

Simplification of Privacy Disclosures: An Experimental Test

Omri Ben-Shahar and Adam Chilton

ABSTRACT

Simplification of disclosures is widely regarded as an important goal and is increasingly mandated in a variety of areas. In the area of data privacy, lawmakers and interest groups developed best-practices techniques to help consumers understand how firms collect and use personal information. Commentators have even advocated going a step further and using simpler disclosures—warning boxes that alert consumers to the least-expected elements. But do these techniques succeed in better informing consumers or preventing unwise behavior? To answer this question, we engaged a leading market research firm to conduct a survey on risky sexual behaviors while randomizing the format of the privacy disclosures provided to the respondents. We find that best-practice simplification techniques have little or no effect on respondents' comprehension of the disclosure, willingness to share personal information, and expectations about their rights. Our results challenge the wisdom of focusing regulatory effort on simplifying disclosures.

1. INTRODUCTION

Mandated disclosure is the most commonly used regulatory device in privacy protection. Disclosure seems like a sensible tool because transparency provides a simple and proportional solution to the underlying problem. If people unwisely surrender much of their personal information because they are unaware that it will be broadly used and widely

OMRI BEN-SHAHAR is the Leo Herzel Professor of Law at the University of Chicago. ADAM CHILTON is an Assistant Professor of Law at the University of Chicago. This paper was prepared for the conference Contracting over Privacy held at the University of Chicago Law School on October 16–17, 2015. We thank Richard McAdams, Emily Buss, Daniel Hemel, William Hubbard, Matthew Kugler, Lior Strahilevitz, and workshop participants at the University of Chicago for helpful comments and Marcos Garcia Dominguez, Lisa Fan, and Patrick Maxwell for helpful research assistance. Financial support for this project was provided by the Coase-Sandor Institute of Law and Economics at the University of Chicago Law School.

[*Journal of Legal Studies*, vol. 45 (June 2016)]

© 2016 by The University of Chicago. All rights reserved. 0047-2530/2016/4502-0018\$10.00

shared, then the solution is to make them aware by requiring firms to disclose the information. With the knowledge that disclosures afford them, people can then make wiser information-sharing choices. On the basis of this theory, laws regulating financial and medical services, for example, require privacy disclosures to be prominent and to provide specific information that is deemed relevant to consumers' decisions (see, for example, Gramm-Leach-Bliley Act, 12 C.F.R. 1016.5; Health Insurance Portability and Accountability Act, sec. 164.520).

But despite the allure and wide embrace of transparency as a guiding principle, it has proven hard to accomplish successfully. Firms, of course, usually comply with the letter of shine-the-light laws and post privacy notices with magnificent detail. But consumers remain largely ignorant of the specific details of the notices and continue to divulge loads of personal information. And even if consumers wanted to inform themselves, they could not feasibly make informed data-sharing choices. According to one estimate, the average person encounters so many privacy disclosures every year that it would take 76 days to read them, and the lost time would cost the economy \$781 billion (McDonald and Cranor 2008).

Privacy disclosures' failure to meaningfully inform consumers poses a challenge: could the formats be reengineered to be more useful? If consumers care about privacy but do not read the disclosures because they are poorly drafted and overloaded, the solution seems inevitable: simplification. If a disclosure is too long, shorten it. If it is too technical, make it more user friendly. If it is poorly presented, improve the formatting.

Although it may sound obvious that complex privacy disclosures should be simplified, exactly how this should be done is not obvious. Privacy disclosures are complex because firms collect many types of personal data and use it in numerous ways. Should people be given less than full information to help them focus better attention on the most essential ingredients? Or should people be given all the information but in a standardized format? In the search for pragmatic ways to present privacy practices more effectively, several simplification strategies have risen to the fore.

One simplification strategy focuses on improving the formatting and organization of disclosures. In the privacy area, it is part of a protocol known as best practices. Proposed by lawmakers, advocacy groups, and privacy experts, best practices are an informal "code of conduct" focusing on "enhanced transparency" to help consumers to meaningfully compare and choose on the basis of, among other things, privacy consider-

ations (National Telecommunications and Information Administration 2013, p. 1). Although some of the best practices address the substance of privacy practices (for example, imploring firms not to collect unnecessary information), the bulk of the best practices deal with formal properties of disclosures (for example, telling firms how to present information clearly and succinctly).

A second simplification strategy is even more ambitious. It looks to the perceived (although debatable) success of disclosure tools like the nutrition data boxes on food packaging as a model for effective disclosure. The gist of this solution is to aggressively trim the disclosure to only a few essential facts—those least expected by consumers—and present them in a standardized, nontechnical, and easily comparable format (Kelley et al. 2009; Ayres and Schwartz 2014).

Academics and policy makers have suggested that these strategies would improve consumers' comprehension and behavior. For example, a study solicited by the government found, largely on the basis of interviews, that "good design techniques, combined with the simplified content, helped consumers better understand the information" (Kleimann Communication Group 2006, p. xi). Similarly, Ayres and Schwartz (2014, p. 605) believe that the warning-box method "might efficiently correct the most serious forms of consumer optimism." But despite these claims about the benefits of simplification strategies, it is not known whether they actually change behavior.

In this paper, we present the results of a survey experiment designed to help answer this question. The experiment focuses on privacy disclosures, but its design is generic enough to at least suggest a more general lesson about the simplification paradigm. We engaged a leading market research firm to field the experiment and told the respondents (deceptively) that they were participating in a survey on risky sexual practices that would be used to help develop a commercial mobile application that improves the utility of existing dating services. During the survey, the respondents were asked a host of questions about their sexual behavior and to provide several key pieces of identifying information—topics that are both known to elicit privacy concerns (Alter and Oppenheimer 2009).

The experimental treatment was the privacy disclosure that the respondents received at the beginning of the survey. Before the respondents were asked any questions, they were presented with a data-privacy disclosure that listed all the ways that their personal information would be collected and shared. The disclosures were intended to arouse discomfort

among the respondents, and they covered exactly the same topics and adopted exactly the same substantive policies. For example, all of the disclosures said that we share the information about their sexual behavior with commercial parties and that we do not monitor how those parties secure the data. Our experiment was thus not about the substance of the privacy disclosures but instead about the impact of the disclosures' formal properties. Specifically, we randomly presented the respondents with disclosures that employed different sets of the best practices that are most commonly found in guides for developing model disclosures.

