A Comprehensive Empirical Study of Data Privacy, Trust, and Consumer Autonomy

JAY P. KESAN,* CAROL M. HAYES** & MASOODA N. BASHIR***

Modern society is driven by data. Data storage is practically unlimited with today's technology, and analytical tools make it easy to find patterns and make predictions in a way that is very useful for private businesses and governments. These uses of digital data can raise considerable privacy issues that are of great concern to consumers. In this Article, we present and analyze the results of an extensive survey that we conducted to explore what people know, what people do, and what people want when it comes to privacy online.

Our survey is the first comprehensive examination of the intersection of knowledge and opinions on topics including online privacy, law, and the data practices of both the private sector and the government. Our survey results indicate that consumers often want more options than the market gives them. Over 80% of our survey participants, for example, indicated that on some occasion they have submitted information online when they wished that they did not have to do so. One of the possible reasons why consumers still participate in these markets is that they do not have any meaningful alternatives. The private sector currently has very little incentive to provide these alternatives because consumers have been responding to unattractive business practices with complacency. Responses to our survey also indicate a low level of trust of the government as a data collector and data keeper.

Our results indicate that significant changes are needed to increase consumer engagement in the online marketplace and improve trust between the government and its citizens. These improvements should begin by empowering users and giving them more control over their digital data, and we present ambitious proposals to this end. The long-term solution that we propose would involve an overhaul of current data privacy laws and the creation of a centralized profile repository that would serve a purpose similar to credit reporting bureaus. Through this repository, consumers could view and challenge most information that private businesses and the government hold about them. Dramatic changes are necessary in order to ensure that consumers can have empowering and engaging experiences in today's world of digital data.

---

* Jay P. Kesan, Ph.D., J.D., Professor and H. Ross & Helen Workman Research Scholar, University of Illinois College of Law.

** Carol M. Hayes, J.D., Research Associate, University of Illinois College of Law.

*** Masooda N. Bashir, Ph.D., Assistant Professor, University of Illinois Graduate School of Library and Information Science.

The authors also wish to thank Boyi Guo, Alejandro Gutierrez, Kevin Hoff, and Gahyun Jeon for their assistance with the survey and data. Some aspects of this research were presented at the Workshop on the Future of Privacy Notice and Choice at Carnegie Mellon and at the Telecommunications Policy Research Conference in Arlington, Virginia.

INTRODUCTION

Twenty years ago, the Internet was a technology at the edge of relevance.[1] Since then, it has transformed virtually every aspect of American culture. Some older ways of doing things have become outdated. What do you talk about at a high school reunion when you have been informed about all of your classmates' day-to-day lives via Facebook? Who needs to read restaurant reviews in the newspaper when diners can read the opinions that local residents dished out on Yelp? Why would you wait in line for eight hours to get tickets to a concert when you can choose your seats from the venue's website and pay Ticketmaster a nominal convenience fee?

The Internet has had an immeasurable effect on the economy. Some sectors, like print journalism, have experienced significant decline, but other sectors have exploded. With the help of Big Data and data brokers, online marketing goes far beyond direct mail techniques from the late twentieth century. Data brokers are described by the Federal Trade Commission (FTC) as "companies that collect consumers' personal information and resell or share that information with others."[2] This practice has significant application to behavioral marketing, as it allows for more information to be used to target potential customers. General Motors used an early form of behavioral marketing in the 1920s by sending surveys to potential customers to find out what the customers wanted.[3] Today, behavioral marketing on

---

1. *See* Chris Higgins, *What the Internet Looked Like in 1995*, MENTAL FLOSS (Mar. 26, 2013, 1:07 PM), http://mentalfloss.com/article/49676/what-internet-looked-1995 [http://perma.cc/2JNE-XSCC].

2. FED. TRADE COMM'N, DATA BROKERS: A CALL FOR TRANSPARENCY AND ACCOUNTABILITY, at i (2014), *available at* http://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf [http://perma.cc/YS4U-2783].

3. *See* DANIEL J. SOLOVE, THE DIGITAL PERSON: TECHNOLOGY AND PRIVACY IN THE INFORMATION AGE 16–17 (2004).

the Internet allows advertisers to buy access to individuals who match a particular consumer profile.[4] This can lead to more successful marketing campaigns and more sales.

The field of Big Data analytics represents a technological and social shift that is rapidly morphing society. Richards and King refer to this "Big Data Revolution" as just the latest stage of the Information Revolution.[5] Data and metadata swirl around us, being stored in greater and greater volumes. The information that is being collected can be used to predict traffic patterns, detect fraud, and stop terrorist attacks before they can happen by taking situational awareness to a superhuman level.[6] The data also enable more mundane predictions. For example, a retailer can identify pregnant shoppers to a reasonable degree of certainty based on purchase data.[7] But these increased capabilities are accompanied by a decrease in the level of control that consumers can exercise over information about themselves, and some have argued that surveillance by private actors is worrisome in a way similar to surveillance by government actors.[8]

The Internet has also transformed government activities and proposals. These effects range from the relatively innocuous, such as how federal agencies communicate and store information,[9] to the controversial, like the widespread surveillance activities of the National Security Agency (NSA) as revealed by Edward Snowden in 2013.[10] Partly in response to public concern over government surveillance, Apple and Google announced in late 2014 that their future products will, by default, use extremely strong encryption that even the companies themselves could not bypass.[11] Many politicians and members of law enforcement have spoken out against this. In January 2015, United Kingdom Prime Minister David Cameron argued that there should always be a way for the government to access even the most securely encrypted information.[12] The U.S. Congress has held multiple hearings to

---

4. Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and a New Concept of Personally Identifiable Information*, 86 N.Y.U. L. REV. 1814, 1851–52 (2011); *see* Paul Lanois, *Caught in the Clouds: The Web 2.0, Cloud Computing, and Privacy?*, 9 NW. J. TECH. & INTELL. PROP. 29, 34 (2010) (noting that user profiles are bought and sold on exchanges that resemble the stock market).

5. Neil M. Richards and Jonathan H. King, *Big Data Ethics*, 49 WAKE FOREST L. REV. 393, 397 (2014).

6. *See id.* at 405–06.

7. *See* Charles Duhigg, *How Companies Learn Your Secrets*, N.Y. TIMES MAG., Feb. 16, 2012, http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?_r=0 [http://perma.cc/6AZ2-27W4].

8. *See, e.g.*, Alexander Tsesis, *The Right to Erasure: Privacy, Data Brokers, and the Indefinite Retention of Data*, 49 WAKE FOREST L. REV. 433, 440 (2014).

9. Federal Information Security Modernization Act of 2014, Pub. L. No. 113-283, 128 Stat. 3073 (to be codified at 44 U.S.C. §§ 3551–58).

10. *See generally* Ewen MacAskill & Gabriel Dance, *NSA Files: Decoded: What the Revelations Mean for You*, GUARDIAN (Nov. 1, 2013), http://www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded [http://perma.cc/KN34-PB5D].

11. Joe Miller, *Google and Apple to Introduce Default Encryption*, BBC (Sept. 19, 2014), http://www.bbc.com/news/technology-29276955 [http://perma.cc/GCW2-3PW6].

12. *See* Alex Hern, *How Has David Cameron Caused a Storm over Encryption?*,

address arguments from law enforcement that unbreakable encryption would lead to more crime and endanger children.[13]

These arguments over unbreakable encryption may cause some computer security professionals to experience déjà vu. In the early 1990s, American policymakers were considering how to ensure that law enforcement would continue to have access to encrypted communications, and proposed a voluntary encryption standard that would leave the encryption key in "escrow" with the United States government.[14] This newest encryption controversy shows that, twenty years later, governments are still trying to figure out how to access more information for potential investigations in an age where criminal activities are becoming more and more digitized. The debates, however, are improving as computer scientists and technologists gain a voice in the political process. The chairman of the previously referenced subcommittee, for example, is freshman Congressman Will Hurd (R-TX), who has a degree in computer science and who worked with a cybersecurity firm before being elected to Congress.[15]

People often protest perceived government overreach and mass surveillance, but data brokers collect the same information in the interest of commerce.[16] Meanwhile, in the realm of cybersecurity, there is a need for increased information sharing between the government and the private sector about potential threats to cyber infrastructure.[17] Internet users are caught in the middle and are often called upon to make decisions with inadequate information. Ubiquitous privacy policies and terms-of-service (TOS) agreements supposedly establish the rights of users vis-à-vis the companies that they are dealing with, but these agreements typically do not regulate the use of the information by third parties like data brokers.[18] Additionally, the contractual

---

GUARDIAN (Jan. 15, 2015, 3:26 PM), http://www.theguardian.com/technology/2015/jan/15/david-cameron-encryption-anti-terror-laws [http://perma.cc/7MH7-5LF4].

13. *Going Dark: Encryption, Technology, and the Balance Between Public Safety and Privacy: Hearing Before the S. Comm. on the Judiciary*, 114th Cong. (2015), *available at* http://www.judiciary.senate.gov/meetings/going-dark-encryption-technology-and-the-balance-between-public-safety-and-privacy [http://perma.cc/Q23A-2NTS]; *Encryption Technology and Potential U.S. Policy Responses: Hearing Before the H. Subcomm. on Info. Tech. of the Comm. on Oversight and Gov't Reform*, 114th Cong. (2015), *available at* https://oversight.house.gov/hearing/encryption-technology-and-potential-u-s-policy-responses/ [https://perma.cc/XV22-VAXA].

14. *See* A. Michael Froomkin, *It Came from Planet Clipper: The Battle over Cryptographic Key "Escrow,"* 1996 U. CHI. LEGAL F. 15, 24–29.

15. *Biography*, CONGRESSMAN WILL HURD, https://hurd.house.gov/about/full-biography [https://perma.cc/N6XM-VVAW].

16. *60 Minutes: The Data Brokers: Selling Your Personal Information* (CBS television broadcast Mar. 9, 2014), *available at* http://www.cbsnews.com/news/the-data-brokers-selling-your-personal-information/ [http://perma.cc/PCV3-H98V].

17. Jay P. Kesan & Carol M. Hayes, *Creating a "Circle of Trust" To Further Digital Privacy and Cybersecurity Goals*, 2014 MICH. ST. L. REV. 1475; *see also Cybersecurity: Tech Firms Urged To Share Data with US*, BBC (Feb. 13, 2015), http://www.bbc.com/news/technology-31440978 [http://perma.cc/WV47-NRD9].

18. Jay P. Kesan, Carol M. Hayes & Masooda N. Bashir, *Information Privacy and Data Control in Cloud Computing: Consumers, Privacy Preferences, and Market Efficiency*, 70 WASH. & LEE L. REV. 341, 457–58 (2013).

language may also impact the users' legal rights under the Constitution and existing privacy statutes.

In the aftermath of the Snowden leaks and during the dawn of greater public awareness of the data broker industry, we conducted a comprehensive study to gather information about what people know, what people do, and what people want when it comes to privacy online. For this study, we collected data using a two-part survey. The first part of the survey explored participants' opinions of topics spanning four categories: (1) online behavior, (2) personal privacy, (3) cloud service providers, and (4) government data collection. The second part of the survey examined the knowledge levels of participants in five areas: (1) economics of the Internet, (2) cloud computing, (3) online security, (4) education records privacy law, and (5) other privacy law. Results from the knowledge survey were scored based on questions answered correctly. 455 people completed the full knowledge survey, and 534 people completed the full opinion survey. We considered partial opinion surveys when examining responses to specific opinion questions, so many questions had over 700 data points, but we did not consider partial knowledge surveys. 210 people completed the opinion survey and the full knowledge survey. This study is the first empirical examination of the intersection of knowledge and opinions about privacy, law, and the data collection activities of the private sector and the government.

The results of our survey will be discussed in Part III. Through examination of these results, some themes begin to emerge. One of the themes is the sense of helplessness that many consumers feel with regard to agreements that they must accept in order to use a service. 81% of our survey participants (n = 727) reported that, on some occasion, they had submitted information online when they wished that they did not have to do so. One of the self-reported reasons that participants gave for not reading privacy policies was that it would not make a difference whether they read the policy or not. When asked whether they approve or disapprove of many common data collection and data use practices among marketing companies online, many of our survey participants indicated strong disapproval—and yet they have not unplugged from this invasive data culture. These findings complement earlier research into consumer behavior[19] and information disclosure policy.[20] As a policy matter, lawmakers should consider if simply mandating that consumers be given information is sufficient to ensure consumer protection,[21] especially in light of the evidence that consumers often do not seem to be making a meaningful choice when agreeing to a website's terms and submitting information online.

Similarly alarming themes emerge when the issue is the collection of information by the government. The Obama administration has emphasized the need for more information sharing between the government and the private sector in order to combat cybersecurity threats,[22] but many of our survey participants were very skeptical of government data collection. Participants who performed better on some

---

19. *E.g.*, Rainer Böhme & Stefan Köpsell, *Trained to Accept? A Field Experiment on Consent Dialogs*, 2010 PROC. SIGCHI CONF. ON HUM. FACTORS COMPUTING SYSTEMS 2403, *available at* http://dl.acm.org/citation.cfm?id=1753689 [http://perma.cc/BL4F-2XWM].

20. *E.g.*, Omri Ben-Shahar & Carl E. Schneider, *The Failure of Mandated Disclosure*, 159 U. PA. L. REV. 647 (2011).

21. *Id.*

22. *Cybersecurity: Tech Firms Urged To Share Data with US*, *supra* note 17.

knowledge sections also indicated a lower level of trust of the government as a data collector and data keeper. Thus, another major issue which policymakers should confront is the lack of trust between the private sector and the government, especially among consumers who are more informed about information technology issues and therefore are more likely to influence their peers' technology-related decisions. Simple requests for information sharing are not enough to improve the relationship between the government and the private sector on this. Policymakers must take steps to actually increase how much the private sector trusts the government to protect individual information.

Our recommendations to address the problems uncovered by our survey include an ambitious and far-reaching proposal. The notice and choice paradigm has largely failed consumers, as the data trade continues to commoditize personal information with no oversight, and the government sows distrust among the population by engaging in mass surveillance. To overhaul the notice and choice paradigm, we suggest the creation of a profile repository to provide a centralized location for consumers to view the nonsensitive information that data brokers and the government hold about them, while also giving consumers the option to challenge or remove some elements of their profiles. These repositories are what we have termed Profile Information Reporting Agencies (PIRAs). PIRAs could be modeled after credit reporting agencies. There would be some information held by the government, such as information about active investigations, which consumers should not be able to view and challenge. However, to ensure transparency, most of the profile-relevant information held by both the private sector and the government should be viewable by the affected consumer.

During the transitional period, we suggest the adoption of practices to increase knowledge and increase consumer engagement. Part of our recommendation involves the use of a privacy agent that is adaptable based on the consumer's level of knowledge. This additional privacy agent would aid in the transition between the current paradigm, where consumers know very little, and our proposed paradigm, where consumers are not only given the right kind of information but are also empowered to view, update, and challenge data that they do not want in their profiles. We also encourage the use of educational programs to increase consumer competence in the online marketplace. Increasing public education and introducing adaptive privacy agents would not completely solve the problems we uncovered, but these steps would greatly assist during the transition to a new data privacy paradigm.

In Part I, we examine issues relating to online privacy, from theory to law, that provided a theoretical foundation for our survey. In the United States, ideas about privacy have evolved on the heels of technological developments, and changes in the law have often followed these changes of ideas. In Part II, we discuss interdisciplinary privacy literature and empirical work. In Part III, we present the results of our surveys examining knowledge, behavior, and opinions about online privacy issues. In Part IV, we set forth our recommendations for improving the status of consumers who find themselves adrift in this giant sea of data without a paddle or a compass.

I. PRIVACY THEORY AND PRIVACY LAW IN THE INFORMATION AGE

In this Part, we discuss the legal and theoretical backgrounds that underlie our survey. We conducted this survey in part because of the way that the Internet has been changing how people view privacy, the law, and even social interactions.[23] It is currently unclear how reasonable the "reasonable expectation of privacy" of Fourth Amendment jurisprudence is in this climate of data brokers and mass surveillance.[24] Our survey was designed to provide insight into consumer opinions and knowledge about this topic, as well as other issues like the patchwork design of federal privacy laws and the importance of information privacy. In the United States, privacy theory and law have evolved with technology. Until recently, the government and private sector had technological limitations on how much information they could collect about individuals and how long they could store this information. As the technological restrictions become less important, the theories and laws that shape privacy should be examined and updated to keep pace with what individuals want. The findings from our research provide more information to shape the future of privacy theory and law.

A survey by the Pew Research Center found that, as of January 2014, 87% of American adults are Internet users.[25] Research by the Kaiser Family Foundation suggests that the average youth between the ages of 8 and 18 spends every permissible waking moment using electronic devices, many of which (like smart phones and computers) are connected to the Internet.[26] A 2010 study by the Nielsen Company found that the average American Internet user is online over 55 hours per month.[27] Today, cyberspace is a major social outlet that is often intertwined with the physical realm.[28] People keep in touch through a variety of electronic messaging technologies, including e-mail, text messaging, other instant messaging over the Internet, and social networking websites.[29] So much information is flowing over the networks, and there are so many possible uses for that information other than simple interpersonal communications.

