Online Privacy and Informed Consent: The Dilemma of Information Asymmetry

Masooda Bashir
Graduate School of Library & Information Science
University of Illinois at Urbana-Champaign
501 E. Daniel St.
Champaign, IL 61820
mnb@illinois.edu

April D. Lambert
Graduate School of Library & Information Science
University of Illinois at Urbana-Champaign
501 E. Daniel St.
Champaign, IL 61820
adlambe2@illinois.edu

Carol Hayes
College of Law
University of Illinois at Urbana-Champaign
504 E. Pennsylvania Ave.
Champaign, IL 61820
cmullin2@illinois.edu

Jay P. Kesan
College of Law
University of Illinois at Urbana-Champaign
504 E. Pennsylvania Ave.
Champaign, IL 61820
kesan@illinois.edu

ABSTRACT

Every day billions of users allow cloud-based internet services to collect, store, and manage their personal information. The use of this information is constrained only by long, wordy privacy agreements that users likely did not read before clicking "Agree." Even if they were to read them, would users understand these policies? We present the results of a two-part privacy survey that assessed users' knowledge and opinions of online privacy issues. We asked users not only what they think, but what they know. Results expose several key knowledge gaps, demonstrating a problem of information asymmetry between users and internet services providers, and strong dissatisfaction with the current system. These findings demonstrate that there is insufficient comprehension and voluntariness in the consent process for users to give informed consent to the collection and management of their personal information, which may in part explain the "privacy paradox."

Keywords

Online privacy agreements; informed consent; privacy paradox; information asymmetry; comprehension; voluntariness.

INTRODUCTION

ASIST 2015,November 6-10, 2015, St. Louis, MO, USA.

Researchers have long noted that there is a difference between what people *say* about the privacy of their personal information and what people actually *do* to protect the privacy of their information. This "privacy paradox" is the subject of much attention, with researchers trying to explain why behavior does not seem to follow from stated interests and intentions (Taddicken, 2014; Nissenbaum, 2009; Barnes, 2006). As this research and theorizing proceeds, billions of people from across the globe use cloud-based internet services provided by companies such as Facebook, Twitter, Amazon, and Google. As a result of the popularity of these and other cloud-based services, an enormous amount of personal information is now being stored and managed in the "cloud" rather than on users' computers. Users have made all of this information available to cloud providers, despite stated concerns about privacy. One reason for this paradox may be that users simply do not, or cannot, understand that they are acting contrary to their stated privacy beliefs and thus cannot give informed consent to the collection and use of their private information.

In the United States, few legal requirements specifically govern the storage and management of personal information, and users are left essentially to protect their own interests. Since the 1970's the international community has developed a set of Fair Information Practices (FIPs) to govern the protection of privacy and transborder flows of private information (Gellman 2014). Most recently revised by the Organisation for Economic Cooperation in 2013. FIPs emphasize five principles: (1) Notice/Awareness, (2) Choice/Consent, (3) Access/Participation, (4) Integrity/Security, and (5) Enforcement/Redress. Though none of the FIPs standards have been adopted as law in the United States, many companies refer to these standards when developing their privacy policies. For the most part, American companies focus their efforts only towards the first two, concluding that notice and consent suffices for fulfilling their privacy obligations (Bashir, Hoff, Hayes & Kesan, 2014).

Most cloud services attempt to satisfy the notice and consent requirements by requiring users to accept contracts called Privacy Policies (PPs) and Terms of Service (ToS) agreements. This moment of consent is one of the few moments where users are given an explicit choice between sharing information and protecting their privacy, but most often the choice is no real choice at all. Either the user agrees to give up all their personal information to the service or they choose not to use the service at all. And service providers do not make it easy for users to make this choice. Disclosure agreements often span several pages and are written in unwieldy "legalese." Several previous researchers have demonstrated that most people do not read these documents before accepting them (McDonald & Cranor 2008, Bakos et al. 2009; Plaut & Bartlett 2011). Even if users were to take the time to read all of the PPs and ToS agreements they encounter, would they be able to understand the policies? Though a great deal of time is saved by users by choosing not to even read the policies, considerable informational asymmetry exists between users and service providers regarding the collection and processing of personal information online. This asymmetry is compounded as online services evolve, policies are revised, and users continue to avoid these agreements. Considering this informational asymmetry, do users of online services really get meaningful and fair notice and consent under current practices?

One potential method for reducing this informational asymmetry is the application of the principles of *informed consent* to online environments (Friedman, Felton, and Millett, 2000). Informed consent can be defined as, "the process by which a fully informed user participates in decisions about his or her personal data" (van der Geest, Pieterson, and de Vries, 2005). To achieve informed consent online, five principles must be upheld: disclosure, competence, comprehension, voluntariness, and agreement (Friedman, Felton, and Millett, 2000). Of the five principles, comprehension and voluntariness are arguably the hardest to achieve under current internet business models and user behavior patterns.

Because most users do not take the time to read and understand privacy disclosures, their comprehension of service providers' policies is likely to be low. Comprehension here refers to an individual's accurate interpretation of the significance of disclosures. A user cannot give informed consent when the length, terminology, or organization of a privacy policy interferes with his comprehension of the contents of the policy. To be fully informed, individuals should be able to restate a disclosure agreement using different words and apply its contents to a different, but similar situation (Friedman, Felton, and Millett, 2000).