After exposing respondents to one of these disclosures, our survey proceeded to ask questions designed to test whether the formal properties of the disclosure influenced behavior in three ways. First, we measured respondents' comprehension of the disclosure by counting the amount of time they spent on the disclosure screen and by asking a series of comprehension questions. Second, we measured respondents' willingness to disclose sensitive personal information by asking questions about risky sexual practices and by requesting identifying information. Third, we measured respondents' expectations about their privacy rights by asking whether they believed they had legal recourses if we violated our privacy policy and whether they were satisfied with the steps taken to protect their privacy. In each of these three tests our results were consistent: altering the formal properties of the privacy disclosures had essentially no effect on respondents' comprehension of our disclosure, willingness to disclose information, or expectations about their privacy rights.

In addition, we conducted a test to examine the effect of a more radical simplification technique proposed by Ayres and Schwartz (2014) and others. To do so, we conducted a separate survey that asked respondents to rank the gravity of our privacy practices. We then used a treatment that was simply a warning label that specified only the five worst, or least expected, privacy practices. Here, too, we found no meaningful effects. Perhaps struck by the novel format of a warning label, respondents spent more time viewing it (although hardly enough to digest it), but they then proceeded to behave similarly to the respondents who were presented with the other disclosures.

Our findings contrast sharply with the more optimistic tone suggested by prior experimental exercises. For example, Kelley et al. (2010, p. 1577) presented people with simplified and full-text versions of privacy disclosures and concluded that standardized short table formats "significantly outperformed" the full-text formats. Similar findings are documented in testing of simplified mortgage forms (see, for example, Kleimann Com-

munication Group 2012). But these results should be taken with caution. They show that if people are sufficiently focused on the cognitive task of reading and learning about a specific issue, going back and forth to re-read the text, their understanding increases by good presentation of the materials. By contrast, our findings suggest that when people are engaged in a real-world task that focuses their attention elsewhere, the incidental presentation of simplified disclosures does not affect their behavior.

Before proceeding, it is important to note a procedure we followed to enhance the credibility of our study. Both of us entered this project skeptical about the value of disclosure, even when simplified. We were concerned that if the results would turn out consistent with our prior beliefs, as indeed they did, our readers would view them with (justified) suspicion. After all, how often do scholars publish empirical results that conflict with their own previously published predictions and conjectures? Thus, as a form of precommitment, we circulated a draft with the design of our experiment prior to conducting the survey and thus before we knew the results. Subsequently, on the basis of comments by participants in a conference to whom we presented the results, we ran another round of the experiment, tweaked to further induce subjects to act with caution and to notice the disclosures. The results of both rounds of the experiment were very similar, and for brevity we present here only the final round.¹

Our paper proceeds as follows. We begin by providing a brief survey of disclosure-simplification techniques and then explain the survey experiment we designed to test whether these techniques influence behavior. We then present the primary results of our experiment. After doing so, we present the results testing a warning-label-style privacy disclosure. Finally, we conclude by briefly discussing the implications of our results for future research and policy.

2. RESEARCH DESIGN

2.1. Best Practices

There is a widespread consensus among privacy experts, lawmakers, and advocates that an important component of a firm's data-privacy practices is providing a clear disclosure of its policies. This consensus is reflected in the Consumer Privacy Bill of Rights released by the White House (2012,

1. The results of our pilot survey are presented in the online appendix. The unpublished first round of the experiment is available in Ben-Shahar and Chilton (2015).

p. 1) that declares, “Consumers have a right to easily understandable and accessible information about privacy and security practices.” The Federal Trade Commission (FTC 2012, p. 60) has deemed accessible transparency to be a “baseline principle” and instructed firms “to present choices to consumers in a prominent, relevant, and easily accessible place . . . and undertake consumer education efforts to improve consumers’ understanding of how companies collect, use, and share their data.” In its widely accepted Fair Information Practice Principles, the FTC requires privacy disclosures to be “clear and conspicuous,” which, together with the requirement that “a disclosure be readily understandable, likely will . . . communicate effectively the information needed by consumers to make an informed choice about the privacy of their information, including whether to transact business with a financial institution.”²

To provide guidance on exactly how to achieve these goals, a number of best-practice protocols have been developed (see, for example, California Office of Privacy Protection 2008, p. 5; National Telecommunications and Information Administration 2013). Some of these protocols are mandated by law, others are voluntary codes of conduct drafted by lawmakers, and some are compiled by private groups. We reviewed the recommendations of several of these guides, and while there is no universal list of best practices, we found six recommendations to be the most common.

Titles. Use clear titles and headers for the specific provisions.

Layered Information. Provide a short-form summary for each provision, followed by the more comprehensive information. The long form should appear in smaller font and may even be posted elsewhere, but in such cases a clear reference or link to it must accompany the short-form summary.

Font. Use easily readable type in a legible size and in a distinct color that contrasts distinctly with the background.

Literary Style. Use active, not passive, language and short sentences with plain, straightforward language.

Examples. When listing categories of personal information that is being collected or shared, give concrete examples, rather than ambiguous statements, of the type of information in each category.

Names. If the notice refers to partner and affiliated companies, provide their names.

2. See Federal Trade Commission, Privacy of Consumer Financial Information; Final Rule, 16 C.F.R. pt. 313.3 (May 24, 2000), for a definition of “clear and conspicuous.”

Other best-practice recommendations are also often made, but they do not directly deal with the presentation format. We thus decided to focus on these six common recommendations when designing our experimental test of the effectiveness of the formal properties of disclosure.

2.2. Experimental Design

To test the impact of the formal properties of privacy disclosures on respondents' behavior, we needed to set the experiment in a context that met three criteria. First, it had to involve questions of a sensitive nature for which respondents may be apprehensive about sharing personal information and for which we could build on rich prior research in wording sensitive questions. Second, it had to allow us to present the subjects with a plausible explanation why a corporation would be interested in their personal information. Third, it had to be a topic for which we could credibly tell respondents that the stakes of the research were nontrivial and, as a consequence, that failure to respond truthfully would be harmful.

Ultimately, we elected to frame our survey as consumer research on risky sexual behavior. We told the respondents that our survey was a study being conducted by a for-profit company that is examining risky sexual behavior to improve matches and reduce sexual harassment in online dating services. This topic satisfies the three criteria: sexual practices are a sensitive topic with rich prior research on survey design (for example, Kays, Gathercoal, and Burhow 2012; John, Acquisti, and Loewenstein 2011), it is reasonable to think that a company developing dating applications would be interested in the sexual practices of respondents, and reducing sexual harassment is a worthy goal.