Data breaches and identity theft are arguably the most visible reasons to protect information privacy. As the technology has evolved, the barriers to identity theft have weakened. The Identity Theft Resource Center identified 783 data breaches in the

---

23. *See* Katherine J. Strandburg, *Home, Home on the Web and Other Fourth Amendment Implications of Technosocial Change*, 70 MD. L. REV. 614, 626 (2011) ("It is not just that 'the Internet is different'; it is that the Internet, like every major advance in infrastructural technology before it, has made everything different.").

24. Kesan et al., *supra* note 18, at 409–10 (2013).

25. *Internet Use Over Time*, PEW RESEARCH CTR., http://www.pewinternet.org/data-trend/internet-use/internet-use-over-time/ [http://perma.cc/MD6T-NQL9].

26. Andrea Cascia, *Don't Lose Your Head in the Cloud: Cloud Computing and Directed Marketing Raise Student Privacy Issues in K-12 Schools*, 261 EDUC. L. REP. 883, 894 (2011).

27. Lanois, *supra* note 4, at 29. About half of that time is spent on social networking, e-mails, games, and instant messaging. *Id.*

28. Strandburg, *supra* note 23, at 639.

29. John Soma, Melodi Mosley Gates & Michael Smith, *Bit-Wise but Privacy Foolish: Smarter E-Messaging Technologies Call for a Return to Core Privacy Principles*, 20 ALB. L.J. SCI. & TECH. 487, 497–502 (2010); *see also* Strandburg, *supra* note 23, at 655–56.

United States in 2014, surpassing the previous record of 662 breaches in 2010.[30] According to the Bureau of Justice Statistics, financial losses resulting from identity theft totaled $24.7 billion in 2012, an amount $10 billion higher than the financial losses from all other property crimes combined.[31]

What holders of consumer information, like data brokers, do with that information is also a threat to information privacy, and this threat is exacerbated by a lack of transparency. The FTC criticizes the lack of transparency of companies that trade in consumer information.[32] The Federal Communications Commission (FCC) has similarly emphasized the importance of transparency for consumer protection in the telecommunications context.[33] Our survey results indicate that many consumers do not trust many of the players in the private markets for Internet services and advertising. This call for greater transparency applies to government activities as well as the private sector. When President Obama took office in 2009, one of his early actions was the issuance of a memorandum about the need for transparency in government.[34] However, following several high-profile leaks of classified information, and in light of some of our findings from this survey, it is clear that the transparency priority should be revisited and operationalized in a way that will facilitate greater trust between the government and the private sector. We posit that increasing transparency could improve the amount that individuals trust online businesses and the government.

We conducted our survey during a tumultuous time in the history of data privacy in the United States, and our hope is that our findings will prove helpful for those who want to empower consumers and promote understanding between the government and its citizens. We begin this ambitious journey by exploring the idea of privacy by briefly retracing the steps of privacy theory development.

A. Privacy Theory

In our survey, one of the topics that we explored was participants' views of privacy as an idea. In academic circles, some characterize privacy as the ability to control data about oneself, while others argue that privacy rights threaten the common good.[35] Julie Cohen says that the value of privacy is its contribution to self-determination, allowing individuals to resist efforts by private companies and governments to reduce them to

---

30. *Identity Theft Resource Center Breach Report Hits Record High in 2014*, IDENTITY THEFT RES. CTR. (Jan. 12, 2015), http://www.idtheftcenter.org/ITRC-Surveys-Studies/2014databreaches.html [http://perma.cc/B5RY-BCPD].

31. *16.6 Million People Experienced Identity Theft in 2012*, BUREAU OF JUSTICE STATISTICS (Dec. 12, 2013), http://www.bjs.gov/content/pub/press/vit12pr.cfm [http://perma.cc/73S6-DFVR].

32. *See* FED. TRADE COMM'N, *supra* note 2, at 3.

33. Kevin Werbach, *The Network Utility*, 60 DUKE L.J. 1761, 1837 (2011).

34. Memorandum on Transparency and Open Government, 2009 DAILY COMP. PRES. DOC. 13 (Jan. 21, 2009).

35. Anita L. Allen, *Privacy-as-Data Control: Conceptual, Practical, and Moral Limits of the Paradigm*, 32 CONN. L. REV. 861, 861–62 (2000) (citing Paul M. Schwartz, *Internet Privacy and the State*, 32 CONN. L. REV. 815 (2000)).

nameless, predictable data points.[36] By allowing survey participants to weigh in on these questions, we hope to contribute additional insights that will be valuable for future privacy discussions.

The concept of privacy has evolved as mass media and technological developments have made private information more accessible. The first major push in privacy theory in the United States is often attributed to Samuel Warren and Louis Brandeis and their seminal work on the subject, *The Right to Privacy*.[37] Warren and Brandeis were especially concerned about the use of private information by the media and the implications of technological developments like new and cheaper photography technologies.[38] Warren and Brandeis supported the idea of applying the common law to protect a right to privacy, which they famously summarized as a "'right to be let alone.'"[39] By the time *The Right to Privacy* was fifty years old, privacy had become a new but minor doctrine in tort law. Only twelve states recognized the right of privacy by common law, and only two recognized it by statute.[40]

For modern privacy scholars, the next major development in privacy theory in the United States was the 1960 publication of William Prosser's article, *Privacy*.[41] Prosser's article is credited with laying the groundwork for the four main privacy torts: intrusion into private affairs; public disclosure of private facts; false light; and appropriation.[42] Like the earlier article by Warren and Brandeis, this evolution in privacy theory occurred in the context of technological improvements that extended the reach of mass media, including television broadcasting.

As the Internet has become more ubiquitous, information privacy concerns have increased. Alan Westin, an early information privacy scholar, defined privacy as "the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others."[43] The work of information privacy scholars holds the most promise for a sustainable theoretical approach to online privacy.

Today, we live in a new technological age, and thus privacy theory should move into a third stage of its development. Prosser's privacy torts are still used by courts, but they fail to address all of the harms that might arise from the use of

---

36. Julie E. Cohen, *What Privacy Is For*, 126 HARV. L. REV. 1904, 1905 (2013).

37. Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193 (1890).

38. Neil M. Richards, *Reconciling Data Privacy and the First Amendment*, 52 UCLA L. REV. 1149, 1198 (2005); Schwartz & Solove, *supra* note 4, at 1819. This position leads to a balancing of the interest in privacy against interests under the freedom of the press clause of the First Amendment. Neil M. Richards & Daniel J. Solove, *Prosser's Privacy Law: A Mixed Legacy*, 98 CALIF. L. REV. 1887, 1892 (2010); Daniel J. Solove, *Identity Theft, Privacy, and the Architecture of Vulnerability*, 54 HASTINGS L.J. 1227, 1229 (2003).

39. Richards & Solove, *supra* note 38, at 1891 (quoting Warren & Brandeis, *supra* note 37, at 213).

40. *Id.* at 1895.

41. William L. Prosser, *Privacy*, 48 CALIF. L. REV. 383 (1960); *see* Paul M. Schwartz & Karl-Nikolaus Peifer, *Prosser's* Privacy *and the German Right of Personality: Are Four Privacy Torts Better than One Unitary Concept?*, 98 CALIF. L. REV. 1925, 1926 (2010).

42. SOLOVE, *supra* note 3, at 58.

43. ALAN F. WESTIN, PRIVACY AND FREEDOM 7 (1967).

information on the Internet.[44] For example, the privacy tort of invasion typically requires the invasion to be of an offensive nature, but a lot of information collection appears largely innocuous.[45] In *Shibley v. Time, Inc.*, the litigation concerned the magazines' sales of their respective subscriber information to advertisers, but these sales were found to not meet the injury requirements for invasion of privacy.[46] Even when a company has experienced a security breach resulting in the loss of employee or customer data, federal courts have been hesitant to conclude that the risk of future identity theft was enough to establish actual harm to confer standing.[47] In our survey we found that being able to control access to even nonsecret personal data was important to a majority of survey participants. Using the data collected in our survey, we attempt to extrapolate the types of harms that survey participants might consider to flow from privacy violations. By identifying the types of harms that might be driving consumers to value privacy, we are attempting to provide empirical support for future developments in privacy theory.

Privacy theorists have frequently attempted to enunciate the actual harms caused by a violation of information privacy expectations. Daniel Solove asserts that harms deriving from privacy violations include dignitary harms, like injuries to reputation, and "architectural" harms, such as an increased risk that a particular harm will occur in the future or the "chilling effect" from law enforcement having too much power over individual expression.[48] Anita Allen notes that privacy can be viewed as essential to protect the values of independence and freedom, and that reduction of privacy is a threat to liberalism.[49] Paul Schwartz suggests that the constitutive nature of information privacy makes the erosion of information privacy a threat to civil society.[50] Julie Cohen argues that violations of privacy can undermine the development of the self.[51]

These abstract ideas of what makes information privacy so important are very academic in nature. By using our survey results to identify the types of harms that might be driving consumers to value privacy, we are attempting to provide empirical support for future developments in privacy theory. As this brief history shows, privacy theories will continue to evolve with technology. Privacy laws must do the same.

---

44. Kesan et al., *supra* note 18, at 382–84.

45. Richards & Solove, *supra* note 38, at 1919.

46. 341 N.E.2d 337, 339 (Ohio Ct. App. 1975); Richards & Solove, *supra* note 38, at 1919.

47. *See* Reilly v. Ceridian Corp., 664 F.3d 38, 42 (3d Cir. 2011), *cert. denied*, 132 S. Ct. 2395 (2012); *but see* Krottner v. Starbucks Corp., 628 F.3d 1139, 1140 (9th Cir. 2010).

48. Daniel J. Solove, *A Taxonomy of Privacy*, 154 U. PA. L. REV. 477, 487–89 (2006).

49. *See* Anita L. Allen, *Coercing Privacy*, 40 WM. & MARY L. REV. 723, 734, 755 (1999).

50. *See* Paul M. Schwartz, *Privacy and Democracy in Cyberspace*, 52 VAND. L. REV. 1609, 1613–14 (1999).

51. *See* Cohen, *supra* note 36, at 1911.

B. Privacy Law

One of the goals of our research is to evaluate how consumers' expectations align with existing laws directed at information privacy concerns. We explored participants' knowledge of some basic topics relating to federal data privacy laws and the Constitution. In the United States, such laws are often a patchwork and can vary from state to state.[52] For example, states vary in how they treat security breaches that result in data theft.[53] Minnesota has a merchant liability statute that allows merchants to be sued if they do not sufficiently protect credit card information and that information is subsequently stolen.[54] On the other hand, many other states merely require companies to notify customers of data breaches, and the relevant statutes do not create any additional duties or entitlements.[55]

1. Federal Laws

Because of the interstate issues raised by Internet communications and the existence of inconsistent state laws, a unified federal approach to data privacy is desirable. Federal privacy law is currently a patchwork of fixes to different issues. Federal statutes often focus on the presence of personally identifiable information (PII), though they often vary in what sort of information is considered "personally identifiable."[56]

While there is not an explicit clause in the U.S. Constitution that states the existence of a general right to privacy, the Supreme Court held in *Griswold v. Connecticut* that privacy was part of the penumbra of rights surrounding the enumerated rights of the Constitution.[57] Much discussion of Supreme Court privacy jurisprudence focuses on decisional privacy; that is, the right of individuals to make decisions free of government intervention.[58] In *Whalen v. Roe*, however, the Supreme Court recognized "the individual interest in avoiding disclosure of personal

---

52. For a discussion of state laws, see Kimberly L. Rhodes & Brian Kunis, *Walking the Wire in the Wireless World: Legal and Policy Implications of Mobile Computing*, 16 J. TECH. L. & POL'Y 25, 50–51 (2011), and Schwartz & Solove, *supra* note 4, at 1831.

53. *See Security Breach Notification Laws*, NAT'L CONF. ST. LEGISLATURES (Oct. 12, 2015), http://www.ncsl.org/research/telecommunications-and-information-technology/security-breach-notification-laws.aspx [http://perma.cc/FJC9-SKL2].

54. Rhodes & Kunis, *supra* note 52, at 50.

55. *See*, *e.g.*, Cooney v. Chi. Pub. Sch., 943 N.E.2d 23, 28 (Ill. App. Ct. 2010) (holding that Illinois's Personal Information Protection Act does not impose upon data holders a duty of care with regard to personal information like social security numbers).

56. *See* Schwartz & Solove, *supra* note 4, at 1827. Statutes concerned with PII include the Cable Communications Policy Act, the Video Privacy Protection Act, and the Gramm-Leach-Bliley Act. *Id.* at 1824, 1829–30.

57. 381 U.S. 479, 484 (1965).

58. Some of the best known examples of decisional privacy cases in the Supreme Court over the last fifty years concern contraception and abortion. *E.g.*, *Griswold*, 381 U.S. at 479 (ruling unconstitutional a Connecticut law that prohibited the use of contraception); Roe v. Wade, 410 U.S. 113 (1973) (citing a right to privacy under the Constitution in prohibiting states from outright banning abortion).

matters,"[59] which has influenced many lower courts in recognizing a constitutional right to information privacy.[60]

a. Sector-Specific Federal Laws

Current federal privacy statutes largely address privacy issues piecemeal, each focusing on a limited category of privacy concerns.[61] The online privacy interests of children under the age of thirteen are protected by the Children's Online Privacy Protection Act (COPPA).[62] Information privacy concerns for health records are addressed by the Health Insurance Portability and Accountability Act (HIPAA).[63] The Family Educational Rights and Privacy Act (FERPA) addresses information privacy in educational institutions.[64] The information privacy of some types of financial records is addressed by the Gramm-Leach-Bliley Act (GLBA),[65] and other types of financial information relevant to consumer credit are addressed in the Fair Credit Reporting Act (FCRA).[66] In our survey, we addressed FERPA as an example of a sector-specific federal privacy law.

Federal statutes addressing information privacy in the private sector are a tangled web, and a unified approach would benefit millions of citizens. Even though existing federal privacy laws are often narrow, some of the sector-specific laws could nonetheless be instructive in considering how a broader data privacy law should or should not be structured. The GLBA, for instance, represents the current failing paradigm of "notice and choice" in that it permits financial institutions to share their customers' nonpublic personal information with the institutions' affiliates but the customers must first be told and have the ability to opt out of such sharing.[67] The GLBA also requires financial entities to provide customers with annual privacy notices.[68]

FERPA, on the other hand, provides much stronger default protections. FERPA was enacted in 1974 and applies to all levels of education that receive federal funding.[69] "Education records" under FERPA include a variety of information held by the school, not just grades. FERPA defines "education records" as including materials which "contain information directly related to a student" and which "are

59. 429 U.S. 589, 599 (1977).

60. Solove, *supra* note 48, at 558.

61. SOLOVE, *supra* note 3, at 71 ("Thus, the federal privacy statutes form a complicated patchwork of regulation with significant gaps and omissions."); *see also* Michael Birnhack & Niva Elkin-Koren, *Does Law Matter Online? Empirical Evidence on Privacy Law Compliance*, 17 MICH. TELECOMM. & TECH. L. REV. 337, 349 (2011).

62. 15 U.S.C. §§ 6501–6506 (2012).

63. Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191, 110 Stat. 1936 (codified as amended in various sections of Titles 26, 29, and 42 of the U.S. Code).

64. 20 U.S.C.A. § 1232g (West 2010 & Supp. 2015).

65. 15 U.S.C. §§ 6801–6809 (2012).

66. 15 U.S.C. § 1681b (2012).

67. SOLOVE, *supra* note 3, at 70–71.

68. Paul M. Schwartz, *Property, Privacy, and Personal Data*, 117 HARV. L. REV. 2056, 2101 (2004).

69. 20 U.S.C.A. § 1232g.

maintained by an educational agency or institution or by a person acting for such agency or institution."[70] FERPA, like many statutes about privacy, focuses on PII.[71]

FERPA holds promise for the future of privacy law because it protects privacy in two ways: (1) by ensuring that students and/or their parents have access to the student's education records and are provided with a way to challenge the contents of the records,[72] and (2) by prohibiting the educational institution from sharing students' education records with other people unless an exception applies or consent is obtained.[73] Unlike the GLBA, FERPA uses an "opt-in" model for information sharing, enhanced with statutory exceptions. For example, before a student turns 18, the parent has broad access to these records, but parents of students who are 18 or older can only obtain the student's education records without the student's consent if the student is considered to be the parent's dependent for tax purposes.[74] FERPA regulations also set forth more specific guidelines for when consent must be obtained prior to the disclosure of information contained in education records.[75]

A broader legislative privacy regime should embrace the dual focus of FERPA, providing consumers with a right to access and challenge compiled information, with a default opt-in model for information sharing and limited enumerated exceptions. A lot of privacy statutes currently focus on allowing consumers to opt out of data collection and sharing, and it is rare for a privacy statute to give the subject of the information this degree of access and control. The results of our study indicate the presence of a general feeling of helplessness that consumers seem to experience in this environment of Big Data. Requiring data brokers and their business partners to adopt an opt-in approach to information sharing and allow consumers to access, challenge, and alter information in their profile could improve transparency and go a long way towards alleviating this sense of helplessness. Such a law could also encourage more participation in the online marketplace, leading to more benefits for both consumers and companies.

b. The Fourth Amendment and Surveillance

In our survey, we also investigated perceptions of privacy protections under the Constitution. The Fourth Amendment declares that people have a right "to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures."[76] In Fourth Amendment jurisprudence, the focus is often on the existence of a "reasonable expectation of privacy."[77] If the government conducts surveillance

70. *Id.* § 1232g(a)(4)(A).