When users perceive that their acceptance of these privacy policies are not fully voluntary, it is also difficult to achieve informed consent online. Voluntariness refers to the degree that an agreement is coercive or manipulative. In contract law, a contract of adhesion exists when one party has all of the power and offers the terms on a "take it or leave it" basis. Adhesion contracts are not automatically unconscionable or coercive, but may be subjected to greater scrutiny. Cloud service providers typically offer only two options: "accept" or "cancel." Thus many ToS agreements and PPs online are adhesion contracts (Friedman, Felton, and Millett, 2000). Further, these agreements differ little between providers (Kesan, Hayes, and Bashir, 2013). A lack of alternative service delivery models (e.g., paying for privacy or accepting fewer features in return for privacy) or competition between providers can undermine the degree of voluntariness of users' consent.

In this paper, we present the results of a two-part survey that evaluated knowledge, opinions, and behavior surrounding comprehension, voluntariness, and informed consent in online disclosure agreements. The survey was designed to serve two main purposes. First, it assessed whether users were able to comprehend and give voluntary consent to online privacy policies. To do so, we included an extensive knowledge section that evaluates what users understand about a variety of online services issues, such as how cloud computing works and what cloud service providers commonly do with personal information stored in the cloud. Second, the survey explored opinions and perceptions related to potential methods that could improve comprehension and voluntariness in online consent agreements. Thus, this survey examines not only users' opinions about these privacy topics, but also assesses the knowledge of the terminology and policies of cloud services providers. We asked users not just what they think, but what they know

RELATED WORK

The literature concerning privacy and public perceptions of privacy has been growing in recent years. Scholars like Daniel Solove (2006) have examined the legal and theoretical foundations of informational privacy concerns. Lorrie Cranor and others have emphasized the practical implications of these concerns in studies of user behavior (Leon et al., 2013; Sunshine, Egelman, Almuhimedi, Atri, and Cranor, 2009). Other researchers, such as Helen Nissenbaum (2009), have examined the problem defined as a "privacy paradox", which refers to a discrepancy between what people say they want in terms of privacy and what they actually do. For example, survey research suggests that public concern about online privacy issues is substantial (Rainie, 2013). A recent Pew survey concludes that 91% of respondents felt that "consumers have lost control over how personal information is collected and used by companies"

(Madden, 2014). Yet behavioral research has shown that the vast majority of people do not read, or even access, terms and conditions before accepting them (Bakos, Marotta-Wurgler, and Trossen, 2009). This disconnect may be in part the result of informational asymmetry between users, cloud service providers, and other companies who profit from personal information online.

Several researchers have called for the application of informed consent principles to online environments in order to increase the ability of users to monitor and control the flow of their personal information. For example, Friedman et al. (2002) proposed four goals for Web browser redesign that would make browsers more conducive to informed consent in regard to cookie management capabilities. Van Der Geest et al. (2005) outlined six items that should be included in user profiling consent forms in order for them to parallel the informed consent process used in the medical settings. In a study of purchasing behavior, Tsai et al. (2011) showed that increasing the saliency of privacy policy information led users to utilize the information more when selecting online vendors.

Our survey draws on insights gained from these types of sources and expands the coverage to significantly more topics. In particular, we chose to emphasize user knowledge because research has yet to address this topic in much detail. We hope to contribute the most detailed empirical information to date about people's knowledge on a variety of important online topics, such as how cloud computing works and what cloud service providers commonly do with personal information stored in the cloud. On the opinion-side, we explore perceptions and actions related to voluntariness in submitting information online. Furthermore, we evaluate user demand for potential legal regulations and alternative service delivery models that, if implemented, could increase comprehension and voluntariness in online privacy agreements.

METHODOLOGY

The online survey was divided into two parts which were administered separately, but could be accessed from the same website. Respondents were randomly presented with either part one or part two listed first when they visited the site. Upon completion of one part, respondents were could complete the other part. Respondents were permitted to terminate their participation at any point, so not all respondents fully completed both parts. Part one, fully completed by 455 people, evaluated knowledge; part two, at least partially completed by 756 people, assessed opinions and online behavior. There were right and wrong answers for almost all of the questions in part one, but the questions in part two were primarily subjective. We thoughtfully constructed the questions with an eye towards minimizing bias. For example, because sufficient comprehension requires that individuals be able to apply the contents of disclosure agreements to different situations (Friedman, Felton, and Millett, 2000), most of the knowledge-based questions were written as third-person scenarios. The knowledge survey questions were organized into five sections based on subject matter: 1. Cloud Computing; 2. Online Security; 3. Economics of the Internet; 4. Educational Records; and 5. Legal Aspects of Online Privacy. Thus, the first two sections focused primarily on technical aspects of the internet while the last two sections addressed primarily legal issues (with a focus on U.S. laws). Section three, on the other hand, assessed knowledge surrounding the data trade for personal information online. The survey contained 42 questions. Of the 42 questions, 28 were multiple-choice and 14 were true/false. The second part of the survey was organized into three sections: 1. Online Behavior; 2. Personal Privacy; 3. Cloud Service Providers. Examples of the types of questions in this part of the survey are included in the Survey Results section below, where relevant.

The survey link was distributed through email and social media in primarily academic settings at a large Midwestern university in the United States. Respondents included university students, staff, and faculty. Survey completion was promoted through a prize drawing for a $10 Amazon gift card. Participants were encouraged to complete both parts of the survey, but many completed just one

SURVEY RESULTS

The results for each part of the survey are discussed separately in this section. Section A covers the knowledge survey while Section B discusses the opinion and online behavior survey.