The survey started by providing all respondents with the same prompt telling them the purpose of the research. On the first screen respondents were told, "This survey is being conducted by a for-profit company in Chicago developing a new commercial application (App) for smartphones that provides users novel search tools across all available listings in dating sites. The results of this survey will be used to design tools that improve relevance of match results and reduce sexual harassment. It is important that you answer all questions honestly." After reading this prompt, the respondents were asked to click "next" to start the survey.

We included this initial prompt for both substantive and practical reasons. The substantive reason is to provide a framing for our survey that would accomplish the goals discussed above while also distracting respondents from the fact that we are studying the formal properties of

privacy disclosures. The practical reason is to allow us to measure the time respondents spent reading the privacy policy disclosures that we presented on the following screen because we were able to use when respondents finished the initial screen as a start time.

After the initial prompt, respondents were directed to a screen presenting a privacy disclosure. This was the key experimental treatment. Respondents were either randomly presented with one of five different privacy disclosures or, as a quasi-control condition, a blank screen. The privacy disclosures were designed to be realistic and were based on dozens of actual disclosures used by major businesses and on multiple best-practice guides. The privacy disclosures were also designed to share four attributes.

First, all of the privacy disclosures began with a prominent title: “Your Privacy Rights.” This is the primary way users are alerted to the subject matter of privacy policies on websites and mobile apps.

Second, all of the privacy disclosures began by informing the respondents of the purpose of the disclosure and what it includes, in the following manner: “We value your privacy. In this page, we explain our data privacy practices. We explain how we collect, use, disclose, and store the information about you that you reveal in the survey. These disclosures of your data and personal information may be done without requesting additional consent from you. Additional disclosures of your information to government agencies will be made to the extent permitted or required by law. Continuing past this page means that you grant us permission to engage in these practices, including the permission to share the information with others as explained below.”³

Third, all of the privacy disclosures covered the same four standard topics. We detailed them under the following headings: “The Information We Collect,” “How We Use the Information,” “Disclosure of Data to Third Parties,” and “Protection of Personal Information.”

Fourth, all of the privacy disclosures adopted the same substantive policy for each of the four topics. These policies were designed to mirror common privacy practices of commercial firms, while still pushing the envelope. For example, respondents were told that the firm would share information with commercial health insurance companies, that it would link their information to other public and private data sources, and that

3. The wording was changed slightly for the disclosures that do not use a clear literary style.

it would retain their information indefinitely. Table 1 reports the substantive policies that the privacy disclosures adopted.

In short, each of the privacy disclosures had the same title, presented the same introduction, covered the same topics, and adopted the same substantive policies. This allowed us to vary the treatment only in style, not in substance.

As we previously noted, our preliminary research revealed that there are six best practices that are most commonly advocated: titles, layered information, clear font, easy-to-understand literary style, examples, and specific names. To test their effects, we developed five versions of a privacy disclosure for our treatment conditions, in addition to a sixth, quasi-control, blank treatment:⁴

Best-Practice Treatment. All six best practices were used.

Organization Treatment. The best-practice treatment was altered by removing the two best practices that are related to document organization: titles and layered information.

Presentation Treatment. The best-practice treatment was altered by removing the two best practices related to ease of reading: clear font and easy-to-understand literary style.

Specification Treatment. The best-practice treatment was altered by removing the two best practices related to the concreteness of the information: using examples and providing specific names for other entities and organizations.

Worst-Practice Treatment. None of the six best practices was followed.

Blank Treatment. This version simply displayed a blank page on the screen that contained a privacy disclosure for the other five treatment groups.

After being presented with one of these six treatments, the respondents were asked to click “next” at the bottom of the disclosure page to continue the survey. The remainder of our survey was designed to test the effects of the privacy disclosures on the behavior of the respondents. We discuss how our survey did so as we discuss our results.⁵

2.3. The Use of Deception

It is important to acknowledge that our experiment directly deceived the respondents. Our decision to use deception, however, is hardly unique.

4. For the exact wording of the privacy disclosures, see part 1 of the online appendix.

5. For the exact wording of the survey questions, see part 2 of the online appendix.

Table 1. Substantive Policies in the Privacy Disclosures

Topic	Policies
The information we collect	We collect information on your sexual practices based on your responses to questions in this survey We collect personal information that can identify you through this online interaction with your computer
How we use the information	We collect and link additional information about you from other public and private data sources We use the information for commercial development of a dating application for smartphones We use the information to develop advertising and marketing tools within the application
Disclosure of data to third parties	We use the information to follow up with you on future communications We share the information with commercial partners to improve the commercial utility of the application We share the data with advertisers and ad-placement companies that help us design the profit opportunities from the mobile application We share the data with commercial health insurance companies that partner with us in developing the data-driven application
Protection of personal information	We retain the information indefinitely We save your personal information on a cloud server with limited access We encourage our partners to take high-security measures in storing your data, but we are unable to verify their security practices

For example, one study found that in the years from 1986 to 1997, between 31 percent and 47 percent of papers in a top social psychology journal used deception in their research design (Hertwig and Ortmann 2008). The pervasiveness of deception in experimental research is based on the belief that it is acceptable when it is essential to the research design and the risks are minimal (see generally Morton and Williams 2010, pp. 500–521).

Distracting respondents from the fact that we were studying privacy disclosures was the only way to ensure that they would react to those disclosures in a normal way. We took three steps to minimize the risks to respondents. First, the sensitive questions they were asked were based on prior survey research on privacy and thus within accepted research standards. Second, all respondents were provided with a debriefing statement after the conclusion of the experiment that informed them about the true purpose of our research. Third, we followed a data security plan to minimize the risk that any sensitive information would be stored or compromised.

3. PRIMARY RESULTS

The experiment was administered online to a nationally representative sample recruited by Survey Sampling International (SSI). A leading market research firm, SSI primarily conducts surveys for corporate clients. The recruited sample had 1,484 respondents that were representative of the US adult population on the basis of gender, age, ethnicity, and census region.⁶ After the respondents were randomly presented with one of the treatments, they were presented with questions designed to measure how the disclosure influenced their comprehension of the privacy policy, willingness to share personal information, and expectations about their privacy rights.

3.1. Comprehension of the Disclosure

Respondents' comprehension of the disclosures was tested in two ways. First, we measured the amount of time that respondents spent on the screen displaying the privacy disclosures. Second, we asked the respon-

6. For information on the demographic breakdown of the sample, the number of respondents who received each treatment, and the demographic balance across treatment groups, see part 3 of the online appendix.