71. *See* Steve Mutkoski, *Cloud Computing, Regulatory Compliance, and Student Privacy: A Guide for School Administrators and Legal Counsel*, 30 J. MARSHALL J. INFO. TECH. & PRIVACY L. 511, 520 (2014).

72. 20 U.S.C.A. § 1232g(a)(1)–(2).

73. *Id.* § 1232g(b).

74. *Id.* § 1232g(b)(1)(H).

75. 34 C.F.R. § 99.30 (2015).

76. U.S. CONST. amend. IV.

77. *See* Kesan et al., *supra* note 18, at 409.

somewhere that there is a reasonable expectation of privacy, a warrant is often necessary to protect against unreasonable intrusion.[78]

The third-party doctrine is the aspect of Fourth Amendment jurisprudence that has been the most controversial in the context of information privacy online. Under the third-party doctrine, a court is likely to conclude that there is no reasonable expectation of privacy in papers and effects that have been turned over to a third party.[79] Bailments of concealed items, however, might not be limited by the third-party doctrine, because the bailee entrusted with concealed items does not necessarily have the authority to view the items or to consent to the search of the items by others.[80] In the context of online services, the breadth of the third-party doctrine was limited by the Sixth Circuit in *United States v. Warshak*, with the court reasoning that the e-mail provider was not a recipient of the communication, but rather was an intermediary.[81] However, *Warshak* still leaves many questions unanswered. In *Warshak*, the Sixth Circuit concluded that the defendant had a reasonable expectation of privacy in his e-mail, but noted that there may have been a different conclusion if the defendant had consented to broad language in terms of what the e-mail provider was allowed to do with the information in his account.[82]

The third-party doctrine was central to the decision of a federal district court in *ACLU v. Clapper* in 2013, which concerned the controversial practice by the NSA of using section 215 of the Foreign Intelligence Surveillance Act to engage in bulk collection of telecommunications metadata.[83] This metadata includes the phone numbers the callers dialed, when they dialed the numbers, the call routing information, and how long the calls lasted. There, District Court Judge Pauley held that telecommunications network subscribers have "no subjective expectation of privacy in telephony metadata."[84] The Second Circuit reversed this decision in May of 2015, reasoning that the NSA's program was not authorized by the language of the statute.[85] The appellate court did not reach a conclusion as to the constitutionality of the program or section 215, but openly acknowledged the "daunting" constitutional issues relating to this type of surveillance and the application of the third-party doctrine in today's world.[86]

---

78. *See* Katz v. United States, 389 U.S. 347, 360–61 (1967) (Harlan, J., concurring).

79. *See* United States v. Miller, 425 U.S. 435 (1976) (finding no reasonable expectation of privacy in financial records disclosed to a financial institution in the ordinary course of business); Couch v. United States, 409 U.S. 322 (1973) (finding no reasonable expectation of privacy in financial records turned over to an accountant for tax return purposes).

80. United States v. James, 353 F.3d 606 (8th Cir. 2003) (overturning a child pornography conviction because the police obtained the evidence from a person entrusted with computer disks with specific instructions to not use the disks).

81. 631 F.3d 266, 288 (6th Cir. 2010).

82. *Id.* at 287.

83. ACLU v. Clapper, 959 F. Supp. 2d 724 (S.D.N.Y. 2013), *vacated*, 785 F.3d 787 (2d Cir. 2015), *and modified*, No. 14-42, 2015 WL 4196833 (2d Cir. June 9, 2015) (staying mandate due to amendments to section 215).

84. *Id.* at 752.

85. ACLU v. Clapper, 785 F.3d 787, 826 (2d Cir. 2015), *modified*, No. 14-42, 2015 WL 4196833 (2d. Cir. June 9, 2015) (staying mandate due to amendments to section 215).

86. *Id.* at 825.

c. Stored Communications Act

The Fourth Amendment is essential for protecting citizens against unreasonable searches by the government, but it does not apply to private entities if they are not state actors,[87] nor is it clear how the Fourth Amendment applies to newer technologies that go far beyond anything that the drafters of the Bill of Rights could have imagined. The Electronic Communications Privacy Act (ECPA) fills in some of these gaps, focusing on protecting the users of modern communications and computing technologies from information privacy infringement by the government or other members of the private sector. The ECPA consists of three federal statutes: the Stored Communications Act (SCA),[88] the Pen Register statute,[89] and the Wiretap Act.[90] The aspect of the ECPA that is the most relevant to our survey is the SCA, which sets forth the circumstances in which data holders can disclose customer information and the circumstances in which the government can compel data holders to disclose that same information. The section about compelled disclosure enumerates some types of data for which a warrant is necessary, while other types of data can be obtained with lesser evidentiary showings. In our knowledge survey, we asked participants some basic knowledge questions about the SCA, and we also asked participants some opinion-based questions on legal protections that should be afforded to different types of digital information.

In our survey, we discovered that most participants think that e-mail should be entitled to the same legal protections as physical mail, which generally cannot be searched without a warrant.[91] Under the SCA, however, the degree of privacy in an e-mail currently depends on whether it is stored on a hard drive or in the cloud.[92] The SCA only requires a warrant to obtain unopened e-mails less than 180 days old,[93] and its application to webmail is unclear because e-mails on webmail services always reside on the provider's servers.[94] The ability to access important communications from anywhere is a huge benefit for mobility and mobile computing, but under the SCA as currently written, e-mails that have been previously read and which are stored in the cloud can be obtained by the government without meeting the high "probable cause" standard for a search warrant.[95]

---

87. *See* Skinner v. Ry. Labor Execs.' Ass'n, 489 U.S. 602, 613–14 (1989) ("The [Fourth] Amendment guarantees the privacy, dignity, and security of persons against certain arbitrary and invasive acts by officers of the Government or those acting at their direction.")

88. 18 U.S.C.A. §§ 2701–2712 (West 2015).

89. 18 U.S.C. §§ 3121–3127 (2012). A "pen register" is a device that records phone numbers dialed, though the language of the statute also applies to other technological means. *Id.* § 3127.

90. 18 U.S.C. §§ 2510–2522 (2012); *see also* Andrew William Bagley, *Don't Be Evil: The Fourth Amendment in the Age of Google, National Security, and Digital Papers and Effects*, 21 ALB. L.J. SCI. & TECH. 153, 167 (2011).

91. *See infra* Part III.B.2.d.

92. Lanois, *supra* note 4, at 45.

93. 18 U.S.C. § 2703 (2012).

94. Bagley, *supra* note 90, at 167–68.

95. For more discussion of the SCA, see Orin S. Kerr, *A User's Guide to the Stored Communications Act, and a Legislator's Guide to Amending It*, 72 GEO. WASH. L. REV. 1208, 1216 (2004), and Kesan et al., *supra* note 18, 399–407.

Section 2703 of the SCA details the procedures that are required before the government can compel an information holder to disclose a third party's digitally stored information. The necessary showing varies based on what type of information is being sought. Depending on the type of information, the government may have to obtain a warrant, or the government may be required to provide notice that such information is being sought through means that require less than a showing of probable cause.[96] When a warrant is not required, some of the information may be compelled with a subpoena, while other types of information must be obtained via the issuance of an order under section 2703(d), which requires a stronger evidentiary showing than a subpoena but less than probable cause.

The constitutionality of 2703(d) orders as they pertain to geolocation data was recently challenged in *United States v. Davis*,[97] where the Eleventh Circuit confronted the question of how much protection historical cell tower location information is entitled to under the Fourth Amendment. In an en banc decision, the court invoked the third-party doctrine in concluding that a search of metadata records held by a third party was not a search for Fourth Amendment purposes, and that Davis lacked a reasonable expectation of privacy in this information.[98] The court further reasoned that even if Davis had a reasonable expectation of privacy, a warrant would not be necessary to compel production of this information. The 2703(d) order, the court asserted, bestowed the information with greater protections than it was entitled to under the Fourth Amendment, essentially raising the bar from a subpoena rather than lowering the bar from a warrant.[99] The SCA, however, was not enacted with the expectation that millions of people would soon be walking around with devices that tracked their locations at all times. The current state of the SCA as it applies to the Internet, cloud computing, and various types of metadata should be reevaluated by Congress to fill in gaps that were not anticipated in 1986 when the SCA was originally enacted. We did not address the privacy of geolocation data in our survey, but this topic is ripe for empirical study by future researchers.

2. Recent Case Law

In the last five years, three significant digital-data cases have been decided by the Supreme Court. The first, *City of Ontario v. Quon*, involved text messages on a government-provided device where investigators obtained the text messages from the service provider without a warrant.[100] The second, *United States v. Jones*, involved the placing of a GPS device on a suspect's vehicle without a warrant.[101] The third, *Riley v. California*, explored whether the search incident to arrest exception to the warrant requirement permitted the detailed examination of cell phone contents

---

96. 18 U.S.C. § 2703.

97. 785 F.3d 498 (11th Cir. 2015), *petition for cert. filed*, 84 U.S.L.W. 3081 (U.S. July 29, 2015) (No. 15-146).

98. *Id.* at 513.

99. *Id.* at 505–06.

100. 130 S. Ct. 2619 (2010).

101. 132 S. Ct. 945 (2012).

by police.[102] The search incident to arrest exception permits police to search items in the immediate vicinity of an arrestee, and *Riley* concerned the limits to this exception.[103]

These three cases reflect hesitation on the part of the Court to draw conclusions about digital privacy in the ether of the Internet. In *Quon*, the Court assumed, but did not conclusively determine, that text messages were entitled to a reasonable expectation of privacy.[104] In *Jones*, the majority did not delve into the digital-privacy issues of using GPS data to track a suspect's location, instead relying on a theory of trespass to find that placing a GPS device on a suspect's car violated the Fourth Amendment when no warrant had been obtained.[105] This narrow ruling in *Jones* leaves undecided the issue of whether citizens have a reasonable expectation of privacy in information about their locations when that information is captured by electronic devices.[106] In *Riley*, the Court unanimously declared that the content of cell phones could not be accessed as part of a search incident to arrest, but the Court limited its analysis to data stored on the cell phone.[107] The Court noted that modern cell phones can hold the equivalent of millions of lines of text and many hours of video.[108] The fact that there may be no line between locally stored information and cloud stored information also contributed to the Court's conclusion.[109] However, the Court did not address what level of process would be necessary for obtaining the contents of cloud storage except to acknowledge it as a complicating factor.

3. The FTC and Data Privacy

Currently, many data privacy issues are addressed by market self-regulation or light-touch regulation from federal agencies like the FTC. The FTC provides guidance and the occasional adjudication of disputes about how companies manage data. The FTC also promotes the adoption of Fair Information Practices (FIPs). FIPs address how to handle and use personal information,[110] often focusing on issues like data quality, transparency, special protections afforded to sensitive data, and enforcement standards.[111] The FTC views FIPs as being based on five core

---

102. 134 S. Ct. 2473 (2014).
103. *Id.* at 2483–84.
104. 130 S. Ct. at 2623.
105. 132 S. Ct. at 950.
106. If the Supreme Court had decided the *Jones* case based on a theory that there is a reasonable expectation of privacy in location information, this would have raised serious constitutional concerns about the use of geolocation data that law enforcement members obtain from service providers or through other means. Because the Supreme Court took a narrower approach, the Eleventh Circuit held in *United States v. Davis* that investigators were not required to obtain a warrant to obtain historical cell site data. 785 F.3d 498 (11th Cir. 2015), *petition for cert. filed*, 84 U.S.L.W. 3081 (U.S. July 29, 2015) (No. 15-146).
107. *See* 134 S. Ct. at 2485.
108. *Id.* at 2489.
109. *Id.* at 2491.
110. Solove, *supra* note 38, at 1266.
111. Richards, *supra* note 38, at 1167.

principles: (1) notice and consumer awareness; (2) consumer choice and consent; (3) access and participation in the process; (4) data integrity and security; and (5) enforcement and redress.[112]

The FTC also uses its adjudicatory powers to issue orders against organizations. When adjudicating these issues, the FTC's primary approach to protecting privacy online focuses on bringing complaints against companies that are not honest about privacy expectations in their statements to customers. TRUSTe, a provider of privacy certifications, recently entered into a settlement with the FTC. The FTC found that TRUSTe failed to conduct the annual recertifications as promised in over one thousand separate incidences between 2006 and 2013.[113] Because TRUSTe certification is an established signal that is used to indicate that a company prioritizes and protects consumer privacy, this is a potentially significant blow to consumer confidence in online interactions.

Other recent FTC decisions include a consent order issued against Epic Marketplace, Inc. and Epic Media Group, L.L.C., which engaged in "history sniffing" without notifying customers;[114] a complaint against Goldenshores Technologies, L.L.C. concerning the company's tracking of device data by a flashlight application without providing notice to consumers who install the app;[115] and a settlement with Snapchat, Inc. over the company's misleading representations that photo messages sent via the app could not be saved.[116] The FTC has also issued orders against Credit Karma and Fandango for distributing applications with security measures that did not comply with the companies' assertions about security in privacy policies.[117]

Recent adjudications show that the FTC prioritizes consumer privacy. However, the FTC is a very busy agency, and the biggest problem with relying on the FTC to protect consumer privacy is that the FTC often relies on the companies to establish their own standards for how to handle consumer information. There are some exceptions, such as the complaints against Goldenshores Technologies and Epic Marketplace, where the absence of a statement about collecting information was the determining factor. In other situations, like the complaints against Credit Karma and Fandango, a failure to follow the standards that the company said they followed was

---

112. FED. TRADE COMM'N, PRIVACY ONLINE: A REPORT TO CONGRESS 7 (1998).

113. Press Release, Fed. Trade Comm'n, TRUSTe Settles FTC Charges It Deceived Consumers Through Its Privacy Seal Program (Nov. 17, 2014), *available at* http://www.ftc.gov/news-events/press-releases/2014/11/truste-settles-ftc-charges-it-deceived-consumers-through-its [http://perma.cc/W38Y-58DE].

114. *In re* Epic Marketplace, Inc., 155 F.T.C. 406 (2013).

115. *In re* Goldenshores Techs., L.L.C., No. C-4446, 2014 WL 1493611 (F.T.C. Mar. 31, 2014).

116. Press Release, Fed. Trade Comm'n, FTC Approves Final Order Settling Charges Against Snapchat (Dec. 31, 2014), *available at* http://www.ftc.gov/news-events/press-releases/2014/12/ftc-approves-final-order-settling-charges-against-snapchat [http://perma.cc/P9BB-AKPK].

117. Press Release, Fed. Trade Comm'n, Fandango, Credit Karma Settle FTC Charges that They Deceived Consumers by Failing to Securely Transmit Sensitive Personal Information (Mar. 28, 2014), *available at* http://www.ftc.gov/news-events/press-releases/2014/03/fandango-credit-karma-settle-ftc-charges-they-deceived-consumers [http://perma.cc/AM8Q-XTWL].

the determining factor. For example, if a company states that information is encrypted using Secure Sockets Layer (SSL) technology, and information is not being encrypted in this specific way, that is a deceptive practice that the FTC can punish. Arguably, all a company really needs to do to avoid FTC liability is to make just enough disclosures about what information they collect and how, while providing general assurances for how information will be handled. Therefore, the current case-by-case approach to consumer privacy at the FTC creates a perverse incentive for generic privacy policies where the company makes few, if any, concrete assertions about its data practices. Clearer actions to protect consumers are necessary to reduce the trust deficit between individuals, online services, and government actors.

C. "Do You Agree to These Terms?"

One of the topics examined in our survey was the interaction that participants have with privacy policies and TOS agreements. These agreements are contracts that practically every Internet user must accept for each website that she uses. TOS agreements set forth terms governing the relationship between a service provider and its customers.[118] Generally, cloud-based services targeting individual users are accompanied by non-negotiable TOS agreements that favor the service provider over the end user.[119] Privacy policies often accompany TOS agreements, and by agreeing to a website's privacy policy, the user consents to the terms contained therein. These terms typically address what information will be collected and with whom that information will be shared.[120] Privacy policies often also address data security issues, like the use of SSL encryption during data transmission.[121]

Under the common law of contracts, forming a contract requires mutual assent. When a contract is not subject to negotiation and is offered by the more powerful party on a "take it or leave it" basis, the contract is often referred to as a contract of adhesion.[122] Privacy policies and TOS agreements typically meet the definition for adhesion contracts.[123] Such contracts are not automatically invalid, but they may be subject to greater scrutiny. If a court finds that the contract is unconscionable, it may

---

118. *See* Joshua A.T. Fairfield, *Nexus Crystals: Crystallizing Limits on Contractual Control of Virtual Worlds*, 38 WM. MITCHELL L. REV. 43, 44 (2011) (referring to terms of use and End User License Agreements (EULAs) as the "social contract of the new millennium," setting forth the rights and redresses of citizens).