Knowledge

Demographics

We received a total of 455 complete responses to the knowledge survey, with slightly more female respondents than male respondents (55% female, 45% male). In terms of age, 37% of respondents were between 18-21 years old (the age of typical undergraduate students in the U.S.), 28% were between 22-29, 16% were between 30-39, and 19% were age 40 or older. The vast majority of respondents were either White (67%) or Asian (19%). Overall, the sample was highly educated; 60% of respondents had completed an undergraduate-level degree and almost half of this figure (29% overall) had also completed a graduate-level degree.

Overall knowledge scores

Knowledge scores were computed by assigning equal weight to all 42 questions, with question subparts assigned a fractional value of that question. We converted raw scores into percentages and then calcul-

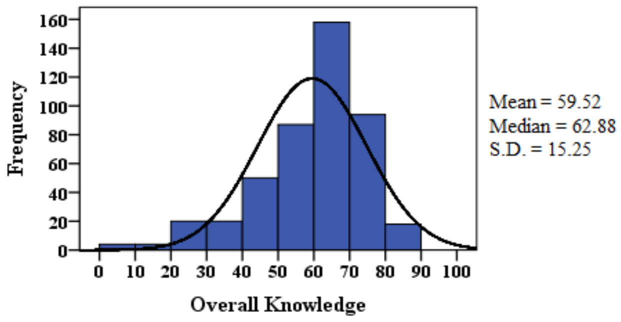

Figure 1. Frequesncy distribution for overall knowledge scores

ated mean scores for each section and overall. Figure 1 above displays the frequency distribution for overall knowledge scores. The middle 50% of scores fell between 52% and 70%.

Figure 2 displays participants' mean knowledge scores for each of the five sections as well as their overall score. Error bars indicate ±1 standard deviations. As can be seen, the mean scores (proportions of correct answers within each section) were highest in the section that focused on the Economics of the Internet and lowest in the sections that focused on Educational Records and the Legal Aspects of Privacy.

When comparing the mean scores above, it is important to note that the specificity of the questions varied by subject matter. For example, the questions in the Educational Records and Legal Aspects of Privacy sections were based on specific provisions of different U.S. laws designed to protect individuals' privacy. In contrast, the Economics of the Internet section included more general questions about how online companies make money on the internet. Performance in one section is not necessarily indicative of performance on another.

We evaluated demographic differences in knowledge by performing one-way between subjects analyses of variance (ANOVAs) using overall and section scores (the section-by-section results are discussed in the corresponding sections below). For highest level of education, we grouped respondents into one of three ordinal categories (non-college graduate, college graduate, and advanced degree). Overall knowledge scores differed significantly across the three groups, F (2, 450) = 21.97, p < .05. Tukey post-hoc comparisons of the three groups indicated that the non-college graduates (M = 54.21, 95% CI [51.70, 56.73]) had significantly lower scores than those with college or advanced degree.

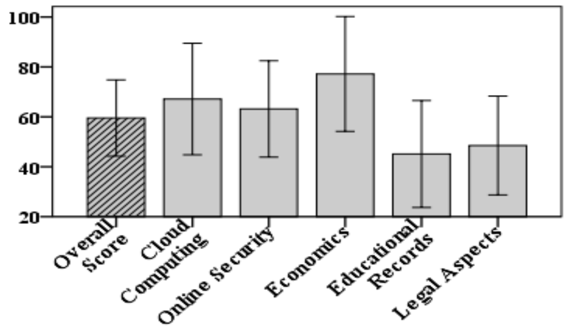

Figure 2. Mean knowledge scores

Additionally, we investigated whether age or education level predicts privacy knowledge, controlling for each other. Results of multiple regression analyses showed that age (b = 2.29, p < .05) and education level (b = 3.08, p < .05) both predicted overall scores. This result regarding age belies any assumption that younger users are more knowledgeable about technology-related subjects.

Knowledge of cloud computing

The first section assessed respondents' basic knowledge of cloud computing and how it works. Figure 3 below presents the frequency distribution of knowledge scores within the section. A breakdown of results in this section by category demonstrates that while respondents knew some of the uses of cloud computing, many failed to understand the breadth of services reliant on cloud computing, betraying a lack of knowledge about the fundamental nature of the technology. In this section, respondents performed worst on a question that asked about examples of cloud services. This was largely driven by the fact that more than half of respondents failed to indicate that Twitter and Netflix offer cloud-based services. On the other hand, respondents were able to answer questions about the benefits of cloud storage. Specifically, about 83% of respondents knew that cloud computing offers individual users a high storage capacity and about 89% knew that cloud computing can be used to back up and recover digital files.