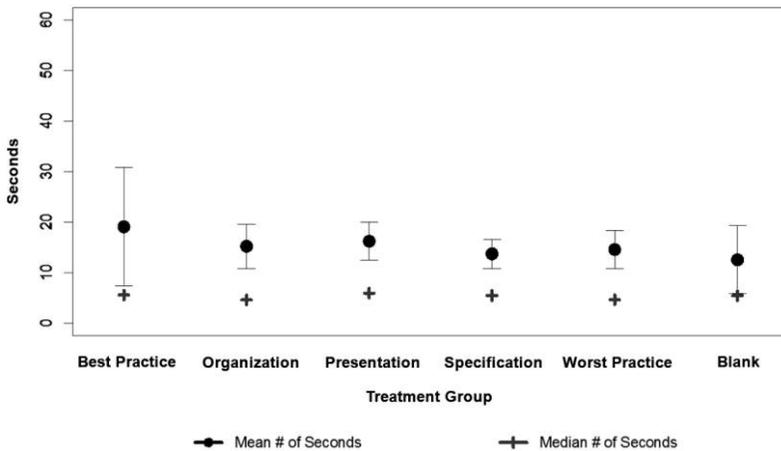


Figure 1. Time spent viewing the disclosure screen

dents questions to measure their comprehension of the privacy disclosures.

3.1.1. Time Viewing the Disclosure Screen. Best practices intend to make disclosures easier for consumers to understand. But how should such ease affect the time spent reading disclosures? People may spend less time reading clearer disclosures because they are easier to follow, more time because they are less daunting and more accessible, or the same time because they always click through disclosure screens as fast as possible. To test which of these effects the formal properties of privacy disclosures have, we measured the number of seconds between when respondents clicked “next” on the initial screen to move to the disclosure screen and when respondents clicked “next” to leave the disclosure screen.

Figure 1 reports the mean number (and 90 percent confidence intervals) of seconds that the respondents in each treatment group spent on the page with the privacy disclosure. Figure 1 shows that the mean number of seconds spent on the disclosure screen ranged from 19.12 for the best-practice group to 12.59 for the respondents who were presented with the blank screen (the blank group). Even though the difference in mean seconds between these two groups is the largest discrepancy between our six treatments, it is not statistically significant at the .1 level ($p = .43$). Since means can be influenced by the presence of outliers in the data, Figure 1 also reports the median number of seconds that respondents in each treatment group spent on the disclosure screen. The median values are almost indistinguishable: all fall between 4.5 and 6.0 seconds.

To put these results in perspective, our privacy disclosures ranged from 499 to 970 words long, and research suggests that the average college-educated adult in the United States can read roughly 300 words per minute (see, for example, Noah 2000). This suggests that, even ignoring the fact that privacy policies like ours have unfamiliar terminology (Jensen and Potts 2004), it should still take over a minute and a half to read even our shortest disclosure. Not only was the median amount of time our respondents spent on the disclosure screen 5.47 seconds (just enough time to aim at and click the “next” button), but just 2.4 percent of the respondents (36 of 1,484) spent a minute and a half on the disclosure screen. In other words, regardless of the formal properties of the disclosure, nearly all respondents clicked through without taking the time to read it.

Of course, it is possible that the respondents who did take the time to read the disclosure responded to our survey differently after doing so. In unreported results, we analyzed the responses for the respondents who spent at least 30, 60, or 90 seconds on the disclosure screen (hereafter, the “readers”).⁷ These readers were more likely to correctly answer the comprehension questions discussed in Section 3.1.2, but these respondents also shared roughly the same amount of personal information and largely had the same attitudes as the other respondents to our survey.

3.1.2. Correct Answers to Comprehension Questions. In addition to measuring the time respondents spent on our privacy disclosures, we also directly asked them questions designed to assess their understanding of the disclosures. At the end of our survey—after all of the other questions discussed in the subsequent sections of this paper—we asked respondents five questions about the contents of the privacy policies.

We asked respondents what our policies are for the information we collect, who we share the data with, how long we retain the data, the security measures our partners have to take, and how we respond to data breaches.⁸ They were presented with four possible answers to each question, and one of the options for each question was that we did not mention a policy on the topic. We then counted how many of these five questions the respondents correctly answered.

Figure 2 reports the mean number (and 90 percent confidence in-

7. For the analysis, see part 5.1 of the online appendix.

8. These five questions were presented in random order. Although the first four questions were all addressed in our privacy disclosures, the fifth question was not. We included this question, however, because it is a commonly covered topic in privacy disclosures, and asking respondents for our policies regarding it provided an additional test of whether our disclosure was actually read.

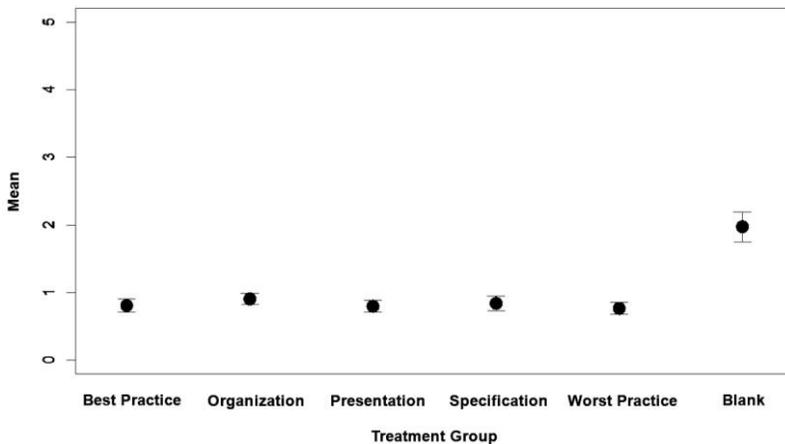


Figure 2. Respondents with correct answers to comprehension questions

tervals) of questions that the respondents correctly answered about the contents of our privacy disclosures. Most notably, regardless of the format of the disclosure, respondents in all five disclosure treatments answered roughly one of the five questions correctly (mean = .82). The best-practice group averaged .81 questions, and the worst-practice group correctly answered .77—a difference that is far from statistically significant ($p = .60$). Put another way, with the exception of the blank group, the average respondent correctly answered fewer questions than would be expected by random chance. The respondents in the sixth treatment—the blank group—answered roughly two of the five questions correctly (mean = 1.97), but since the blank group did not receive a disclosure, the correct answers for them (“We did not mention a policy on the topic”) were different than they were for the other treatment groups.