119. Mark H. Wittow & Daniel J. Buller, *Cloud Computing: Emerging Legal Issues for Access to Data, Anywhere, Anytime*, J. INTERNET L., July 2010, at 7; Bagley, *supra* note 90, at 163 ("Google's profit model is based on offering free services to consumers in exchange for their consent to non-negotiable terms of service.").

120. Kesan et al., *supra* note 18, at 449–50.

121. *See* Konstantinos K. Stylianou, *An Evolutionary Study of Cloud Computing Services Privacy Terms*, 27 J. MARSHALL J. COMPUTER & INFO. L. 593, 603 (2010).

122. BLACK'S LAW DICTIONARY 390 (10th ed. 2014) (defining adhesion contract).

123. Solove, *supra* note 38, at 1234–35 (arguing that the idea that users give informed consent to these terms is a fiction, due to the total lack of negotiation).

be unenforceable.[124] Future research could examine unconscionability in the context of consumer helplessness in the online data market.

II. EMPIRICAL RESEARCH AT THE INTERSECTION OF LAW, SOCIAL SCIENCE, AND TECHNOLOGY

We conducted our survey to examine perceptions and knowledge relating to online privacy and the surrounding technologies, business practices, and laws that shape modern privacy norms. In the previous Part, we discussed the legal and theoretical contexts of privacy online and how our survey was shaped by these issues. In this Part, we will provide more details about the interdisciplinary work that has explored similar topics through the lenses of disciplines including psychology and engineering, and how our survey contributes to the literature.

Most prior research in this area addresses policymaking only tangentially, but some researchers have emphasized the potential for regulation. Turow et al. released a report in 2009 about the results from their survey examining opinions about tailored advertising.[125] In their survey, Turow and his team explored topics including participants' opinions about tailored content and possible laws to apply to these practices.[126] McDonald and Cranor also explored the policy implications of consumer perceptions and knowledge of tailored advertising in 2010.[127] The work of Huang and Bashir on global privacy norms also addresses the possibility of regulation to protect privacy as a human right.[128]

Over the last couple of decades, the study of human-computer interaction (HCI) has grown in the academic literature. HCI is an interdisciplinary field that combines psychology and other social sciences with technical fields including computer science.[129] Some research in the area of HCI focuses on perceptions of service trustworthiness, including the protection of user privacy[130] and trust in automation.[131] HCI research may also involve topics such as usability and design methodologies.[132]

---

124. *See, e.g.*, Bragg v. Linden Research, Inc., 487 F. Supp. 2d 593, 605 (E.D. Pa. 2007); People v. Network Assocs., Inc., 758 N.Y.S.2d 466 (N.Y. Sup. Ct. 2003).

125. Joseph Turow, Jennifer King, Chris Jay Hoofnagle, Amy Bleakley & Michael Hennessy, *Americans Reject Tailored Advertising and Three Activities That Enable It* (Sept. 29, 2009) (unpublished manuscript), *available at* http://ssrn.com/abstract=1478214 [http://perma.cc/Q7ZL-CTE2].

126. *Id.* at 11.

127. Aleecia M. McDonald & Lorrie Faith Cranor, *Beliefs and Behaviors: Internet Users' Understanding of Behavioral Advertising* (Aug. 16, 2010) (unpublished manuscript presented at Telecommunications Policy Research Conference, Sept. 2010), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1989092 [http://perma.cc/6GVK-CXUR].

128. Hsiao-Ying Huang & Masooda Bashir, *Is Privacy a Human Right? An Empirical Examination in a Global Context*, 2015 INT'L CONF. ON PRIVACY, SECURITY & TR. 77.

129. Gary M. Olson & Judith S. Olson, *Human-Computer Interaction: Psychological Aspects of the Human Use of Computing*, 54 ANN. REV. PSYCHOL. 491, 492 (2003).

130. *Id.* at 500.

131. *E.g.*, Kevin Anthony Hoff & Masooda Bashir, *Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust*, 57 HUM. FACTORS 407 (2015).

132. *See* Mark S. Ackerman & Scott D. Mainwaring, *Privacy Issues and Human-Computer Interaction*, *in* SECURITY AND USABILITY: DESIGNING SECURE SYSTEMS THAT PEOPLE CAN USE

Analysis of the effectiveness of user interfaces may be accomplished using several techniques, including surveys or laboratory settings.[133] When studies are in controlled laboratory settings, the researchers may simply use software to track clicks and note how much time elapses between clicks,[134] or they may use eye-tracking technology to determine what aspects of the interface draw the attention of users.[135] The researchers may also talk with the subjects as part of the data-gathering process in the lab.[136]

Some empirical research in HCI literature asserts the existence of a privacy paradox, where there is a disconnect between what people say they want in terms of privacy and what they actually do.[137] But there are many concepts relating to privacy, so simply stating that a paradox exists with regard to "privacy" is not necessarily helpful. Awad and Krishnan examined this issue by focusing on data transparency and found that people who desire data transparency are less likely to allow companies to profile them for personalized offerings.[138] Another element to evaluate is the user's trust of a service. In a study of privacy behaviors on Facebook, Xu, Wang, Wisniewski, and Grossklags found that participants who indicated a high level of trust in Facebook were more willing, compared to the low trust group, to share more information with Facebook apps by default.[139]

One of our reasons for examining knowledge in our survey was to determine whether being more knowledgeable about online activities is related to attitudes or behaviors concerning privacy. Research by Park, Campbell, and Kwak indicated a correlation between being knowledgeable about privacy issues and actually taking steps to protect privacy.[140] In a random survey of users of an instant messaging service, Paine, Reips, Stieger, Joinson, and Buchanan found that 56% of subjects

---

381 (Lorrie Faith Cranor & Simson Garfinkel eds., 2005).

133. *See* Jesper Kjeldskov & Connor Graham, A Review of Mobile HCI Research Methods (2003) (unpublished manuscript), *available at* http://www.researchgate.net/publication /221270890_A_Review_of_Mobile_HCI_Research_Methods [http://perma.cc/6BWX-FFAR].

134. *See* Lorrie Faith Cranor, Praveen Guduru & Manjula Arjula, *User Interfaces for Privacy Agents*, 13 ACM TRANSACTIONS ON COMPUTER-HUM. INTERACTION 135, 162 (2006).

135. *See* M. Oya Çinar, *Eye Tracking Method To Compare the Usability of University Websites: A Case Study*, *in* HUMAN CENTERED DESIGN: FIRST INTERNATIONAL CONFERENCE PROCEEDINGS 671 (Masaaki Kurosu ed., 2009).

136. *See* Cranor et al., *supra* note 134, at 162.

137. *See* Nathaniel Good, Rachna Dhamija, Jens Grossklags, David Thaw, Steven Aronowitz, Deirdre Mulligan & Joseph Konstan, *Stopping Spyware at the Gate: A User Study of Privacy, Notice and Spyware*, 2005 PROC. SYMP. ON USABLE PRIVACY & SECURITY 45; Yong Jin Park, Scott W. Campbell & Nojin Kwak, *Affect, Cognition and Reward: Predictors of Privacy Protection Online*, 28 COMPUTERS HUM. BEHAV. 1019, 1024 (2012).

138. Naveen Farag Awad & M.S. Krishnan, *The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness To Be Profiled Online for Personalization*, 30 MIS Q. 13 (2006).

139. Heng Xu, Na Wang, Pam Wisniewski & Jens Grossklags, Default and Context: Investigating Facebook Users' Privacy Perceptions and Behaviors of Installing Third-Party Apps (2014) (unpublished manuscript presented at Workshop on the Future of Privacy Notice and Choice, Carnegie Mellon Univ.), *available at* https://www.cylab.cmu.edu/news_events /events/fopnac/pdfs/xu.pdf https://perma.cc/4YJH-JDY4].

140. Park et al., *supra* note 137, at 1024.

reported that they were concerned about privacy online, and 73% said that they take privacy-protecting steps while using the Internet.[141] Our research expands on previous findings like these by comparing privacy behaviors to knowledge in other areas, like cloud computing and online security.

HCI literature on privacy sometimes examines privacy policies and how to make these policies more effective. A number of the possible solutions proposed are technological rather than regulatory in nature.[142] One possibility is to use privacy tools that set privacy preferences at the level of the user, and then evaluate those preferences against the actual privacy practices of websites.[143] Such tools could rely on self-reported preferences through the use of a survey or could approximate privacy preferences.[144] Lorrie Cranor has discussed and evaluated other standardized alternatives to text-based privacy policies, including privacy nutrition labels and icons.[145] Patil has explored the potential for using filters as a method of privacy preference specification.[146] In light of the growth of the Internet of Things, Jutla and Bodorik have also considered privacy agents that could assist users of these kinds of devices.[147]

It is generally accepted that people do not read TOS or privacy policies, which is understandable considering that most of these documents span several pages and are often written in unwieldy "legalese." If people did suddenly start reading privacy policies and TOS agreements, McDonald and Cranor note that the

---

141. Carina Paine, Ulf-Dietrich Reips, Stefan Stieger, Adam Joinson & Tom Buchanan, *Internet Users' Perceptions of 'Privacy Concerns' and 'Privacy Actions*,*'* 65 INT'L. J. HUM.-COMPUTER STUD. 526, 533–34 (2007).

142. *See, e.g.*, Susana Alcalde Bagüés, Luis A. Ramon Surutusa, Mikel Arias, Carlos Fernández-Valdivielso & Ignacio R. Matías, *Personal Privacy Management for Common Users*, 3 INT'L J. SMART HOME 89 (2009); Sarah Spiekermann & Lorrie Faith Cranor, *Engineering Privacy*, 35 IEEE TRANSACTIONS SOFTWARE ENGINEERING 67 (2009).

143. Jan Kolter & Günther Pernul, *Generating User-Understandable Privacy Preferences*, 2009 INT'L CONF. ON AVAILABILITY, RELIABILITY, & SECURITY 299, *available at* http://www.computer.org/csdl/proceedings/ares/2009/3564/00/3564a299.pdf [http://perma.cc/NUZ4-AWWF]; *see also* Yun Shen & Siani Pearson, PRIVACY ENHANCING TECHNOLOGIES: A REVIEW (Hewlett Packard Labs., Technical Report No. 2011-113, 2011), *available at* http://www.hpl.hp.com/techreports/2011/HPL-2011-113.pdf [http://perma.cc/5BRL-VYX6].

144. Keith Irwin & Ting Yu, *Determining User Privacy Preferences by Asking the Right Questions: An Automated Approach*, 2005 PROC. ACM WORKSHOP ON PRIVACY ELECTRONIC SOC'Y 47 *available at* http://dl.acm.org/citation.cfm?id=1102209 [http://perma.cc/V3UA-ZP7L].

145. Lorrie Faith Cranor, *Necessary but Not Sufficient: Standardized Mechanisms for Privacy Notice and Choice*, 10 J. ON TELECOMM. & HIGH TECH. L. 273 (2012).

146. Sameer Patil, Helping Jonny Manage Privacy: Exploring Filters as Interfaces for Privacy Preference Specification (2014) (unpublished manuscript presented at Workshop on the Future of Privacy Notice and Choice, Carnegie Mellon Univ.), *available at* https://www.cylab.cmu.edu/news_events/events/fopnac/pdfs/patil.pdf [https://perma.cc/MQ7Z-UEY4].

147. Dawn N. Jutla & Peter Bodorik, Privacy's 7Cs and the Crowded Augmented Reality User: A Position Paper (2014) (unpublished manuscript presented at Workshop on the Future of Privacy Notice and Choice, Carnegie Mellon Univ.), *available at* https://www.cylab.cmu.edu/news_events/events/fopnac/pdfs/jutla.pdf [https://perma.cc/Z4ET-26TW].

national opportunity costs could potentially be very high, perhaps into the hundreds of billions of dollars, considering how many websites the average user visits every year and how many hours would be required.[148] To make the process less time consuming, Sadeh's research team examined the possibility of automatically extracting information from privacy policies based on the topics that users are most concerned about, and then presenting the information to users in a salient manner.[149]

This is not to say that privacy policies are a lost cause for the goal of increasing consumer engagement. Some research indicates that just the presence of the policy may increase user preferences for a website. Jensen, Potts, and Jensen, for instance, found a correlation between the presence of TRUSTe seals or privacy policies and user preference for a particular website.[150] Wu, Huang, Yen, and Popova found that the content of privacy policies can influence how users interact with websites where the users must provide personal information.[151] Other research indicates that the presentation of these policies matters. Tsai, Egelman, Cranor, and Acquisti note that the effect on consumer decision making from these policies is not consistent, but that in situations where privacy information is presented in a prominent and intuitive manner, consumers may be willing to pay more money in order to purchase a product from a website that has stronger privacy protections.[152] Some research suggests that a multilayered approach to privacy notices may increase the saliency of these notices, though Good's research team also found that the use of vague language in short notices gave users an impression of increased security, compared to more detailed notices.[153]

Our research builds on previous empirical research in several ways. Previous research that focused on knowledge generally emphasized knowledge about privacy, but our research took a multilateral approach to knowledge. We evaluated participants' knowledge in five areas: cloud computing, online security, economics of the Internet, FERPA, and other privacy laws. We chose these areas because they allow us to examine the interaction between some types of more technical knowledge and privacy perceptions. Our separate analyses of the sections of the knowledge

---

148. Aleecia M. McDonald & Lorrie Faith Cranor, *The Cost of Reading Privacy Policies*, 4 J.L. & POL'Y FOR INFO. SOC'Y 543 (2008).

149. Norman Sadeh et al., Towards Usable Privacy Policies: Semi-Automatically Extracting Data Practices from Websites' Privacy Policies (2014) (unpublished manuscript presented at Workshop on the Future of Privacy Notice and Choice, Carnegie Mellon Univ.), *available at* https://cups.cs.cmu.edu/soups/2014/posters/soups2014_posters-paper20.pdf [https://perma.cc/JS6A-X82L].

150. Carlos Jensen, Colin Potts & Christian Jensen, *Privacy Practices of Internet Users: Self-Reports Versus Observed Behavior*, 63 INT'L. J. HUM.-COMPUTER STUD. 203 (2005).

151. Kuang-Wen Wu, Shaio Yan Huang, David C. Yen & Irina Popova, *The Effect of Online Privacy Policy on Consumer Privacy Concern and Trust*, 28 COMPUTERS HUM. BEHAV. 889, 896 (2012).

152. Janice Y. Tsai, Serge Egelman, Lorrie Cranor & Alessandro Acquisti, *The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study*, 22 INFO. SYS. RES. 254, 255, 257 (2011).

153. Good et al., *supra* note 137, at 50.

survey gave us the opportunity to examine how different levels of performance in different areas of knowledge relate to different sets of ideas.

Our research also takes a more comprehensive approach to privacy, considering opinions about how the private sector handles data as well as how the government handles data. The data practices of private companies are quite broad, and as the leak of documents about the NSA's PRISM program indicates, the government has similar capabilities to store and track data about citizens.[154]

In this Part and the previous Part, we provided the theoretical background for our survey as drawn from law, social science, and technology. In the following Part, we examine several of these issues through the lens of our survey results. The results of our study indicate a need for transparency and accountability among those who benefit from the big data environment. Our results reveal that there is a low level of trust of the government and its handling of data, and also a very low approval rating for the practices of big data professionals in the private sector. We therefore join the policymakers and scholars who call for more transparency as big data becomes more ingrained in our everyday lives.[155]

III. KNOWLEDGE, BEHAVIOR, AND OPINIONS REGARDING ONLINE PRIVACY

Our study examines a comprehensive range of topics relating to digital data collection. Wherever possible, we attempted to avoid relying on the word "privacy," instead presenting survey participants with more concrete descriptions of activities and concerns. We approached this project with several research questions.

First, we wanted to explore consumers' priorities with regard to digital data. How do they make decisions about which services to use? How important are ideas like security and privacy when they make these decisions? Do consumers trust online marketing companies to act in the consumers' best interests? Whom do consumers trust to protect information, and whom do consumers not trust to protect information? The results were examined with an eye to answering these and similar questions about consumer behavior and privacy preferences.

Second, we wanted to explore consumers' levels of knowledge concerning a range of topics that are relevant to digital data and information privacy. How knowledgeable are users about data in the online marketplace? How knowledgeable are users about the *risks* relating to data in the online marketplace? How knowledgeable are users about their legal rights relating to digital data? We analyzed each knowledge section separately because we recognize that there are many different facets of knowledge that contribute to a person's understanding of digital data and how to use it.