Surprisingly, considering the low rate of correct responses on the question asking for examples of cloud services, 75% of respondents were able to correctly answer a question that essentially asked, "What is cloud computing?" This finding contrasts with a 2012 survey in which 51% of respondents answered "yes" to a question about whether stormy weather interferes with cloud computing (Wakefield Research, 2012). This discrepancy may be due to the fact that our

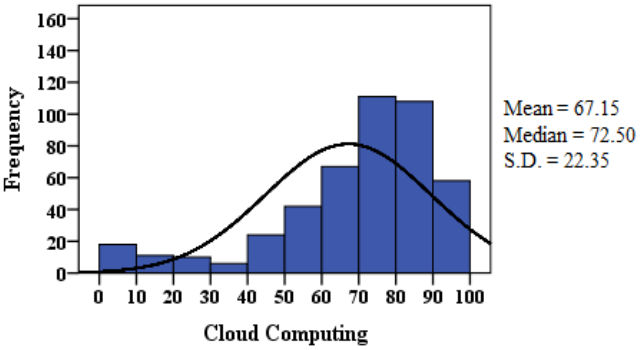

Figure 3. Frequency distribution for cloud computing knowledge.

survey was distributed in an academic setting while the other survey was given to a nationally representative sample of the U.S. adult population. Nevertheless, one quarter of our respondents got the question wrong, which suggests that cloud computing is still a perplexing concept to some people.

Overall, the response patterns in this section suggest that users are generally aware of the existence of cloud computing and some of its uses, such as backing up digital files. However, users are less knowledgeable about other aspects of cloud computing, such as its fundamental nature and the full scope of services that depend on cloud computing (e.g., storing movie and television show preferences on Netflix). This is problematic because if users do not understand the basics of the technology's design, how can they be expected to give informed consent? Without this basic aspect of comprehension, the consent process has significantly less meaning.

Knowledge of online security

Our goal in this second section was to assess what people understood about basic issues related to cyber security. Figure 4 presents the frequency of knowledge scores within the section. The specific question on which participants performed worst in this section dealt with the risks of using public Wi-Fi. Only 60% of respondents correctly identified the primary risk associated with using a public WiFi network (i.e., someone else using the network could potentially see all the websites that one visits and the information that one transmits). Additionally, 76% correctly answered questions about everyday security risks and where popular online services store customer data, while about 84% of respondents correctly answered a question that asked about the definition of encryption.

Participants also performed poorly on three questions that asked about where different aspects of personal information are stored online. Most notably, nearly a quarter of respondents did not know that free webmail services, such as Gmail, almost always store customer information on their own servers, which are connected to the internet. This figure is quite similar to the percentage of respondents (25%) who failed to correctly answer the first question from the cloud computing section that asked about the nature of cloud computing. Thus, it seems that about one quarter of our sample is confused about cloud computing and its application to webmail services. Again, this shows that many millions of people lack a basic comprehension about cloud computing that is necessary to interpret the significance of the terms within the privacy disclosures of cloud service providers.

Knowledge of the economics of the Internet

The questions in this section addressed how companies make money on the internet, with a focus on behavioral advertising. The questions here were less technical in nature compared to the first two sections, so the higher scores in this section are not indicative of performance in other sections. See Figure 5 for frequency distribution.

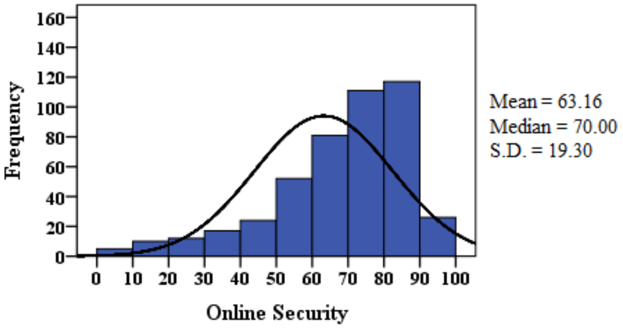

Figure 4. Frequency distribution for online security knowledge.

Although on average respondents performed better in this section than any other, the results far from indicate that people have sufficient knowledge to provide informed consent allowing private companies to profit from their personal information online. For example, in one of the two questions that asked about how free websites make money, 44% of respondents did not know that free websites could profit by selling user information directly to marketing companies. In one of three questions that asked about how targeted advertisements are created, 66% of respondents correctly answered that advertising companies could use emails sent and received on free webmail accounts to personalize advertisements. On average 68% of respondents correctly answered questions about what online advertising companies can do to collect personal information about users. These response patterns highlight a deficiency of comprehension in regard to the data trade for personal information on the Internet.

Knowledge of the legal aspects of online privacy

In the last two sections of the knowledge survey, we explored knowledge regarding the Family Educational Records and Privacy Act (FERPA), a federal law that governs the privacy of educational records in the United States, and other American privacy laws. We asked about FERPA, in addition to more general privacy laws, because the survey was deployed on a university campus, and we

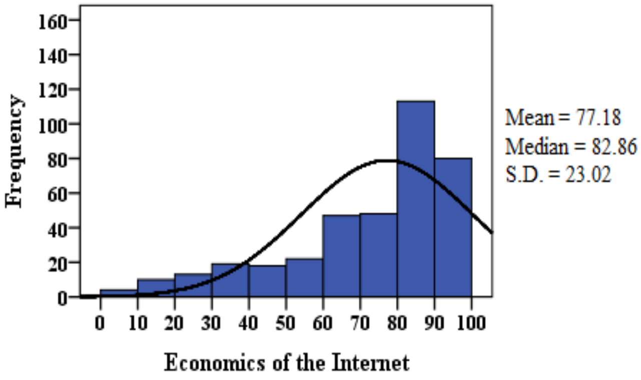

Figure 5. Frequency distribution for economics of the Internet

wanted to assess knowledge regarding a privacy issue highly relevant to many respondents. Participants performed the worst in the FERPA section, averaging only 2.7 correct responses out of 6, while they averaged about 4.4 correct answers out of 9 in the legal section. Due to space limitations, more specific results will not be discussed in this section.