Among the readers, only a very modest improvement was observed.⁹ In that group, about half of the improvement was obtained for all treatments, and the majority of that effect was obtained by spending 30 seconds—hardly enough to read the text. This suggests that the (small) improved comprehension is likely due to a selection effect, whereby the respondents who took longer to read the disclosures are more sophisticated or experienced, rather than to better delivery of the notice. In

9. People who spent at least 90 (60, 30) seconds on the disclosure screen answered correctly 1.25 (1.48, 1.56) questions on average across all treatments. Their scores improved to 1.75 (2.33, 1.71) under the best-practice treatment. See part 5.1 of the online appendix.

all, these results suggest that for the great majority of people, the formal properties recommended for best-practice disclosures do not improve comprehension of their content.

3.2. Sharing Personal Information

Disclosures are aimed at changing behavior. In the privacy context, the behavior in question is the sharing of personal information. Thus, the ultimate measure of success for best-practice disclosures is whether they change respondents' willingness to share personal information. We tested this in two ways. First, respondents were asked a series of sensitive personal questions about their risky sexual behavior. Second, they were asked to provide personal identifying information.

3.2.1. Risky Behaviors Disclosed. We asked the respondents a series of questions about their risky sexual behavior that were either taken from or based on questions used in other research on this topic (for example, John, Acquisti, and Loewenstein 2011). We identified 10 questions on the subject that could plausibly be related to efforts to prevent the spread of sexually transmitted diseases (although, admittedly, some questions are more relevant to the spread of sexually transmitted diseases than others). The respondents were asked to answer each of these questions with either yes or no.¹⁰

1. Have you ever had sex with someone you met the same day?
2. Have you ever had sex with someone you met the same day without using a condom?
3. Have you ever cheated on your partner?
4. Have you ever had anal sex?
5. Have you ever had sexual thoughts about a member of your same sex?
6. Have you ever used sex toys?
7. Have you ever taken nude pictures of yourself or of a sexual partner?
8. Have you ever been diagnosed with a sexually transmitted disease?
9. Have you ever lied about how recently you were tested for sexually transmitted diseases?

10. The questions were presented in random order, and the respondents viewed one question per screen. If they attempted to move to the next screen without answering, they were asked, but not required, to provide a response. Respondents who either answered no or advanced without answering were coded the same for our analysis.

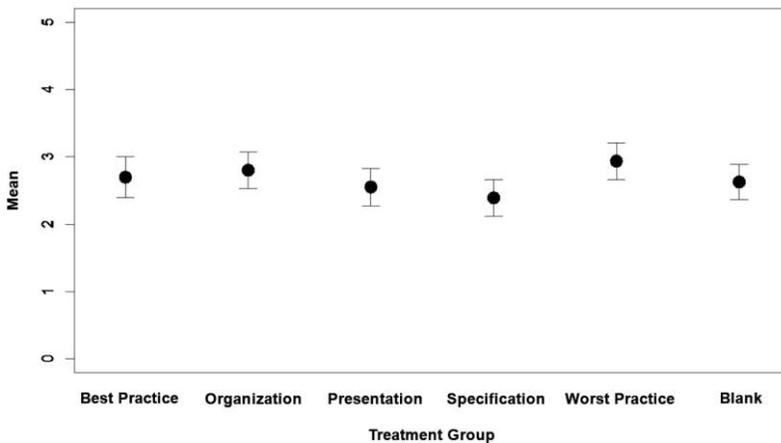


Figure 3. Risky behaviors disclosed

10. Have you ever neglected to tell a partner about a sexually transmitted disease from which you were suffering?

Figure 3 presents the mean number (and 90 percent confidence intervals) of yes responses to these 10 questions.¹¹ Respondents in all six treatment groups answered that on average they had engaged in between two and three risky behaviors. The mean number of risky behaviors ranges from 2.39 for the specification group to 2.94 for the worst-practice group. None of the treatment groups, however, reported a number of risky behaviors that was different from the blank group at the .1 level of statistical significance. And among the small subset of readers, no measurable differences were observed.¹² These results suggest that the formal properties recommended for best-practice disclosures do not increase willingness to divulge sensitive information.¹³

11. Part 5.2 of the online appendix breaks down the tendency to share each of the 10 risky sexual behavior questions by treatment group. Some behaviors are more often shared, but it does not appear that there are systematic differences across treatment groups.

12. Readers who spent at least 90 seconds disclosed on average a slightly higher number of risky behaviors (3.50), but there are too few respondents in this group to test which disclosure treatments drive this (statistically insignificant) increase. See part 5.1 of the online appendix.

13. Since it is possible that privacy-sensitive subjects would react to the disclosures by terminating their participation in the survey, we also examined how many respondents dropped out between being shown the survey and the end of our questions on risky sexual behavior. The average dropout rate for the six treatments was 3.5 percent, and there were

3.2.2. Identifying Information Provided. After the questions about risky sexual behavior, the respondents were asked a series of standard demographic questions.¹⁴ After that, they were further asked to provide us with five pieces of personal identifying information: the county they took the survey in, the zip code where they took the survey, their phone number, their email address, and their mailing address. For each question, the respondents had to choose between providing the information by filling in a blank answer space or clicking “I prefer not to say.” We then recorded the number of times that the respondents provided an answer. In general, we could not verify whether the responses were accurate (although we were able to conduct some verification of the zip codes provided).¹⁵

Figure 4 reports the mean number (and 90 percent confidence intervals) of times that respondents filled in an answer in the space provided.¹⁶ The mean number of questions that respondents provided an answer for was roughly 2.5 for all six of the treatment groups, ranging from 2.41 for the organization group to 2.58 for the specification group. Once again, none of the treatment groups had average responses that were different from the blank group at the .1 level of statistical significance. These results suggest that the formal properties recommended for best-practice disclosures do not increase willingness to share identifying information.

3.3. Expectations of Privacy Rights

It is possible that even if they are not read or used, the presence of disclosures affects people’s perceptions about the disclosers. For example, in the privacy context, the presence of a privacy notice and its format may alter people’s perceptions about their legal rights or about how seriously their privacy is being taken (Pan and Zinkhan 2006; Proctor, Ali, and Vu

not any statistically significant differences between the six treatment groups. Using best practices, in other words, did not translate into a higher rate of cautionary dropping out.

14. These included age, gender, ethnicity, and the region of the country where they lived. For the responses to these questions, see part 3 of the online appendix.

15. We verified whether the zip-code responses matched the zip codes collected by the survey software. Although there are a large number of discrepancies, there are not any noticeable trends across treatment groups. Further, the great majority of people (84 percent) indistinguishably across all treatments provided zip codes that were in the vicinity of the location in which they took the survey. See part 5.3 of the online appendix.