Third, we wanted to explore how behavior and priorities relating to digital data interact with knowledge relevant to digital data and information privacy. In the psychology literature, many scholars have noted the existence of a disconnect between what consumers say that they want in terms of privacy and their actual

---

154. Glenn Greenwald & Ewen MacAskill, *NSA Prism Program Taps in to User Data of Apple, Google and Others*, GUARDIAN (June 7, 2013, 3:23 PM), http://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data [http://perma.cc/M3TA-68SV].

155. *See* FED. TRADE COMM'N, *supra* note 2; Richards & King, *supra* note 5, at 396.

privacy-oriented behaviors.[156] Does level of knowledge affect the interaction between privacy desires and privacy choices? We evaluated opinion questions alongside each knowledge section individually to obtain a more nuanced perspective, given the many types of knowledge that can be relevant to privacy concerns.

A. Methodology

In order to examine what people know, what people want, and what people do online, we conducted a two-part online survey focusing on users' knowledge and opinions. Our research team composed nine distinct sets of questions. Four sets of questions were for the opinion survey, and five sets of questions were for the knowledge survey. Some of the questions were patterned after similar questions asked in surveys conducted by other researchers. We formatted the survey using LimeSurvey and hosted the survey on University of Illinois webservers. The survey link was distributed partially through social media websites like Facebook, but the highest number of participants responded when we distributed the link to an e-mail list at the University of Illinois at Urbana-Champaign (UIUC). The UIUC campus is very large, with over 32,000 students at the undergraduate level and over 12,000 students at the graduate level, and includes seventeen separate colleges.

Each survey collected demographic information from respondents, and completion of the demographic information was entirely voluntary. For both parts of the survey, a majority of respondents were between 18 and 25 years old. Some scholars have questioned whether age is a factor in attitudes towards privacy. Hoofnagle, King, Li, and Turow examined this issue and found that Americans both young and old tend to be in agreement with regard to a number of online privacy issues.[157] Thus, while the 18- to 25-year-old demographic was perhaps overrepresented, we do not believe that this age distribution skewed our results.

The opinion survey contained 58 questions. The questions were written in multiple formats. In analyzing the responses to the opinion survey, we included incomplete surveys but made sure to note the number of respondents when reporting the results of a particular question. The questions on the opinion survey fell into one of four categories: (1) online behavior, (2) personal privacy, (3) cloud service providers, and (4) government surveillance and privacy law.

The knowledge survey contained 42 questions. Of the 42 questions, 28 were multiple-choice and 14 were true/false. The questions were organized into five sections based on subject matter: (1) economics of the Internet, (2) cloud computing, (3) online security, (4) FERPA and education records privacy, and (5) privacy law. The cloud computing and online security sections focused primarily

156. *See* Good et al., *supra* note 137, at 45; Park et al., *supra* note 137, at 1024.

157. Chris Hoofnagle, Jennifer King, Su Li & Joseph Turow, *How Different Are Young Adults from Older Adults When It Comes to Information Privacy Attitudes & Policies?* (Apr. 14, 2010) (unpublished manuscript), *available at* http://ssrn.com/abstract=1589864 [http://perma.cc/5VWQ-L5ZA].

on technical aspects of the Internet. The economics of the Internet section focused on how companies make money online, and some of these questions alluded to the data broker industry. One reason that we chose to examine knowledge of FERPA instead of one of the other sector-specific federal privacy statutes was because we anticipated collecting data primarily from people affiliated with UIUC. For most college students, FERPA is directly relevant to their experiences with how the university handles their data, and students must often consent to the disclosure of their education records when they participate in extracurricular activities through the university. Law students chosen for the law review, for instance, must generally consent to such disclosures before their membership is official.

Because of the length of the survey and the importance of the final percentages, we only included fully complete surveys in our analysis of the knowledge survey. After every series of questions, there was an additional question: "Please rate your level of confidence in your answers to the above set of questions." In the analysis, we used this confidence ranking question in order to quickly tell the difference between respondents who had left a question blank because they did not know the answer and respondents who never reached a section because they did not complete the whole survey. Using this method, we determined that 455 people had completed the full knowledge survey.

B. Opinion Survey

1. Demographics

As noted above, we counted the completed answers on incomplete opinion surveys in order to evaluate specific questions. 756 people completed at least one question in the opinion survey. 99% of respondents to both surveys indicated that they live in the United States. There were significantly more female respondents than male respondents (62% female, 36% male). In terms of age, 43% of respondents were between 18 and 21 years old (the age of typical undergraduate students in the United States); 25% were between 22 and 29;[158] 14% were between 30 and 39; and 17% were age 40 or older. For religion, 55% reported a specific affiliation while 33% said "None." For highest level of education, 49% reported having a high school degree/GED or some college experience, 27% reported having an undergraduate-level degree, and 23% reported having a graduate-level degree.

2. Results

Questions on the opinion survey were presented in a variety of styles. Some questions began with a paragraph of background information and then asked for opinions. One question provided information about practices of private companies with regard to consumer information and then asked respondents if they were

158. The categories for the age demographics were: (1) younger than 18; (2) 18–21; (3) 22–25; (4) 26–29; (5) 30–39; (6) 40–49; (7) 50–59; (8) 60–69; (9) 70 or over.

previously aware that those practices existed and whether they approved or disapproved of these practices. Another question asked respondents to use the Likert scale and rank, from 1–10, how highly they prioritized certain things relating to online privacy, like the ability to control the access that others have to their information online.

a. Online Behavior

The questions in this first section gathered information about respondent's behavior on the Internet. Our sample consisted of many frequent Internet users, with about 83% of respondents reporting that they use the Internet for at least three hours each day. A study by the Pew Internet and American Life Project concluded that about 69% of Internet users in the United States already use webmail, other software programs located solely online, or online data storage.[159] Our survey sample, on the other hand, was more immersed in the Internet than Pew's sample. 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. Because this survey was conducted online, there may be some selection bias.

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 the 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 religious 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, and therefore some people may not have perceived their responses as counting for the purposes of the question. If the latter is the more accurate interpretation of this finding, it is problematic in a way similar to unconscious data sharing. 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.[160] Moreover, a recent study by Kosinski, Stillwell, and Graepel found that Facebook "likes" could

159. John B. Horrigan, *Use of Cloud Computing Applications and Services*, PEW RESEARCH CTR. (Sept. 12, 2008), http://pewresearch.org/pubs/948/cloud-computing-gains-in-currency [http://perma.cc/88BV-JAH7]; Wittow & Buller, *supra* note 119, at 5. A majority of those responding in the Pew study also indicated that they were very concerned about the use of their personal data by cloud providers. Horrigan, *supra*.

160. Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701, 1717–18 (2010) (providing an example of AOL search queries being used to identify individuals).

accurately predict personal attributes including religious views.[161] Online anonymity is not the same thing as invisibility.

As we accidentally discovered with regard to religious affiliation, Internet users may sometimes not realize that they are potentially disclosing personal information that could be tracked, especially if the website appears to be anonymous. There are also indications that in some situations where users *are* aware that they are disclosing the information, they may not want to be making such disclosures. The responses to the question displayed in Figure 1 below illustrate the sometimes coercive nature of agreements in online environments. In our survey, more than 4 out of every 5 respondents indicated that, on some occasion, they had submitted information online when they wished they did not have to. 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.

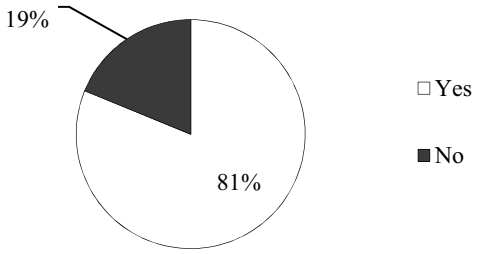

Figure 1. "Have you ever submitted information online, but wished that you did not have to?" ($n = 727$)

In order to evaluate respondents' behavior in relation to online consent agreements, we asked questions about how often people read TOS agreements and privacy policies (PP). Both readership questions utilized five-point Likert scales where the five points were assigned to "never read," "rarely read," "sometimes read," "usually read," and "always read." When examining the means of the agreement variables, the average amount that people read TOS agreements is between rarely and sometimes (M = 2.04, SD = 1.05). The average amount that people read privacy policies is also between rarely and sometimes (M = 2.12, SD = 1.08). For both questions, age (For TOS: $r = .24$, $p < .01$) (For PP: $r = .23$, $p < .01$) and educational level (For TOS: $r = .17$, $p < .01$) (For PP: $r = .19$, $p <. 01$) were positively associated with readership. Multiple regression results indicated that age significantly predicted individuals' tendency to read TOS agreements even after controlling for education level ($b = 1.58$, $p < .05$). This was also true for reading the privacy policy ($b = .18$, $p < .05$). Overall, these responses are not particularly surprising, as it has been well known for many years that most Internet users do not usually read these documents.

---

161. Michal Kosinski, David Stillwell & Thore Graepel, *Private Traits and Attributes Are Predictable from Digital Records of Human Behavio*r, 110 PROC. NAT'L ACAD. SCI. 5802 (2013).

For users who indicated that they do not always read the website policies, we asked them to select from a list of reasons why they do not. The question was in the format of "select all that apply" and they could choose from: "The document is too long," "The document uses a lot of legal terms that I don't understand," "The document is not relevant to me," "The document is not interesting," "The document is the same as every other of its type," and "I trust the company/website without reading the document." For both privacy policies and TOS agreements, the most commonly reported reason was that the documents were too long (81% for privacy policies, 53% for TOS agreements). These findings are consistent with those from a 2011 study that also found "too long" as the most commonly cited reason for individuals not reading click-through agreements.[162] We also provided a prompt to allow respondents to type in other reasons. While most participants did not fill in this prompt, several participants who did input a separate answer basically said that they do not read these documents because the owner of the website was going to do what it wanted with the information anyway. In this prompt, one participant wrote "There's nothing I can do about it if I want to use their services."

Responses to the question displayed in Figure 2 indicate that slightly less than half of our respondents had ever made a conscious decision not to use a website strictly because of the website's TOS agreement or privacy policy. Not surprisingly, we found that people who read TOS agreements and privacy policies more often were more likely to indicate that they had refused to use a website strictly because of the website's privacy policy or TOS agreement. Importantly, these correlations were very strong (For PP readership, $r = .50$, $p < .01$) (For TOS readership, $r = .50$, $p < .01$), suggesting that these documents might affect behavior more if consumers were more aware of what they said.

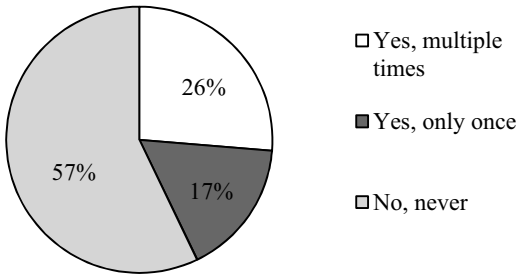

Figure 2. "Have you ever decided not to use a website strictly because of the website's Privacy Policy or Terms of Service agreement?" ($n = 707$)

b. Personal Privacy

The questions in this category were designed to assess respondents' general beliefs about personal privacy, both online and offline. We also wanted to

162. Victoria C. Plaut & Robert P. Bartlett, III, *Blind Consent? A Social Psychological Investigation of Non-Readership of Click-Through Agreements*, 36 LAW & HUM. BEHAV. 293 (2011).

explore the extent to which these beliefs relate to online behavior and opinions about cloud service providers.

This section began with questions exploring the nature of privacy. The first three questions explored the concept of privacy as a right, as a privilege, and as only being important if one has something to hide. The results are shown in Figure 3.

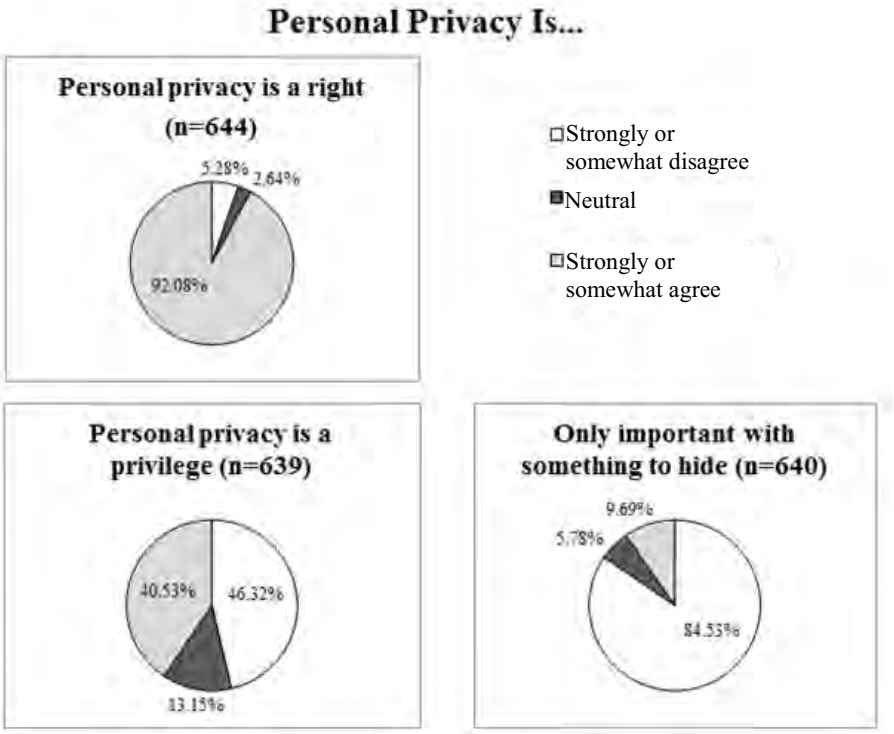

Figure 3. Nature of Privacy

The full wording of the privacy-as-right statement was "Privacy is a right to which everyone is entitled," and the privacy-as-privilege statement was "Privacy is a privilege to which people are entitled only if they follow the law." The apparent discrepancy between privacy-as-right and privacy-as-privilege is likely due to rights and privileges not being seen as dichotomous by the general public as they often are by legal professionals. As the third pie chart shows, a clear majority of participants reject the formulation of privacy as only being important to people who have something to hide.

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 participants who had achieved higher levels of education ($r = .09$, $p < .05$) were more likely to express stronger agreement. Additionally, responses to this question were correlated with some aspects of a participant's online behavior. Respondents who care more about their personal privacy were more likely to indicate that they read privacy policies ($r = .15$, $p < .01$) and that they have

previously refused to use a website because of the website's privacy policy or TOS agreement ($r = .17$, $p < .01$).

We also included two 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.

Another set of questions explored respondents' feelings about data collection online. The set of questions below addresses 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. The results are listed in Table 1.

Table 1. What should advertising companies be allowed to do?

Should online marketing and advertising companies be allowed to . . .	Yes (%)
. . . track consumers' online activity without asking for permission? ($n = 633$)	11
. . . participate in a data trade in which consumers' 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 consumers? ($n = 627$)	52
. . . track consumers' interactions with advertisements to determine which ones are the most relevant to individual consumers? ($n = 627$)	52

The distribution of responses provides valuable insight into the opinions that many people may have on these topics. The strongest majority (89%) against a particular practice concerned online tracking without consumers' permission. Responses to the last two questions concerning targeted advertising were more evenly split between support and opposition. Responses to both targeted advertising questions were positively correlated with Internet use per day (for "Use demographic profiles . . .": $r = .08$, $p < .01$; for "Track consumers' interactions . . .": $r = .09$, $p < .01$), suggesting that people who use the Internet more are more likely to support targeting advertising. This may indicate that individuals who use the Internet more are more likely to experience benefits from targeted advertisements, or that they experience more advantages than disadvantages from having personalized web browsing experiences.

To explore how participants characterized privacy, we ended this section with a series of questions asking participants the importance of several types of abilities related to privacy. The results are shown in Table 2.

Table 2. Importance of different privacy-relevant abilities

Likert Scale	Able to make decisions online without pressure from others	Control access to all of your information online	Keep others from accessing your secret information online	Be protected from unreasonable government search online
1	14	11	11	16
2	14	9	11	9
3	15	11	4	18
4	15	18	7	14
5	49	38	20	30
6	19	15	16	19
7	46	37	11	20
8	83	89	41	40
9	52	69	72	51
10	339	348	449	427
Total	*646*	*645*	*642*	*644*

As Table 2 shows, a majority of respondents chose 10, indicating very high importance for all of the scenarios. By depicting responses of 10 separately, Figure 4 provides a useful visualization of the trends for the four scenarios.