Opinions and Online Behavior

Demographics

A total of 756 individuals completed the opinion survey. There were significantly more female respondents than male respondents (62% female, 36% male). In terms of age, 43% of respondents were between 18-21 years old (the age of typical undergraduate students in the U.S.), 25% were between 22-29, 14% were between 30-39, and 17% were age 40 or older. The vast majority of respondents were either White (68%) or Asian (17%). For highest level of education, 49% reported high school/GED, 27% reported an undergraduate-level degree, and 23% reported a graduate level degree.

Online behavior

The questions in this first section gathered information about respondents' behavior on the Internet. The sample appeared to consist of mostly frequent Internet users, with about 83% of respondents reporting that they use the Internet for at least three hours each day. In terms of online activity, about 99% of respondents indicated that they had used a webmail service before, 97% had used a social networking site, and 81% had used an online file storage service. This indicates that almost all of our respondents had previously used cloud-based services.

We also asked a question about the different types of information participants had previously provided online. Notably, over 99% of respondents indicated that they had provided their credit card information online, 78% said they had provided information about their religion, and 66% indicated that they had provided their social security number. Upon further examination of the "information about your religion" aspect of this question, we found a fascinating discrepancy between those who said "no" and responses to our own demographic question about religious affiliation. Of the 160 people (22% of the total sample) who said they had never before provided information about their religion online, 89 reported a specific affiliation in our survey. This suggests that some people are not fully aware of all of the personal information they provide online—they may do so unconsciously. It should be noted, however, that our survey was anonymous, so some respondents may not have perceived their responses as counting for the purposes of the question. Whatever the reason, it is problematic in that it demonstrates unconscious data share. Anonymity on the internet sometimes lulls users into the belief that their privacy is secure, but research has shown that it is not extremely difficult to de-anonymize collections of data (Ohm, 2009). Online anonymity is not the same thing as invisibility.

The responses to the question "Have you ever submitted information online, but wished you didn't have to?" illustrate the coercive nature of "voluntariness" in online environments: 81% of respondents (n=727) responded positively. Although some private companies may claim otherwise, the reality is that people are often compelled to provide personal information in order to use a website or online service. In our survey, more than 4 out of every 5 respondents indicated that this had happened to them. There was a significant positive correlation with age (r = .09, p < .05), suggesting that older people are more likely to have had regrets about submitting information online.

In order to evaluate respondents' behavior in relation to online consent agreements, we asked questions about how often people read Terms of Service (ToS) agreements and Privacy Policies (PPs). The results were in line with that of other researchers who have demonstrated that users rarely read such policies (Bakos, Marotta-Wurgler, and Trossen, 2009; Plaut and Bartlett, 2011; McDonald and Cranor 2008).

Responses to the question "Have you ever decided not to use a website strictly because of the website's Privacy Policy or Terms of Service Agreement?" indicate that slightly less than half of our respondents (43%, n=707) had ever made a conscious decision not to use a website because of the website's ToS agreement or PP. Not surprisingly, we found that those who read ToS agreements and PPs more often were more likely to indicate that they had refused to use a website strictly because of the website's PP or ToS agreement. Importantly, these correlations were very strong (r = .50, p < .01 for PP readership; r = .50, p < .01 for ToS readership), suggesting that the relevance of these documents could be significantly increased if people read them more.

The final question in this section focused on decision-making on ecommerce websites. We wanted to determine the factors that people care about most when shopping online. The most important factors were reputation (85%) and price (85%) and the least important factors were the contents of a website's PP (12%) and ToS agreement (10%). This finding provides further evidence that most people are either unaware of, or disinterested in, the significance of these documents.

Opinions about personal privacy

The questions in this section were designed to assess respondents' general beliefs about personal privacy, both online and offline. We also wanted to explore the extent to which these beliefs relate to online behavior and opinions about cloud service providers.

In general, our sample appeared to have strong opinions about their personal privacy. Over 92% of respondents somewhat or strongly agreed with the statement, "Personal privacy is important to me." Females (r = .10; p < .05) and more educated people (r = .09, p < .05) were more likely to express stronger agreement. Additionally, responses to this question proved influential in regard to online behavior. Respondents who care more about their personal privacy were more likely to indicate that they read PPs (r = .15, p < .01) and that they have previously refused to use a website because of the website's PP or ToS agreement (r = .17, p < .01).

We also included two agreement questions that assessed whether or not respondents subjectively thought that their online behavior was influenced by privacy and security concerns. About 74% and 77% of respondents reported that online privacy concerns and online security concerns, respectively, influenced their online behavior. These responses seem to contradict the responses to the question displayed in Figure 7, which shows that only 43% of respondents indicated that they have ever decided not to use a website strictly because of the contents of the website's PP or ToS agreement. This discrepancy is indicative of the "privacy paradox," which refers to a disconnect between user preferences for privacy and user behavior in relation to privacy (Nissenbaum, 2009).

We also explored respondents' feelings about data collection online. The set of questions in Table 1 below report approval rates in regard to what people think online marketing and advertising companies *should* be allowed to do. We instructed participants to answer the questions based solely on their own opinion, disregarding all current laws.