16. Part 5.2 of the online appendix provides the mean number of yes responses to each of the five requests for identifying information by treatment group. Disaggregating the data reveals that there are differences in the responses by question. For example, roughly 84 percent of respondents provided a zip code, but only 26 percent of respondents provided a mailing address. Once again, it does not appear that there are systematic differences across treatment groups.

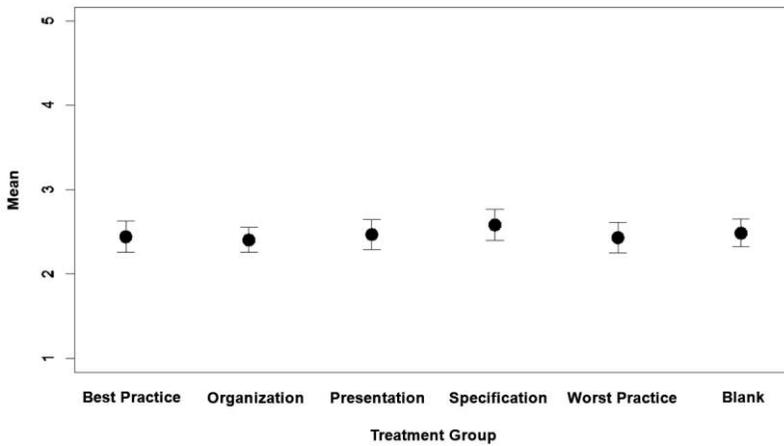


Figure 4. Respondents providing identifying information

2008). To gain insight into this, we asked the respondents several questions on their expectations about their legal rights and reported satisfaction with our survey.

3.3.1. Expectations about Legal Rights. To see if the format of disclosure influences legal expectations, the respondents were asked two questions (in random order). They were asked if they believed they would have a legal claim if we disclosed their data to a third party in a manner consistent with our privacy policy. They also were asked if they believed they would have a legal claim if we disclosed their data to a third party in a manner not consistent with our privacy policy. For both questions, answers had to be provided on a 5-point scale ranging from “very unlikely” (scored as a 1) to “very likely” (scored as a 5), with “neutral” (scored as a 3) in between.

Figure 5A reports the mean responses (and 90 percent confidence intervals) for whether respondents believed they had a legal claim if we disclosed their data in a manner consistent with our privacy policy. In all treatment groups, the responses were roughly neutral (mean = 2.98), without any statistically significant differences across disclosure treatments or the blank group.

Figure 5B reports the mean responses (and 90 percent confidence intervals) for whether respondents believed they had a legal claim if we disclosed their data in a way not consistent with our privacy policy. More



Figure 5. Beliefs about legal claims

respondents thought that it was likely that they had a legal claim (mean = 3.57). The responses were best practice, 3.61; organization, 3.57; presentation, 3.50; specification, 3.52; worst practice, 3.49; and blank, 3.73. Unlike our prior tests, for this question three of the treatment groups had average responses that were different from the blank group at a statistically significant level: the presentation group ($p = .03$), the specification group ($p = .06$), and the worst-practice group ($p = .04$). That said, the differences between the best-practice group and these groups were not statistically significant. Although this is weak evidence that the presence of privacy disclosures may lull respondents into thinking that disclosers are acting legally and that no legal violation is being committed, it does not support the more nuanced notion that a best-practice disclosure would create different expectations than a poorly presented disclosure.

3.3.2. Reported Satisfaction. Finally, even if best practices do not change consumers' comprehension or decisions to share personal information, clearer disclosures may increase consumer satisfaction. We tested this by asking respondents two questions (in random order): "[h]ow satisfied are you that we take your privacy seriously?" and "[h]ow satisfied are you with your experience taking this survey?" For both questions, answers had to be provided on a 5-point scale ranging from "very dissatisfied" (scored as a 1) to "very satisfied" (scored as a 5).



Figure 6. Reported satisfaction

Figure 6A reports the mean levels of satisfaction (and 90 percent confidence intervals) that respondents had about whether we were taking their privacy seriously. Overall, the respondents were roughly satisfied that we were taking their privacy seriously (mean = 4.01). The responses were nearly identical across all six of the treatment groups, and none of the treatments had an average response that was different from the blank group at a statistically significant level. This is interesting, in part because our privacy disclosures announced a number of aggressively unpleasant policies that should diminish people's satisfaction. We already established that the content was largely ignored, and we now see that the format did not affect the overall impression that people gleaned from their exposure to the disclosure page.

Figure 6B reports the mean levels of satisfaction (and 90 percent confidence intervals) that respondents had taking our survey. As with the prior question on privacy, the respondents were roughly satisfied with their experience taking our survey (mean = 3.95), the responses were nearly identical across all six of the treatment groups, and none of the treatments had an average response that was different from the blank group at a statistically significant level. In other words, providing a privacy disclosure—no matter the formatting—did not change the respondents' attitudes about a survey asking highly intimate questions about their sexual history.

4. SIMPLIFICATION BY A WARNING LABEL

4.1. Overview

Our results thus far do not identify any statistically significant differences between the treatments groups' comprehension and behavior (and only very weak and unintended effects on legal expectations). In short, despite the intrusive information they were asked to share and the adverse privacy practices we pretended to have, the respondents who received the best-practice disclosure and those who received the worst-practice disclosure responded indistinguishably.

Perhaps this suggests that although best practices are widely supported, they may not go far enough. They still require more time to review and contemplate than most people are willing to spend. More radical forms of simplification may therefore be deemed necessary, and several templates along these lines have been offered. They all adopt the premise that disclosures should provide very few facts in a very short format and that they should focus on policies that are easiest for people to understand, process, and compare. Nutrition data labels are viewed as a model for these kinds of designs. One prominent version of this approach is the warning-label proposal in Ayres and Schwartz (2014)—a disclosure of legal terms containing only a handful of facts that are most surprising and disadvantageous to consumers.

Despite the fact that these warning-label-style disclosures have some intuitive appeal, we are unaware of research that has tested whether they lead consumers to behave differently than a more conventional best-practice disclosure would. We therefore added a treatment to our survey that would allow us to test the impact of warning-label disclosures. Since Ayres and Schwartz (2014) focus on privacy disclosures, we used their approach as our model.

4.2. Research Design

A short warning label needs to include only the items of utmost importance, and those might vary across firms. Accordingly, Ayres and Schwartz (2014) propose that firms should conduct research to learn which aspects of their privacy policies are least expected. In that spirit, we administered a short preliminary survey to our colleagues and students, asking them to examine the 12 terms in the privacy policy we developed for our experiment (as listed in Table 1). They were asked to rank the terms on a scale from “strongly unexpected” to “strongly expected” on the basis of what



Figure 7. Warning-label-style disclosure

they would expect for a survey on sexual practices being conducted by a corporation.¹⁷

On the basis of the responses to this preliminary survey, we created a warning label containing only the five items that were ranked as the most unexpected. Figure 7 presents the warning-label disclosure we designed, which is almost identical to the one Ayres and Schwartz (2014) propose. We then ran our experiment—exactly as Section 2.3 describes—but, instead of a standard privacy disclosure, presented some respondents with a warning-label disclosure.