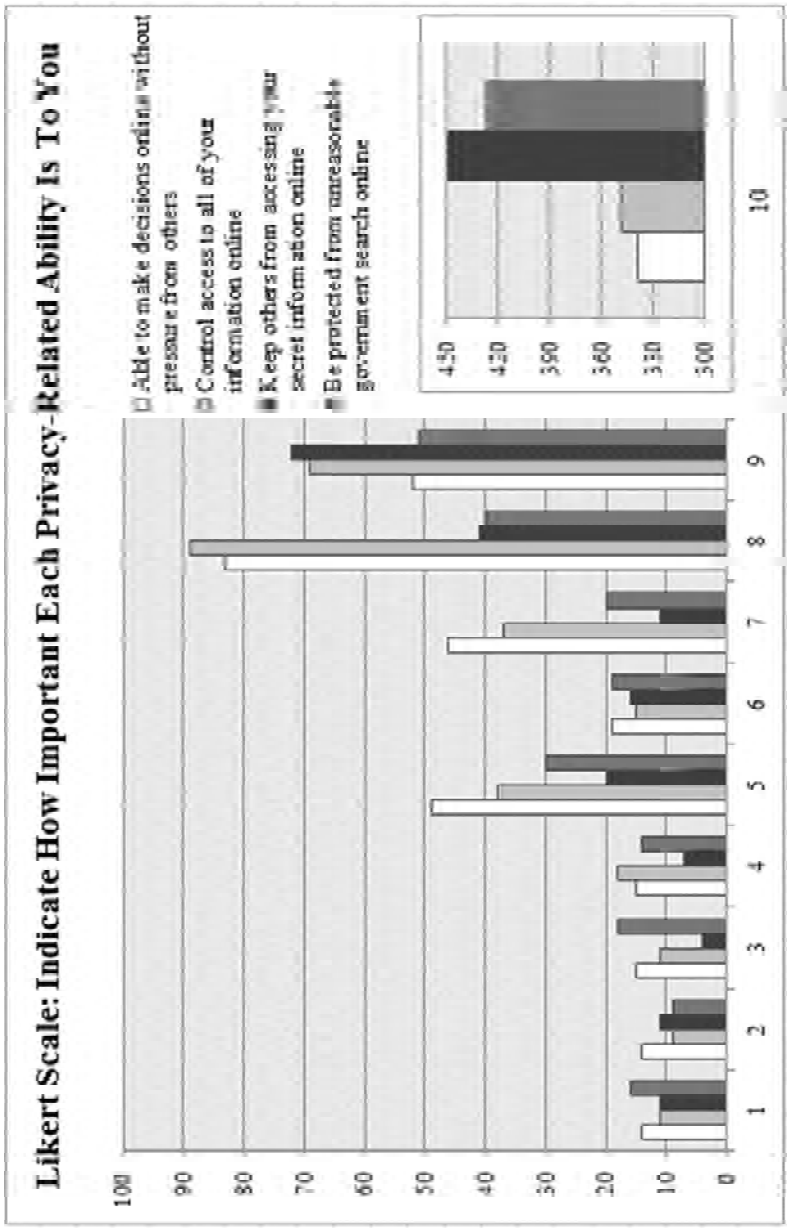

Figure 4. Different interpretations of privacy

These questions were designed to examine participants' opinions about some of the different approaches to privacy in the academic literature. The ability to make decisions without government intervention is part of decisional privacy, which the Supreme Court has acknowledged in cases like *Griswold v. Connecticut*.[163] The wording that we chose for this question was broader than just government intervention, and essentially explores whether telling intrusive parties to "mind their own business" is viewed as being an exercise of a right to privacy. By asking participants to indicate the importance of preventing access to secret information and also controlling access to nonsecret information, we explored the concepts of "privacy as secrecy" and "privacy as data control."

This question provides additional insight into the idea of what kind of harm might result from a privacy violation. The revelation of secret information, for example, seems to be a more important situation to prevent than the general data control implications from distributing nonsensitive information. This focus on privacy as secrecy may indicate a desire to avoid reputational harm from disclosures. This interpretation is made more likely in light of the findings in Figure 3, where participants largely disagreed that privacy was only valuable for people who have something to hide. Comparing the responses to that question with the responses to this question, we conclude that the privacy as secrecy paradigm is not reliant on the secrets actually being something that the target wants to *hide*, just something that the target does not want to *disclose*.

Likewise, because the "something to hide" rationale was largely rejected, the prevention of unreasonable government searches can be understood as a method for preserving individual freedom. This purpose of privacy was, on average, considered more important than the self-determination ability to make one's own decisions without undue pressure from others. A stronger preference for protecting information from the government and maintaining the secrecy of sensitive information may therefore signal a preference for using privacy to preserve individual freedom and avoid reputational harm.

c. Cloud Service Providers

The third category of opinion questions assessed what people think and know about how cloud service providers collect, maintain, and distribute customer information. Because one of our main goals was to learn more about the experience of consumers, we focused on examples of cloud computing services that would be considered Software as a Service (SaaS) instead of Platform as a Service (PaaS) or Infrastructure as a Service (IaaS).[164]

Some of the questions in this section explored issues similar to those explored by McDonald and Lowenthal, who discovered that consumers were often surprised by the breadth of data collection online.[165] The first set of questions displayed below

---

163. 381 U.S. 479 (1965) (ruling unconstitutional a Connecticut law that prohibited the use of contraception).

164. *See generally* Kesan et al., *supra* note 18, at 361.

165. Aleecia M. McDonald & Tom Lowenthal, Nano-Notice: Privacy Disclosure at a Mobile Scale, Panel Discussion at the Workshop on the Future of Privacy Notice and Choice (June 27, 2014), at 5, *available at* https://www.cylab.cmu.edu/news_events/events/fopnac

was based on our previous empirical analysis of the TOS agreements and privacy policies of nineteen leading cloud service providers.[166] In this set of questions, we assessed whether respondents were previously aware of different ways in which cloud service providers, according to provisions commonly found in their TOS agreements or privacy policies, may use customer uploaded information. Additionally, we evaluated whether respondents approved or disapproved of each method once made aware. The results are provided in Table 3.

Table 3. Practices of cloud service providers

Statement	Previously aware (%)	Approve (%)
Some cloud service providers may reproduce or modify works or content uploaded by users. (n = 528)	29	6
Some cloud service providers may publish, publicly display, or distribute content uploaded by users. (n = 529)	46	12
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

The response patterns above reveal significant gaps in people's awareness of common themes within the privacy policies and TOS agreements of leading cloud service providers. Not surprisingly, privacy policy readership was positively related to the "previously aware" aspect of all three questions (for first question: $r = .18$, $p < .01$) (for second question, $r = .13$, $p < .01$) (for third question, $r = .19$, $p < .01$).

The next part of this section of the survey introduced the concept of trust and invited participants to indicate whether they thought that certain potential recipients could be trusted to protect confidential information. Table 4 shows the findings from this series of questions.

Table 4. Who could be trusted to protect your confidential information?

Entity	Could be trusted (%)
Law enforcement agencies (n = 559)	55.5
Government, non-law enforcement (n = 560)	43.9
Internet service providers (n = 566)	16.4
Advertising agencies (n = 569)	4.6
Search engines (n = 567)	8.3
Cloud service providers (n = 561)	12.3
Your close friends (n = 572)	69.8
Your employer (n = 559)	54.0
Your family (n = 573)	84.6
Your health care provider (n = 560)	65.4
Your school/university (n = 561)	64.9

/pdfs/mcdonald-slides.pdf [https://perma.cc/8NDE-QJ5T].

166. Kesan et al., *supra* note 18.

As Table 4 shows, participants often do not trust Internet-based companies to protect their confidential information. On the other hand, the rate at which they trust law enforcement is comparable to the rate at which they trust their employer. Meanwhile, health care providers and educational institutions are trusted at a rate slightly lower than close friends. More interesting details emerge when these responses are compared with performance on the knowledge survey, which we explore in Part III.C.

After asking about trust levels for various people and entities, we asked participants about whether they would approve of a cloud service provider giving the participants' personal information to law enforcement and other potential recipients without the participants' knowledge, and then whether they would approve of the business selling the same information without the participants' knowledge. As Table 5 shows, the approval rate for both tended to be low.

Table 5. Approval rates for giving away or selling consumer information

	Approval rate	
	Give away (%)	Sell (%)
Law enforcement	31.1	7.7
Other government agency	23.4	7.5
ISPs	6.4	3.3
Advertisers	7.7	7.3
Search engines	7.4	6.1
Your employer	8.7	4.4

The highest approval rates were for giving this information to law enforcement or to non-law enforcement government agencies, but the approval rate dipped sharply when the question concerned selling information. 31% of participants ($n = 566$) would approve of the provider giving their personal information to law enforcement, but only 7.7% would approve of the provider selling information to law enforcement. From a privacy harm perspective, this would not make sense if the harm is solely the unauthorized acquisition of personal information by the government. The approval level plummeted between the information being given and the information being sold. While we do not know what the participants' reasoning for this changed view might be, we propose that participants might have perceived the privacy violation to be more egregious in the case of purchasing the information instead of receiving the information for free. This rationale has echoes of Cohen's self-determination approach to privacy, as this sort of economic exchange could be understood as the commodification of the consumer.

The sale of information by cloud providers illustrates a disconnect between what people experience and what they would prefer online. 85.8% of respondents to the opinion survey ($n = 572$) answered "No" to the question "Do you think cloud service providers should be allowed to sell information about users' activities on their websites?" Furthermore, 84.35% of respondents ($n = 556$) answered "Yes" to a question that asked if they believe that cloud service providers sell this information to other organizations. This is an example of a situation where the consumers are often fully aware that this activity goes on, and they do not want it to, but they feel powerless to change what the businesses do.

Consumer engagement with services could potentially be improved if service providers adopted a different business model that gave users more options for

controlling their own privacy settings. Figure 5 displays the results of a scenario-based question that explored preferences for alternative business models. Respondents were given a choice between three options: (1) receiving a free service with fewer features where customers have more control over their data, (2) paying for a service that gives customers more features and more control over their data, or (3) receiving a free service with more features and less control over their data. These three choices represent three possible business models that companies might follow, with the third option representing the current dominant model. Essentially, the decision between the second and third options boils down to whether the respondents prefer to pay for a service with money or with personal information. The results are presented in Figure 5.

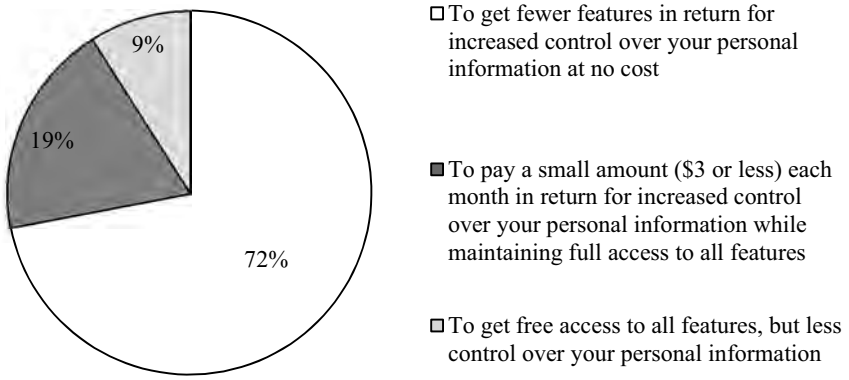

Figure 5. "If a cloud service provider were to adopt a new policy that gives users different options for their privacy settings, would you prefer:" (n = 554)

As this figure shows, only 9% of respondents indicated that they would prefer to have more services, at no cost, with less control over their personal information—in other words, they prefer the current status quo for data privacy. A majority of the survey respondents indicated that they would prefer to receive a free service with fewer features but have more control over their data. This result suggests that, when it comes to services online and how the services use consumer information, the modern consumer may want more options than the market currently provides. Most importantly, it indicates that one of the most common online business models, where users get a lot of features in exchange for allowing their information to be used for targeted advertisements, may be contrary to the actual desires of consumers.

d. Government Surveillance and Privacy Laws

Many of the opinion questions were about more general Internet privacy issues, but we also used the survey to collect opinions about laws that apply to privacy online. Asking nonlawyers what they think the law should be provides an insight into some of the very practical aspects of governance that practitioners, academics, and policymakers sometimes lose sight of. The first question of this section concerned government data collection practices, and the results are shown in Figure 6.

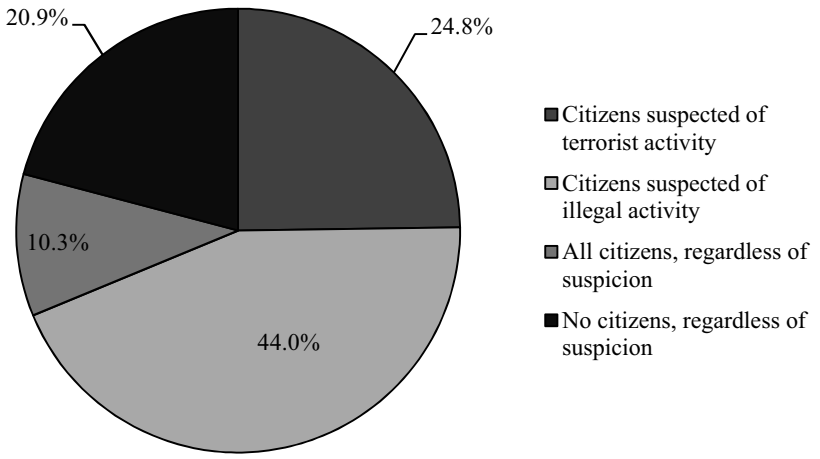

Figure 6. "The government should be able to collect and analyze private information about . . ." (*n* = 541)

As Figure 7 shows, a majority of respondents indicated that they believe that U.S. law currently does not provide strong enough protections for privacy online.

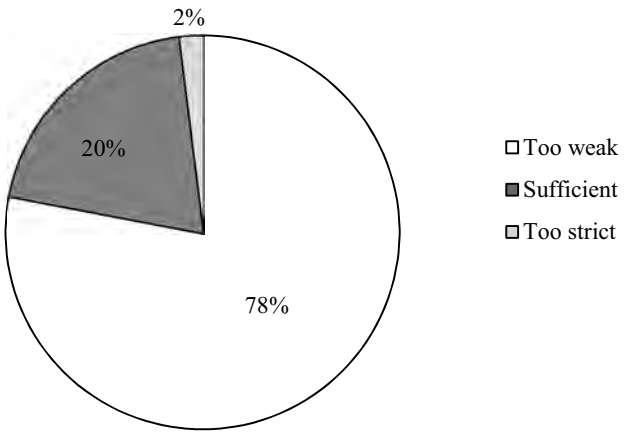

Figure 7. "The current U.S. laws designed to protect individuals' privacy online are . . ." (*n* = 495)

Because of the digital surveillance controversy that came to light in 2013, we also examined respondents' reactions to privacy issues that arise between private citizens and the government. While the average citizen may not be familiar with the intricacies of Fourth Amendment jurisprudence, a majority of the respondents to our survey believe that digital "papers and effects" should be protected just like their physical counterparts. Responses also indicated that people often favor requiring a search

warrant, especially for the contents of communications. This may indicate a desire for more accountability and transparency in government. As the below figure shows, a majority of respondents believe that the government currently has too much power to collect private information based on public safety justifications.

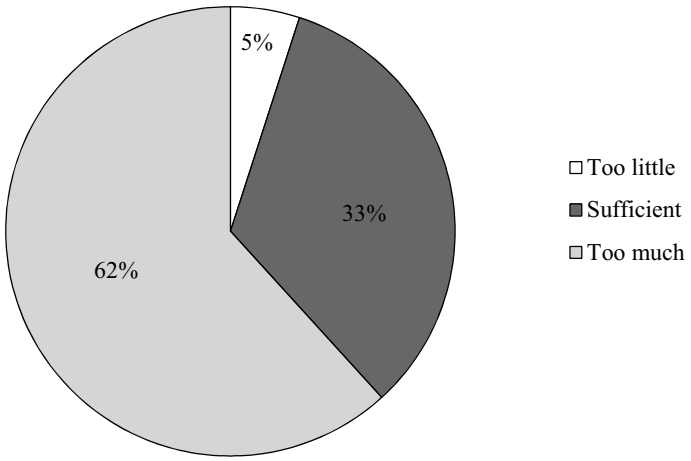

Figure 8. "Which of the following best describes your opinion regarding the amount of power the U.S. government has to collect and analyze private information about citizens in the name of public safety?" (n = 545)

We also presented participants with a scenario about Bill and a hypothetical search of his Gmail account. In these questions, participants were asked to address the issues of obtaining the consent of an account holder and notifying an account holder after a search, both with and without a search warrant. We drafted these questions with the goal of gaining insight into the opinions of laypersons about some of the notice and consent issues that arise in the context of online information searches. As Table 6 indicates, the presence or absence of a warrant had a noticeable effect on whether the participants thought that Bill's consent should be obtained first or if he should be notified afterward.

Table 6. Warrants, consent to search, and notification about past searches (n = 546–550)

	Consent	Percentage (%)	Notification	Percentage (%)
No warrant	Google should obtain	48.2	Google should notify	52.7
	Government should obtain	37.6	Government should notify	36.9
	Neither should obtain	14.3	Neither should notify	10.4
Warrant	Google should obtain	21.2	Google should notify	28.6
	Government should obtain	19.9	Government should notify	48.6
	Neither should obtain	58.9	Neither should notify	22.9

We found the responses to law-oriented questions to be very enlightening because, while very few of the respondents were legal professionals, everyone is affected by how the law is applied. Whether they are legal professionals or laypersons, reasonable people may differ about what sort of application of the law sounds "right." Developing an idea of what sounds "right" to the most people could potentially help in shaping future policy decisions on a massive scale. This potential is especially salient in the realm of digital privacy, where simply applying old laws to new technology often falls short.

