Average response rate (%)	Subjects providing email (%)	Subjects answering all questions	Subjects publishing all questions	Subjects publishing no question
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%)	