Should online marketing and advertising companies be allowed to . . .	Yes (%)
track users' online activity without asking for permission? (n=633)	11
participate in a data trade in which users' online activity on multiple websites is combined to create demographic profiles? (n=627)	27
use demographic profiles to display advertisements that are relevant to individual users? (n=627)	52
track users' interactions with advertisements to determine which ones are the most relevant to individual users? (n=627)	52

Table 1. Opinions about Online Marketing and Advertising

The distribution of responses above provides valuable insight into the opinions that many people may have on these topics. The strongest majority (89%) against a particular practice was found for the issue concerning online tracking without users' permission. Responses to the last two questions concerning targeted advertising were more evenly split between support and opposition. Interestingly, responses to both targeted advertising questions were positively correlated with internet use per day (r = .08, p < .01 for "Use demographic profiles…"; r = .09, p < .01 for "Track users' interactions…"), suggesting that people who use the internet more are more likely to support targeting advertising. One potential explanation for this finding is that individuals who use the internet more are more likely to experience benefits from targeting advertisements.

Opinions about cloud service providers

This section assessed what people think and know about how cloud service providers collect, maintain, and distribute customer information. The first set of questions displayed in Table 2 below was inspired by our previous empirical analysis of the ToS agreements and PPs of nineteen leading cloud service providers (Kesan, Hayes, and Bashir, 2013). We assessed whether respondents were previously aware of how cloud service providers can use customer uploaded information due to provisions commonly found in their privacy agreements. Additionally, we evaluated whether respondents approved or disapproved of each method once made aware.

Statement	Previously Aware (%)	Approve (%)
Q1. Some cloud service providers may reproduce or modify works of content uploaded by users. (n=528)	29	6
Q2. Some cloud service providers may publish, publicly display, or distribute content uploaded by users. (n=529)	46	12
Q3. Some cloud service providers may search through content uploaded by users to find keywords that can be used as the basis for displaying targeted advertisements (n=528)	59	32

Table 2. Awareness and Approval Rates of Cloud Service Providers Use of Users' Content

The responses patterns above reveal significant gaps in people's awareness of common themes within the PPs and ToS agreements of leading cloud service providers. Not surprisingly, PP readership was positively related to the "previously aware" aspect of all three questions (r = .18, p < .01for Q1; r = .13, p < .01 for Q2; r = .19, p < .01 for Q3, respectively). This finding provides further support to the argument that comprehension is severely lacking in the online consent process, particularly when considering the fact that 99% of our sample had used a webmail service before and 95% had uploaded personal photos online.

Disapproval rates for each method were staggeringly high, indicating that a significant majority of our respondents are displeased with the current system. When asked for opinions about U.S. laws designed to protect individuals' privacy online (n=495), 78% responded that they were too weak, 20% thought they were sufficient, and 2% indicated that current U.S. laws were too strict. These responses indicate an acknowledgment that the current system that depends on users to protect their own privacy is not working for many users. Many users are in fact open to more governmental regulatory oversight of the online privacy environment.

Voluntariness is a major issue in the consent process for cloud service providers. Voluntariness could potentially be improved if service providers adopted a different business model that gave users more options for controlling their own privacy settings. If a cloud service provider were to adopt new policies that give users different options for privacy settings, 72% (n=554) stated that they would prefer to get fewer features in return for increased control over their personal information for no cost. Only 9% stated that they would prefer free access to more features in return for less control over their personal information. We believe that these results are an important indicator of users' need for alternative consent models. This finding supports that from a recent survey in Brazil (Dos Santos Brito, Durao, Garcia, and de Lemos Meira, 2013), which showed that there may be a legitimate demand for alternative business models.

DISCUSSION

One of the primary motivations behind our survey was to assess users' knowledge levels and opinions to determine if they were actually giving *informed consent* in online environments, with an emphasis on comprehension and voluntariness in the consent agreements of cloud service providers. This may, at least in part, explain the privacy paradox. Without evidence of informed consent by users, there cannot be fair and meaningful notice and consent regarding the collection and use of personal information by cloud services providers.

Due to the academic setting in which the survey was distributed, we received a high proportion of responses from highly educated individuals, including university students, staff, and faculty. For example, in the knowledge survey, 29% of respondents had a graduate-level degree, which is almost three times the national average of 10.3% in the United States (U.S. Census Bureau). Thus, the knowledge of respondents in our sample is likely greater than that of the "average" person, especially due to the fact that respondents needed Internet access in order to take the survey. The results of the knowledge survey should be interpreted with this in mind. Also, significantly more respondents completed the opinion and online behavior part of the survey than the knowledge section. Perhaps this was because respondents were frustrated with their lack of knowledge and quit the survey early.

Our key results are as follows: (a) respondents lack sufficient comprehension about several key privacy issues commonly stated in online consent agreements; (b) the vast majority of respondents have experienced a situation in which they felt coerced into submitting information online; (c) there is a legitimate demand for alternative service delivery models concerning privacy online; (d) a majority of users expressed support for legislation that would require all cloud service providers to use standardized consent agreements; and (e) there is evidence of unconscious data share behavior by respondents. These key results provide insight into whether users have the capacity to give informed consent regarding their privacy in the online environment.

Comprehension

Comprehension is deficient in the online consent process because people are often unaware of how websites and cloud service providers handle their personal information. Our knowledge data clearly shows that many people lack sufficient comprehension about several important issues related to cloud computing, online security, and the data trade for personal information online. For example, 25% of respondents were unable to correctly answer a basic multiple-choice question about the fundamental nature of cloud computing, 23% were unaware that free webmail services store customer information in the cloud, and 44% did not know that free websites could make money by selling user information directly to marketing companies.. Furthermore, because the knowledge-based questions were either true/false or multiple-choice, respondents were likely able to guess their way to higher knowledge scores by eliminating improbable answer options.