Taken together, the results of the opinion survey provide interesting insights into consumer behavior. Over 92% of respondents to our survey indicated that privacy was somewhat or very important to them. 81% of respondents have experienced information-submission regret at some point, and a somewhat surprising 43% of respondents have decided at least once to not use a service based solely on that service's privacy policy or TOS agreement. A majority of survey participants rejected the idea that privacy is only valuable to those who have something to hide, but they also supported the secrecy paradigm of privacy. This result suggests that they do not view secrecy as being inextricably connected with the act of hiding something. A majority of respondents also think that U.S. law does too little to protect privacy and that the U.S. government has too much power to collect and analyze private information. Having explored what people do and what people want with regards to privacy, we turn now to the knowledge survey to estimate a baseline for what people know.

C. Knowledge Survey

1. Demographics

We received a total of 455 complete responses to the knowledge survey, with 99% of these responses coming from people living in the United States. There were slightly more female respondents than male respondents (55% female, 45% male). In terms of age, 37% of respondents were between 18 and 21 years old, 28% were between 22 and 29, 16% were between 30 and 39, and 19% were age 40 or older. Overall, the sample was highly educated; 60% of respondents had at least completed an undergraduate-level degree, and almost half of this figure (29% overall) had also completed a graduate-level degree. Our sample may also have been more technically inclined than the general population. Approximately half of respondents who identified their major field of study in the knowledge survey listed their major as engineering, mathematics, or a field of hard science.

Because the knowledge survey was long and we only counted the completed surveys, the data is subject to some self-selection bias on the part of the participants. Respondents who completed the entire survey may have had different traits than respondents who stopped answering questions halfway through the survey. For example, this self-selection bias might have affected some of the demographic differences between the opinion survey and the knowledge survey. In the opinion survey, all responses were counted, whether the respondent completed the full survey or not, and approximately 50% of respondents had an undergraduate- or graduate-level degree. On the other hand, 60% of

respondents who completed the full knowledge survey had an undergraduate or graduate-level degree.

2. Results

a. Overall Knowledge Scores

Knowledge scores were computed by assigning equal weight to all 42 questions. We converted raw scores into percentages and then calculated mean scores for each section and overall. Figure 9 below displays the frequency distribution for overall knowledge scores. The middle 50% of scores fell between 52% and 70%.

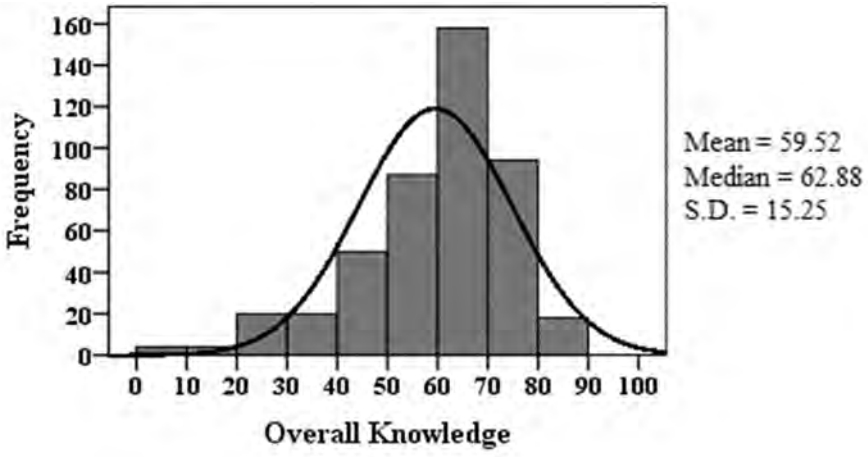

Figure 9. Frequency distribution for overall knowledge scores

We included five categories of questions in the knowledge portion of the survey: cloud computing (CC); online security (OS); the economics of the Internet (EI); education records and FERPA (ER); and other privacy law issues (PL). Compared to the mean performance in other sections, respondents did not perform as well on the ER and PL sections. Figure 10 below displays participants' mean knowledge scores for each of the five sections as well as their overall score. Error bars indicate ±1 standard deviations.

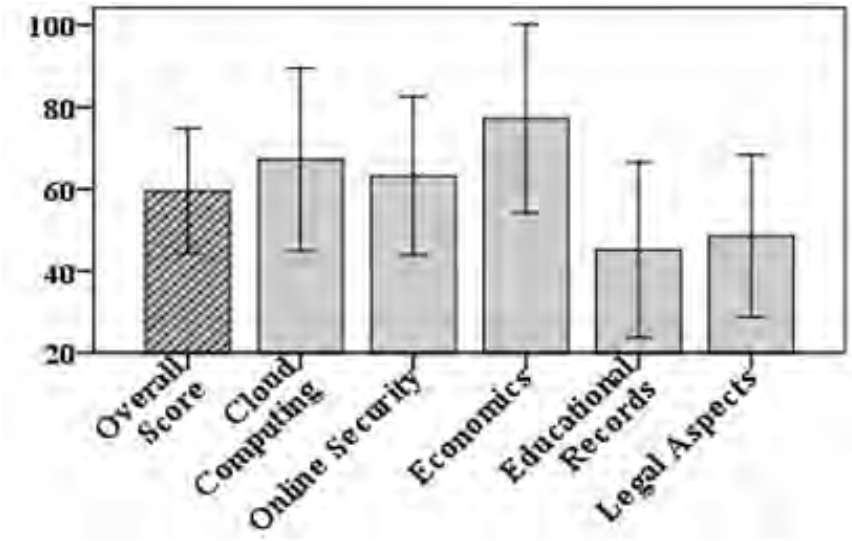

Figure 10. Mean knowledge scores by section and overall

We note that the subject matter of the different sections inherently affected their difficulty. For example, several of the questions in the ER and PL sections were based on specific provisions of different U.S. laws designed to protect individuals' privacy. In contrast, the EI section included more general questions about the data trade for personal information online.

In order to evaluate demographic differences in knowledge, we performed 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 degrees.

Additionally, we investigated whether age or education level predicts privacy knowledge, controlling for each other. Results of multiple regression analyses indicated that age (b = 2.29, $p < .05$) and education level (b = 3.08, $p < .05$) both predicted overall scores.

b. Cloud Computing

In the first section of the knowledge survey, we examined participants' understanding of concepts relating to cloud computing. The term "cloud computing" can mean different things to different people.[167] On a general level,

167. For more discussion of cloud computing in the context of online privacy, see *id.*

"cloud" is used as a metaphor for the "ethereal Internet" and the virtual platform that it provides.[168] The National Institute for Standards and Technology (NIST) currently defines cloud computing as "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (*e.g.*, networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."[169]

The first section of our knowledge survey assessed respondents' basic knowledge of cloud computing and how it works. Figure 11 below presents the frequency distribution of knowledge scores within the section.

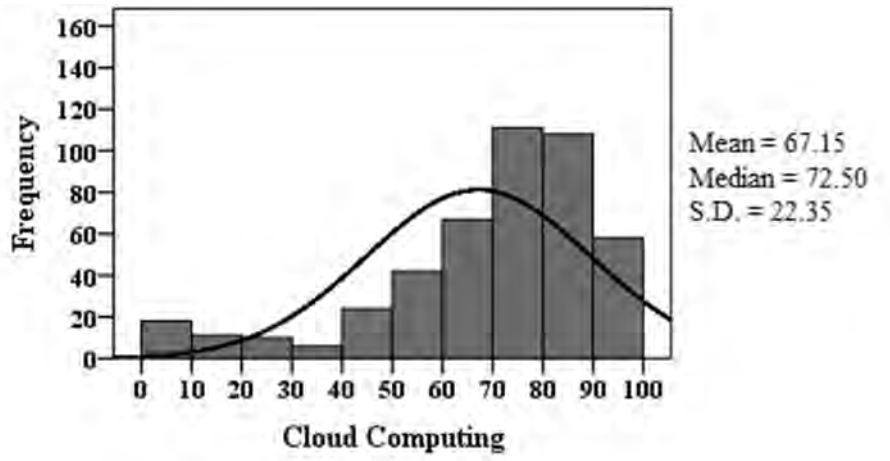

Figure 11. Frequency distribution for cloud computing knowledge

Table 7 summarizes respondents' performance in this section based on subject matter. The "question categories" are condensed versions of the overarching research questions we investigated. Note that certain categories consist of more than one question. The analysis of question categories is presented at the question-level view, while the above figure presents the data at the level of the participant.

---

at 354–61.

168. David A. Couillard, Note, *Defogging the Cloud: Applying Fourth Amendment Principles to Evolving Privacy Expectations in Cloud Computing*, 93 MINN. L. REV. 2205, 2216 (2009); *see* Wittow & Buller, *supra* note 119, at 1.

169. PETER MELL & TIMOTHY GRANCE, NAT'L INST. OF STANDARDS AND TECH., SPECIAL PUB. NO. 800-145, THE NIST DEFINITION OF CLOUD COMPUTING 2 (2011), *available at* http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf [http://perma.cc/8LMG-A35B].

Table 7. Cloud computing knowledge by subject matter

	Number of questions	Average percentage (%) correct
How can individual users benefit from cloud computing?	2	86
How does cloud computing work?	3	78
What is cloud computing?	1	75
How are cloud services distributed to businesses?	3	70
What are some common examples of cloud-based websites and services?	1	8

As can be seen above, respondents performed best on the two questions that asked about the benefits of cloud computing for individual users. More 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. On the other hand, respondents performed worse on the question that asked about examples of cloud services. This was largely driven by the finding that more than half of respondents failed to indicate that Twitter and Netflix are cloud-based websites, though it is appropriate to give some leeway on this topic since the understanding of what constitutes a cloud service is not always concrete even for professionals.

Our results indicated that 75% of participants could correctly identify what cloud computing is. This indicates that a quarter of our respondents still find cloud computing to be a perplexing concept, but respondents to our survey seem to have more understanding of cloud computing than respondents to a similar survey in 2012. In that survey by Wakefield Research, 54% of respondents claimed to not use cloud computing, but their responses to questions about the services they use indicated that 95% of them in fact do use cloud computing.[170]

Overall, the response patterns in this section suggest that consumers are generally aware of how cloud computing can be used to back up digital files instead of solely storing them on a local machine. However, consumers are less knowledgeable about other aspects of cloud computing, such as its fundamental nature and how it can be used for things other than backing up files (e.g., storing video preferences and viewing habits on Netflix, or perpetual storage of tiny tidbits of thoughts using Twitter).

c. Online Security

Our goal in the second section was to assess what people understood about basic issues related to cybersecurity. There are many risks involved in accessing the Internet, including viruses, phishing, and identity theft. Remotely stored data that

170. WAKEFIELD RESEARCH, CITRIX CLOUD SURVEY GUIDE 1 (2012), *available at* http://s3.amazonaws.com/legacy.icmp/additional/citrix-cloud-survey-guide.pdf [http://perma.cc/4VXJ-VD4X].

are not intended for public access are likely to be encrypted, password protected, or have unlisted links.[171] Data, especially data not considered "sensitive," are typically stored in an unencrypted format.[172] In this section, we examined participants' knowledge of topics like secure wireless connections, computer viruses, and encryption. Figure 12 below presents the frequency of knowledge scores within the section.

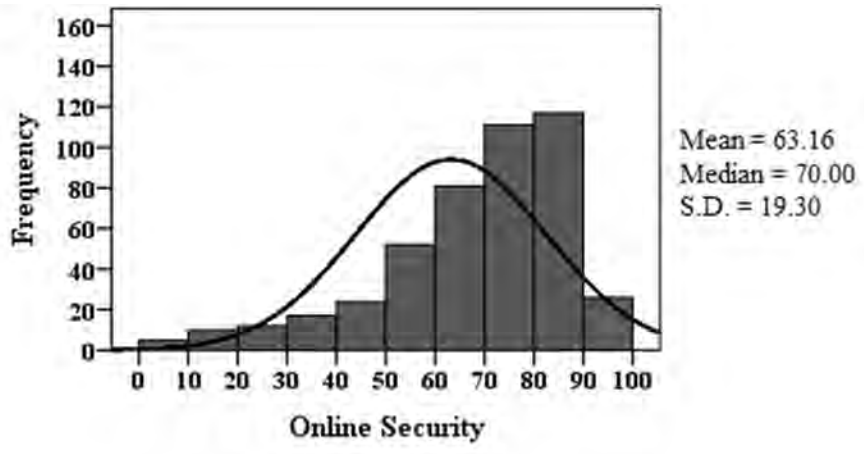

Figure 12. Frequency distribution for online security knowledge

Table 8 displays question-level data based on the subject matter of questions.

Table 8. Online security knowledge by subject matter

	Number of questions	Average percentage (%) correct
What does encryption mean?	1	84
What are cookies and what type of data do they collect and store?	2	80
How should you handle a questionable phishing e-mail?	1	80
What are some everyday online security risks?	3	76
Where are different types of personal information stored online?	3	76

The table above shows that performance in this section was relatively consistent across subject matters. On the high end, about 84% of respondents correctly

171. Couillard, *supra* note 168, at 2217.

172. *See* Stylianou, *supra* note 121, at 605. Because Google does not encrypt stored e-mails, for example, Google's software can scan e-mail content for key words for the purpose of targeted advertising. *Id.*

answered the question that asked about the definition of encryption. On the low end, about 76% correctly answered the questions about everyday security risks and where popular online services store customer data.

The specific question that participants performed worst on 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 Wi-Fi network—that someone else using the network could potentially see all the websites that one visits and the information that one transmits.

Approximately one in four respondents did not select the correct response to the three questions that asked about where different types of personal information are stored online. Most notably, only 77% of respondents knew 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 (75%) who correctly answered 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.

d. Economics of the Internet

The providers of cloud services may take a variety of approaches to service provision, from different cost models to different user interfaces to different treatment of user data. There are many cloud services that are already provided for free, and these services can remain profitable by relying on ad support.[173] Companies that do so often use customer information to generate targeted advertisements, which some criticize as effectively "monetiz[ing] their users' private data."[174] The questions in this section of our knowledge survey focused on 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. Overall, respondents performed the best in this section. The frequency distribution for this section is given below in Figure 13.

173. *See* George Jiang, Note, *Rain or Shine: Fair and Other Non-Infringing Uses in the Context of Cloud Computing*, 36 J. LEGIS. 395, 415 (2010).

174. Christopher Soghoian, *Caught in the Cloud: Privacy, Encryption, and Government Back Doors in the Web 2.0 Era*, 8 J. TELECOMM. & HIGH TECH. L. 359, 396 (2010).

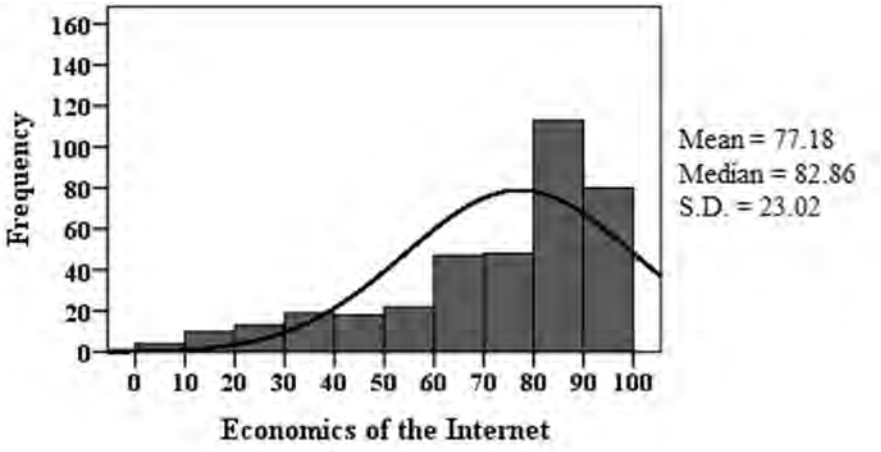

Figure 13. Frequency distribution for economics of the Internet knowledge

Table 9 displays the average percentage of correct responses for each of the three subject matter categories in this section.

Table 9. Economics of the Internet—knowledge by subject matter

	Number of questions	Average percentage (%) correct
How are targeted advertisements created/what are they based on?	3	82
How do free websites make money through advertisements?	2	82
What can online advertising companies do to collect information about Internet users?	2	68

Although on average respondents performed better in this section than any other, the results indicate that when it comes to the business models of companies online, many people are still confused. For example, on average only 68% of respondents correctly answered the questions about what online advertising companies can do to collect personal information about users.

The responses to individual questions within the first two categories in Table 6 also stood out. In one of the two questions that asked about how free websites make money, only 56% of respondents knew that free websites could make money by selling user information directly to marketing companies. Additionally, in one of three questions that asked about how targeted advertisements are created, only 66% of respondents correctly knew that advertising companies could use e-mails sent and received on free webmail accounts to personalize advertisements.

e. Educational Records

In the fourth section, we explored knowledge regarding FERPA, a federal law that governs the privacy of educational records in the United States. This has become an

increasingly important topic in recent years as many educational institutions have transitioned towards storing students' educational records online. Participants performed the worst in this section, averaging only 2.7 correct responses out of six questions. Figure 14 illustrates the frequency distribution for scores in this section.