4.3. Results

Since the results in Section 3 are nearly identical for all five of our privacy disclosures, Figures 8–10 simply report the comparison of the results for the best-practices treatment and the warning-label treatment.

Figure 8 reports the results for the two tests we designed to measure respondents' comprehension of the disclosures. Figure 8A shows that the amount of time spent on the disclosure screen was nearly identical for the respondents who received the best-practice treatment (mean = 19.12 seconds; median = 5.61 seconds) and the respondents who received the

17. For information about the preliminary survey, see part 6 of the online appendix.

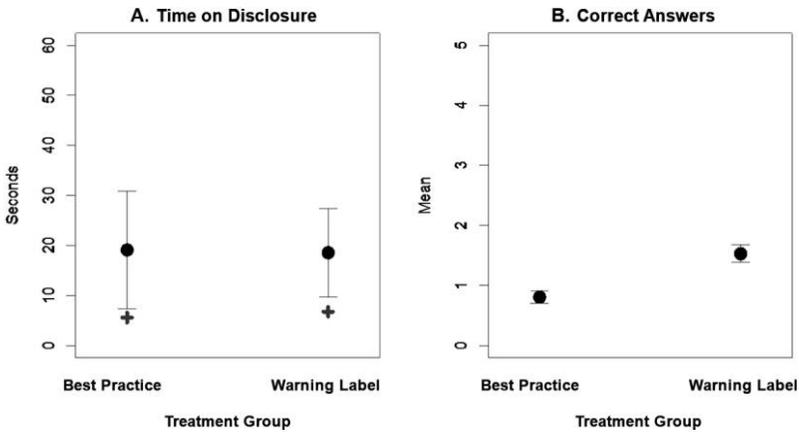


Figure 8. Comprehension of the warning label

warning-label treatment (mean = 18.56 seconds; median = 6.78 seconds). Figure 8B, however, shows that there was a difference in the number of correct answers to the comprehension questions. While respondents who received the best-practice disclosure averaged .81, the respondents who received the warning-label treatment averaged 1.53 correctly answered questions. This difference is highly statistically significant ($p < .001$).

Figure 9 reports the results of our two tests that measured respondents' willingness to share personal information. Figure 9A shows that the numbers of risky behaviors that respondents disclosed are comparable. On average, the best-practice group reported having engaged in 2.70 risky behaviors, and the warning-label group reported having engaged in 2.51 risky behaviors. This difference is not statistically significant. Figure 9B shows the numbers of answers the respondents in both groups provided to the questions asking for identifying information. On average, the best-practice group provided answers to 2.44 questions asking for identifying information, and the warning-label group provided answers to 2.72. This result is statistically significant at the .1 level ($p = .07$). In other words, this evidence suggests that respondents were slightly more likely to provide identifying information when we provided them with a warning label that listed only the most unexpected ways we would use their data.

Figure 10 reports respondents' answers regarding attitudes about our survey in terms of their legal expectations in case we shared their data with a third party and their overall satisfaction. For all four questions,

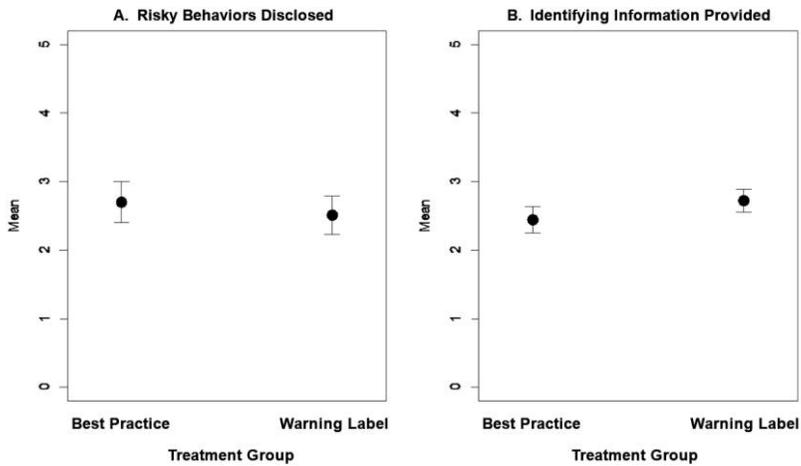


Figure 9. Willingness to share personal information

the differences among the best-practice group and warning-label group were minor and far from statistically significant.

4.4. Discussion

The warning-label group did show some improved understanding of the terms. That said, the respondents were far from fluent: only 1.53 of 5 answers were correct (compared with an average of less than 1 for other treatments). This is likely attributable to the ease of reading a five-item warning box relative to even a well-designed disclosure form. But this small improvement did not translate into any detectable difference in behavior. Only very minor differences were recorded in the willingness to share information despite the fact that the warning label cautioned people about the problematic uses of the information.¹⁸

These findings suggest different lessons compared with prior literature testing nutrition-box modules of privacy disclosures (Kelley et al. 2010). Like prior work, we also see some improved understanding (although the effect is more modest). But unlike prior work, we see no change in the primary conduct. When disclosures are presented in the context of another task, their presentation format is found to be irrelevant.

18. We also detected a modest and statistically significant increase in the subjects' dropout rate under the warning-label treatment. For the best-practice treatment, 3.4 percent of the respondents dropped out of the survey before the end of the questions on risky sexual behavior, and this rate increased to 6.3 percent under the warning-label treatment.