In part two (the opinion and online behavior survey), we found that although 99% of our sample had used a cloud service before, most people were unaware of common provisions within the ToS agreements and PPs of leading cloud service providers. When people were explicitly told how some cloud service providers handle content uploaded by users, they overwhelmingly disapproved. This suggests that people would make different decisions about where to store their personal information if they were made explicitly aware of the differences between the privacy practices of different cloud service providers. Further support for this hypothesis came from the finding that although 43% of respondents had ever refused to use a website solely because of the contents of the website's PP or ToS agreement, individuals were significantly more likely to have done so when they indicated that they more frequently read PPs and ToS agreements. This also accords with recent findings by the Pew Research Center, which showed that 61% of respondents "would like to do more" to protect their online privacy, yet very few were taking steps to protect themselves, such as using encryption or other steps to anonymize their online footprint (Madden, 2014). Despite their desire to do more to protect themselves, many Americans do not take available steps to do so, perhaps because they lack the knowledge to be able to recognize and implement these steps.

Restructuring privacy policies, clarifying policy language, or a legislative requirement for standardized policies are possible solutions for improving the demonstrated deficiency in comprehension. For example, multilayered policies that summarize the document at the top layer and give increasing amounts of detail at lower layers may lead to increased comprehension (Good et al., 2005; Iachello and Hong, 2007). Additionally, the language of the agreements should be accessible, because even short warnings can be easily ignored or misunderstood if users do not know what warnings mean (Sunshine, Egelman, Almuhimedi, Atri, and Cranor, 2007). Another recent study showed that many Americans would favor greater government restrictions on the information collected by advertisers (Madden, 2014); perhaps standardized agreements would be a possible government restriction. This would greatly reduce the possible opportunity costs of reviewing multiple policies. In our survey, about 60% of respondents expressed support for such a law. Another potential solution could be Knowledge-based Individual Privacy Plans, or KIPPS, that could assess gaps in knowledge and generate individualized plans for managing privacy (Kesan, Hayes, and Bashir, 2015). Other researchers have suggested the implementation of privacy labels (Kelley, Cranor, and Reeder, 2009) or icons (Holtz, Zwingelberg, and Hansen, 2011). All of these solutions seek to "cure" information asymmetry by providing simpler, more standardized information for users who are forced to make decisions despite not fully comprehending their privacy options and consequences. They are all attempts to achieve informed consent.

Voluntariness

Voluntariness is also a troubling issue online today because people are often presented with a forced choice dilemma: accept a cloud service provider's terms and conditions or do not use the service. While such a forced choice may not meet the level of coercion that would invalidate a contract under the law, it cannot be seen as meaningfully voluntary. Our survey results illustrate this perception of coercion, as the vast majority of respondents (81%) indicated that in at least one incident, they had submitted information online when they wished they were not required to do so.

Voluntariness could be improved if websites and/or cloud service providers adopted alternative service delivery models that gave users more options for controlling their own privacy settings. Our survey results suggest that people may be open to different possibilities. When given a choice between paying for more privacy, accepting fewer features in return for more privacy, or having less privacy in return for more features, only 9% of respondents selected the third option (which is analogous to the current business model). Thus, 91% of respondents believe that the current system is inadequate. Although our survey question was largely hypothetical, parallel findings from another recent survey (Dos Santos Brito, Durao, Garcia, and de Lemon Meira, 2013) suggest that the demand for alternative service delivery models is worthy of further exploration.

CONCLUSION

Our survey offers valuable insights into what people do, know, and want when it comes to the availability of personal information on the internet. In both the knowledge and opinion sections, the data provide clear evidence of an informational asymmetry between users and service providers with regard to the contents of disclosure agreements. Such informational asymmetry undermines the comprehension and voluntariness of informed consent to online privacy agreements. Without informed consent meaningful notice and consent cannot be achieved, and the privacy paradox perpetuates. Additionally, cloud services providers are not truly in compliance with the fair information practices to which they claim to subscribe. Our survey provides stark evidence that not only do users not read privacy policies, they likely would not be able to understand them if they did. Online privacy problems will likely grow in number and severity in the coming years if the informational asymmetry between users and cloud service providers is not addressed and fair information practices are not adopted. Thus, further research is needed to replicate and extend the trends discussed in this paper and to identify new areas of focus such as social, educational, and cultural factors that may influence comprehension and voluntariness in online consent agreements.

ACKNOWLEDGMENTS

Thank you to Boyi Guo for statistics support.

REFERENCES

Bakos, Y., Marotta-Wurgler, F., & Trossen, D. R. (2009). Does anyone read the fine print? Testing a law and economics approach to standard form contracts. In *Testing a Law and Economics Approach to Standard Form Contracts (October 6, 2009). CELS 2009 4th Annual Conference on Empirical Legal Studies Paper* (pp. 09–40). Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1443256

Barnes, S. (2006). A Privacy Paradox: Social networking in the United States. *First Monday*. 11.9. Retrieved from http://firstmonday.org/ojs/index.php/fm/article/viewArticle/1394/1312%2523

Bashir, M., Hoff, K.A., Hayes, C.M., Kesan, J.P. (2014). Knowledge-based Individual Privacy Plans (KIPPs): A Potential Tool to Improve the Effectiveness of Privacy Notices. Carnegie Mellon University, CyLab Workshop on the Future of Privacy Notice and Choice, June 27th, 2014. Pittsburgh, PA. Available at https://www.cylab.cmu.edu/news_events/events/fopnac/pdfs/bashir.pdf

Day, G. S., & Brandt, W. K. (1974). Consumer research and the evaluation of information disclosure requirements: The case of truth in lending. *Journal of Consumer Research*, 21–32.

Dos Santos Brito, K., Cardoso Garcia, V., Araujo Durao, F., & Romero de Lemos Meira, S. (2013). How people care about their personal data released on social media. In 2013 Eleventh Annual International Conference on

Privacy, Security and Trust (PST) (pp. 111–118). doi:10.1109/PST.2013.6596044

Friedman, B., Felton, E., & Millett, L. I. (2000). *Informed consent online: A conceptual model and design principles*.

Friedman, B., Howe, D. C., & Felten, E. (2002). Informed consent in the Mozilla browser: implementing Value-Sensitive Design. In *System Sciences, 2002. HICSS. Proceedings of the 35th Annual Hawaii International Conference on* (p. 10–pp). IEEE. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=994366

Gellman, R. (2014). *Fair Information Practices: A Basic History*. Retrieved from http://bobgellman.com/rg-docs/rg-FIPShistory.pdf

Good, N., Dhamija, R., Grossklags, J., Thaw, D., Aronowitz, S., Mulligan, D., & Konstan, J. (2005). Stopping spyware at the gate: a user study of privacy, notice and spyware. In *Proceedings of the 2005 symposium on Usable privacy and security* (pp. 43–52). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1073006

Holtz, L. E., Zwingelberg, H., & Hansen, M. (2011). Privacy policy icons. In*Privacy and Identity Management for Life* (pp. 279-285). Springer Berlin Heidelberg.

Iachello, G., & Hong, J. (2007). End-user privacy in human-computer interaction. *Foundations and Trends in Human-Computer Interaction*, *1*(1), 1–137.

Kelley, P. G., Bresee, J., Cranor, L. F., & Reeder, R. W. (2009, July). A nutrition label for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security* (p. 4). ACM.

Kesan, J.P., Hayes, C.M., & Bashir, M.N. (2013). Information Privacy and Data Control in Cloud Computing: Consumers, Privacy Preferences, and Market Efficiency. *Washington and Lee Law Review*, *70*, 341.

Kesan, J.P., Hayes, C.M., & Bashir, M.N. (2015). A Comprehensive Empirical Study of Data Privacy, Trust, and Consumer Autonomy. Retrieved from SSRN: http://ssrn.com/abstract=2576346

Leon, P. G., Ur, B., Wang, Y., Sleeper, M., Balebako, R., Shay, R., … Cranor, L. F. (2013). What matters to users?: factors that affect users' willingness to share information with online advertisers. In *Proceedings of the Ninth Symposium on Usable Privacy and Security* (p. 7). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2501611

Madden, M. (2014). Public Perceptions of Privacy and Security in the Post-Snowden Era. Retrieved from http://www.pewinternet.org/2014/11/12/public-privacy-perceptions/

McDonald, A. M., & Cranor, L. F. (2008). Cost of reading privacy policies, the. *ISJLP*, *4*, 543.

Nissenbaum, H. (2009). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press.

Ohm, P. (2009). *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization* (SSRN Scholarly Paper No. ID 1450006). Rochester, NY: Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=1450006

Plaut, V. C., & Bartlett III, R. P. (2012). Blind consent? A social psychological investigation of non-readership of click-through agreements. *Law and Human Behavior*, *36*(4), 293.

Rainie, L. (2013, September 5). Anonymity, Privacy, and Security Online. Retrieved from http://www.pewinternet.org/2013/09/05/anonymity-privacy-and-security-online/

Solove, D. J. (2006). A taxonomy of privacy. *University of Pennsylvania Law Review*, 477–564.

Sunshine, J., Egelman, S., Almuhimedi, H., Atri, N., & Cranor, L. F. (2009). Crying Wolf: An Empirical Study of SSL Warning Effectiveness. In *USENIX Security Symposium* (pp. 399–416). Retrieved from https://www.usenix.org/legacy/events/sec09/tech/full_papers/sec09_browser.pdf?origin =publication_detail

Taddicken, M. (2014). The 'Privacy Paradox' in the Social Web: The impact of privacy concerns, individual characteristics, and the perceived social relevance on different forms of self-disclosure. *Journal of Computer-Mediated Communication,* 19.2: 248-273.

Tsai, J. Y., Egelman, S., Cranor, L., & Acquisti, A. (2011). The effect of online privacy information on purchasing behavior: An experimental study. *Information Systems Research*, *22*(2), 254–268.

US Census Bureau, D. I. D. (n.d.). Educational Attainment. Retrieved September 4, 2014, from http://www.census.gov/hhes/socdemo/education/

Van der Geest, T., Pieterson, W., & de Vries, P. (2005). Informed consent to address trust, control, and privacy concerns in user profiling. *Privacy Enhanced Personalisation, PEP*, 23–34.

Wakefield Research. (2012). *Citrix cloud survey*. Retrieved from https://s3.amazonaws.com/legacy.icmp/additional/citrix-cloud-survey-guide.pdf