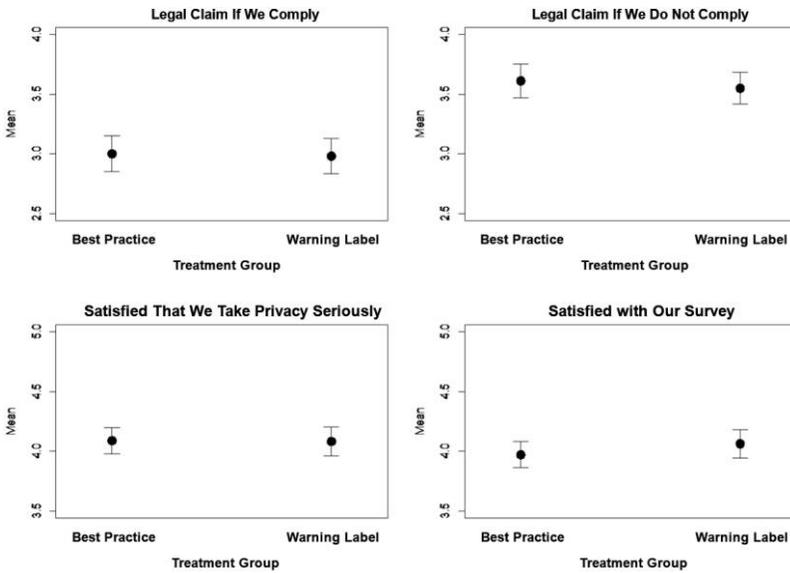


Figure 10. Attitudes about our survey

It is possible that the disappointing effect of the warning label merely reflects current numbness among subjects unmotivated to use privacy notices of any kind. It might be that in an environment where warning labels become the norm, people would learn to utilize them. Experimental methodology necessarily falls short of testing such market-wide effect, but some observational data cast doubt about this conjecture (Agarwal et al. 2015).

5. CONCLUSION

Simplification of disclosures is one of the most widely pursued regulatory techniques of our time. The simplification paradigm makes much sense in theory, but proof of its success has so far been elusive. Our experimental results contribute to a skeptical view of the merits of simplification. Of course, there are limitations to our approach. For example, it may be the case that people who agree to take an online survey responded to privacy disclosures differently than the general public would. In addition, we tested only two simplification formats—a best-practice presentation and a warning label—when there are surely more formats that could be

tried. But with those caveats in mind, we found that that the simplification of disclosures did not change people's understanding of them or their ensuing behavior in any meaningful direction.

On further reflection, these results may not be so surprising. As explained elsewhere, "simplicity's failure grows out of mandated disclosure's concern with complex and unfamiliar issues. Complexity can rarely be described simply to people unfamiliar with it" (Ben-Shahar and Schneider 2014, p. 135). Firms' privacy practices are only one complex topic in a host of issues that disclosures address. If simplification fails in this context, is there a reason to expect it to perform better in other areas?

REFERENCES

- Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel. 2015. Regulating Consumer Financial Products: Evidence from Credit Cards. *Quarterly Journal of Economics* 130:111–64.
- Alter, Adam L., and Daniel M. Oppenheimer. 2009. Suppressing Secrecy through Metacognitive Ease: Cognitive Fluency Encourages Self-Disclosure. *Psychological Science* 20:1414–20.
- Ayres, Ian, and Alan Schwartz. 2014. The No Reading Problem in Consumer Contract Law. *Stanford Law Review* 66:545–609.
- Ben-Shahar, Omri, and Adam S. Chilton. 2015. "Best Practices" in the Design of Privacy Disclosures: An Experimental Test. Working paper. University of Chicago Law School, Chicago.
- Ben-Shahar, Omri, and Carl E. Schneider. 2014. *More Than You Wanted to Know: The Failure of Mandated Disclosure*. Princeton, NJ: Princeton University Press.
- California Office of Privacy Protection. 2008. *Recommended Practices on California Information-Sharing Disclosures and Privacy Policy Statements*. Sacramento: California Office of Privacy Protection.
- FTC (Federal Trade Commission). 2012. *Protecting Consumer Privacy in an Era of Rapid Change: Recommendations for Businesses and Policymakers*. Washington, DC: FTC.
- Hertwig, Ralph, and Andreas Ortmann. 2008. Deception in Experiments: Revisiting the Arguments in Its Defense. *Ethics and Behavior* 18:59–92.
- Jensen, Carlos, and Colin Potts. 2004. Privacy Policies as Decision-Making Tools: An Evaluation of Online Privacy Notices. Pp. 471–78 in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, edited by Elizabeth Dykstra-Erickson and Manfred Tscheligi. New York: Association for

- Computing Machinery.
- John, Leslie K., Alessandro Acquisti, and George Loewenstein. 2011. Strangers on a Plane: Context-Dependent Willingness to Divulge Sensitive Information. *Journal of Consumer Research* 37:858–73.
- Kays, Kristina, Kathleen Gathercoal, and William Buhrow. 2012. Does Survey Format Influence Self-Disclosure on Sensitive Question Items? *Computers in Human Behavior* 28:251–56.
- Kelley, Patrick Gage, Joanna Bresee, Lorrie Faith Cranor, and Robert W. Reeder. 2009. A Nutrition Label for Privacy. *Proceedings of the Fifth Symposium on Usable Privacy and Security*, art. 4.
- Kelley, Patrick Gage, Lucian Cesca, Joanna Bresee, and Lorrie Faith Cranor. 2010. Standardizing Privacy Notices: An Online Study of the Nutrition Label Approach. Pp. 1573–82 in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: Association for Computing Machinery.
- Kleimann Communication Group. 2006. *Evolution of a Prototype Financial Privacy Notice: A Report on the Form Development Project*. Rockville, MD: Kleimann Communication Group.
- . 2012. *Know before You Owe: Evolution of the TILA-RESPA Disclosures*. Report presented to the Consumer Financial Protection Bureau. Rockville, MD: Kleimann Communication Group. http://files.consumerfinance.gov/f/201207_cfpb_report_tila-respa-testing.pdf.
- McDonald, Aleecia M., and Lorrie Faith Cranor. 2008. The Cost of Reading Privacy Policies. *I/S: A Journal of Law and Privacy for the Information Society* 4:540–65.
- Morton, Rebecca B., and Kenneth C. Williams. 2010. *Experimental Political Science and the Study of Causality: From Nature to the Lab*. New York: Cambridge University Press.
- National Telecommunications and Information Administration. 2013. Short Form Notice Code of Conduct to Promote Transparency in Mobile App Practices. Redline draft of July 25. Washington, DC. https://www.ntia.doc.gov/files/ntia/publications/july_25_code_draft.pdf.
- Noah, Timothy. 2000. The 1,000-Word Dash. *Slate*, February 18. http://www.slate.com/articles/briefing/articles/2000/02/the_1000word_dash.html.
- Pan, Yue, and George M. Zinkhan. 2006. Exploring the Impact of Online Privacy Disclosures on Consumer Trust. *Journal of Retailing* 82:331–38.
- Proctor, Robert W., M. Athar Ali, and Kim-Phuong L. Vu. 2008. Examining Usability of Web Privacy Policies. *International Journal of Human-Computer Interaction* 24:307–28.
- White House. 2012. *Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy*. Washington, DC: White House.