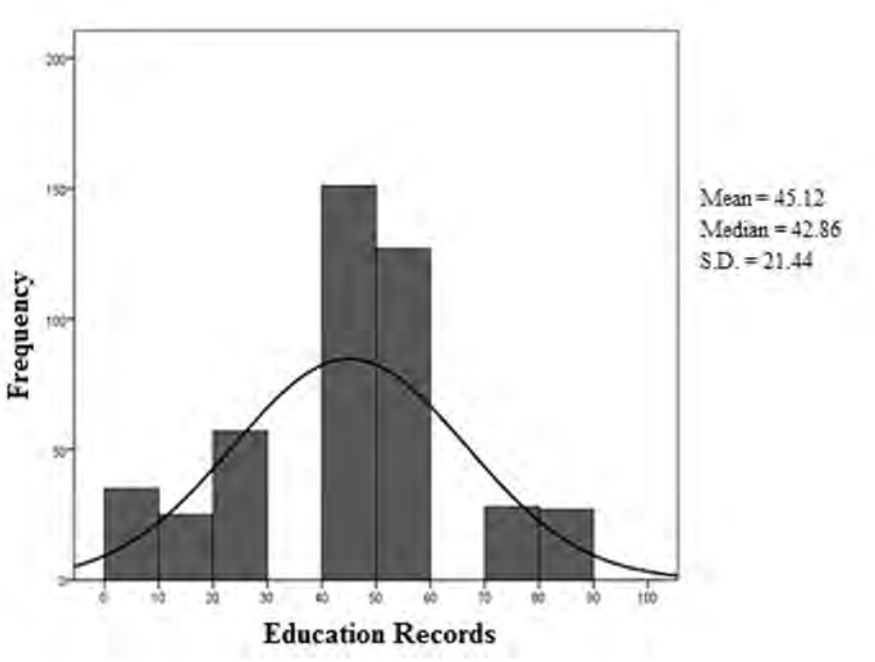

Figure 14. Frequency distribution for FERPA knowledge

Table 10 displays, at the question level, the average percentage of correct responses for each of the four subject matter categories in this section.

Table 10. Educational records—knowledge by subject matter

	Number of questions	Average percentage (%) correct
What information is included in "educational records"?	1	59
Is the student's consent needed before this type of information can be shared? (Multiple choice)	2	15
Is the student's consent needed before this type of information can be shared? (True or false)	2	80
Can parents challenge a minor's inaccurate educational records?	1	55

The first question that we asked in this section was a simple self-reporting question where we asked the participants if they had "ever heard of the Family

Educational Rights and Privacy Act (FERPA)." Out of the 455 completed surveys, 59.1% of respondents said "Yes," 39.2% said "No," and 1.8% selected "Don't know." Considering that most of the survey participants were likely to be affiliated in some way with UIUC, the fact that less than 60% of respondents were even familiar with the name of the law seems surprisingly low. College students must often sign paper forms or click through web forms to indicate their approval when their education records need to be shared with another party, and these forms often reference FERPA by name or acronym. The low level of awareness indicates that students may not be paying much attention to these forms.

FERPA regulates the disclosure of education records, and there are several exceptions that allow information to be shared without the student's consent. To examine participants' knowledge of FERPA, we presented them with three multiple-choice scenario questions involving a student named Tom. In one question, we asked participants how the parents of Tom, a university student, could see Tom's grades if Tom was unwilling to show his grades to anyone. The correct answer was that his parents would need to show that they claim Tom as a dependent for tax purposes, and only 17% selected that option. On the other hand, 55% selected the incorrect option that said that Tom's parents could never get access to his grades. In a second question, Tom was transferring to Case College, and the question asked about what Case College would need to do to access Tom's education records. Because transferring schools is another exception to the consent requirement, the correct answer is that Case College would not need Tom's consent. Only 12% gave the correct answer, while 68% incorrectly answered that Case College would need Tom's consent before they could obtain Tom's educational records.

The section also contained three true/false questions, on which respondents performed much better. When given the statement "Students who are 18 or older must generally give their consent before a school can share their educational records with someone else," 78% correctly marked "True." In response to the statement "A potential employer can visit an applicant's university and look at the applicant's grades without making a formal request or getting the applicant's permission," 81% of respondents correctly answered "False."

We concluded the FERPA section with a few opinion questions, which participants were asked to answer according to a Likert scale that ranged from "Strongly disagree" to "Strongly agree." Over 80% of respondents indicated that the privacy of their education records was important. The second statement was "If given a choice, I would prefer to receive my grades in the mail or in person instead of having my educational records stored online." Only 23% of respondents marked that they agree with that statement "somewhat" or "strongly," and 46% of respondents disagreed with that statement "somewhat" or "strongly." The remaining 31% of respondents selected "neutral." Because over 80% said that privacy in their education records was important, this suggests that most respondents trust the Internet as a medium for transmitting this information.

f. Privacy Laws

The questions in the final section concerned respondents' knowledge of laws that govern online privacy in the United States. The mean score was only slightly higher than the ER section, with participants averaging about 4.4 correct answers out of 9.

This was generally expected as the questions here tapped into somewhat specific legal knowledge. Respondents' legal scores were also positively correlated with scores in all of the other sections. For the EI section, the correlation was .43 ($p < .05$); for the CC section, the correlation was .37 ($p < .05$); for the OS section, the correlation was .47 ($p < .05$); and for the ER section, the correlation was .30 ($p < .05$). Figure 15 below illustrates the frequency distribution of scores for this section.

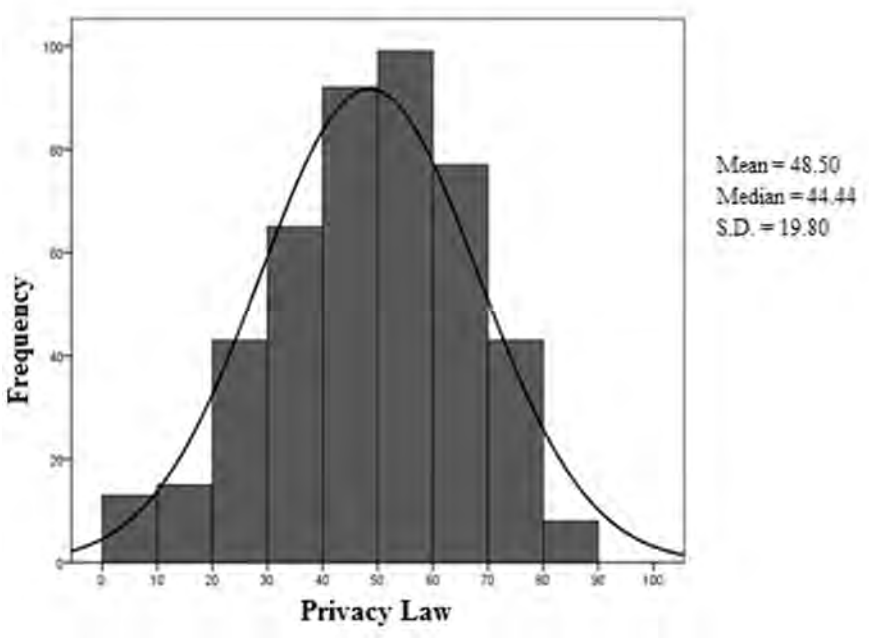

Figure 15. Frequency distribution for privacy law knowledge

Table 11 below displays the average percentage of correct responses for each of the six subject matter categories in this section.

Table 11. Privacy law—knowledge by subject matter

	Number of questions	Average percentage (%) correct
What is the First Amendment?	1	88
Freedom of speech	2	39
What is the Fourth Amendment?	1	72
Unreasonable searches	1	30
Stored Communications Act	3	35
Can you amend Facebook's privacy policy by posting on your Facebook wall?	1	68

The first two questions concerned the First and Fourth Amendments to the U.S. Constitution. The questions were multiple choice and asked respondents to identify the correct description of the respective amendment. 88.5% of respondents correctly

identified the First Amendment as applying to the freedom of speech and freedom of the press. Less than 5% of respondents gave an incorrect response, and 7.3% marked "Don't know." The responses to the Fourth Amendment question indicate that participants are less familiar with the Fourth Amendment than the First Amendment, but performance was still fairly good. 72% selected the correct description ("protects citizens from unreasonable searches and seizures by the government"); approximately 10% of respondents gave an incorrect response; and 16.1% selected "Don't know."

After these basic questions, the difficulty of the questions increased, and performance declined. The fourth question was a true/false question, and participants were asked to evaluate the statement "Under the U.S. Constitution, 'freedom of speech' means that private citizens cannot attempt to censor each other." 39.8% of respondents incorrectly marked "True," and 48% marked "False." In the sixth question, we asked participants to identify the correct description of the Stored Communications Act. 27.7% selected the correct option, and 54.6% marked "Don't know." This is roughly the response level that we were expecting because the Stored Communications Act, in spite of its importance to online privacy, is not often discussed by laypersons.

The questions about the SCA were surface-level questions, and we did not ask participants to differentiate between types of service or between the types of legal documents that might be required. In the final three questions in this section, we presented a scenario about Bob, who is being harassed on Facebook by someone with the e-mail address of sneakypeach79@gmail.com. The first two questions concerned the SCA and, specifically, how it applied when someone wanted to identify whom the Gmail account belongs to. Under the SCA, companies can disclose basic subscriber information to private citizens without any legal documents compelling that disclosure, but if the government wants basic subscriber information, they must present the company with a legal order like a subpoena. 55.6% of respondents incorrectly guessed that Bob would have to file a police report and have the police request the information from Google, and only 5% correctly identified that Bob could request that information himself. 20.8% of respondents marked "Don't know." However, respondents were more accurate when the question concerned what the government would have to do if government actors, and not Bob, wanted to identify the owner of the Gmail account. 72.2% of respondents correctly answered that the government would need to procure a legal order signed by a judge, like a subpoena or a search warrant. This time, only 14.6% checked "Don't know."

D. Interaction of Knowledge and Opinions

For the final part of our study, we created a separate database to represent survey participants who completed both surveys. This allowed us to conduct a more thorough analysis of the interaction between knowledge and opinion responses. The demographics for this sample roughly mirrored the demographics for the knowledge survey. Figure 16 illustrates that the education level of participants in this database was comparable to the level found in the knowledge survey.

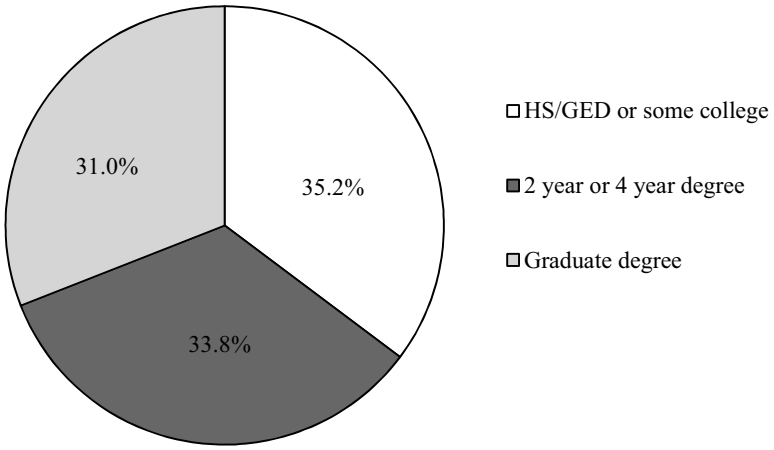

Figure 16. Highest level of education completed (n = 210)

1. Methodology

To create a database of survey participants who took both surveys, we used the Internet Protocol (IP) address that was automatically tracked during the initial data-collection process. In keeping with requirements of protecting participant information, the IP addresses were deleted from the database as soon as the group was identified. There were 210 participants who took the opinion survey and also completed the full knowledge survey. Because some participants left some opinion questions blank, the number of participants for each analysis varies.

The order in which the surveys were taken varied. During the time that the data was collected, we changed the order that the surveys were listed on the survey website, so that sometimes the opinion survey was listed first, and other times the knowledge survey was listed first. Because the opinion survey included some explanatory information that was also tested on the knowledge survey, we evaluated the sample to detect any differences between the scores for the knowledge survey for participants who took the opinion survey first and those who took the knowledge survey first. No significant differences emerged to suggest that taking the opinion survey first led to higher knowledge scores. The only significant difference was in the privacy law section, where the mean score for those who completed the knowledge survey first was higher than the mean score for those who completed the opinion survey first (p = .0008).

To evaluate performance on knowledge sections alongside the participants' responses to the opinion survey, we used frequency analyses and quartile analyses. For the quartile analyses, we compared the responses of those who scored in the bottom quartile in a particular knowledge section with those who scored in the top quartile in the same section. This allowed us to identify

opinions and behaviors that were more common in the lowest and highest scorers.

2. Results

We present the results based on the opinion section categories. When we conducted the quartile analysis, we set a threshold of ten percentage points difference between the lower quartile and upper quartile for the purpose of our initial analysis. For the purpose of this discussion, we primarily focus on quartile differences that are greater than 20 percentage points but occasionally discuss quartile differences greater than 10 percentage points as well.

a. Online Behavior

The first opinion category that we analyze in this smaller sample is the online behavior category of opinion questions. Like the larger sample, this smaller sample was also composed of heavy Internet users. There was no significant difference between those who scored in the first quartile and those who scored in the fourth quartile on any of the knowledge categories.

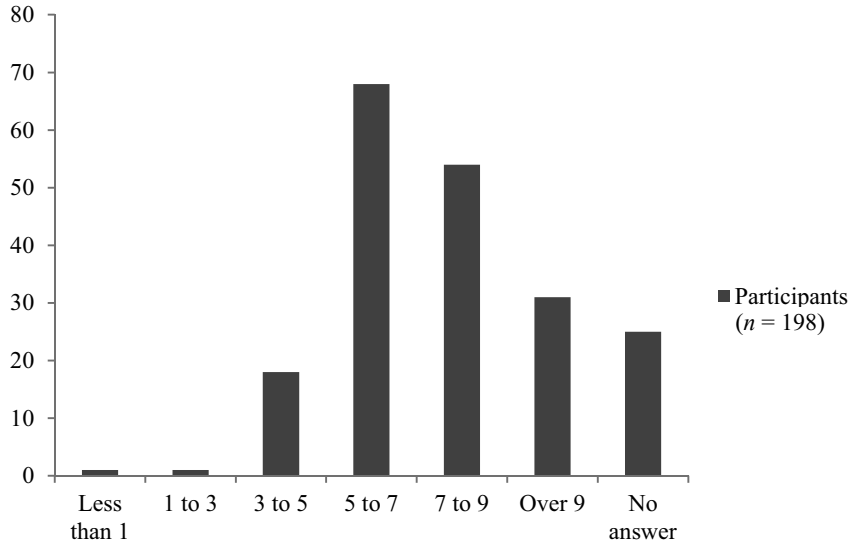

Figure 17. Hours per day on the Internet

One of the connections that we discovered between knowledge and behavior was that those who scored in the highest quartile of the cloud computing section were more likely to have submitted certain categories of sensitive information online, compared to those who scored in the lowest quartile of the cloud computing section. This is shown in Table 12.

Table 12. Quartile analysis: Disclosing information online

	Comparing to cloud computing scores	
Information disclosed	Percentage (%) of bottom quartile	Percentage (%) of top quartile
Social security number	64.7	77.4
Bank account number	54.9	81.1
Mother's maiden name	66.0	84.9
Income information for tax purposes	54.9	79.3

For the disclosure of income information for tax purposes, we also found that 53.06% of those in the lower quartile of EI scores had disclosed this information, compared to 72.55% of those in the upper quartile. In general, we observed a higher reported rate of disclosure of sensitive information for those who scored in the upper quartile of the CC section compared to those who scored in the lower quartile; this difference was not as pronounced for other knowledge sections. We interpret this as indicating a connection between being knowledgeable about cloud computing services and concluding that the benefits of these services are worthwhile. As Table 13 shows, those who scored higher on the cloud computing section were also more likely to use some types of cloud services compared to their lower scoring counterparts.

Table 13. Quartile analysis: Types of services used

	Comparing to cloud computing scores	
Online service used	Percentage (%) of bottom quartile	Percentage (%) of top quartile
Backed up some or all of a computer's hard drive to an online site	29.4	43.4
Used an online file storage service	71.2	94.3

There were also notable differences between the quartiles for other subsections as well. In examining the total knowledge score, we found that 68% of scorers in the lower quartile have used an online file storage service, compared to 89.29% of scorers in the upper quartile. There were similar findings for the EI scores (65.38% of those in the lower quartile, 80.39% of those in the upper quartile) and OS scores (71.95% of those in the lower quartile, 88.33% of those in the upper quartile).

As we noted above, a majority of respondents to the larger survey indicated that they had previously submitted information online when they wished that they did not have to. This was also found in the smaller sample of those who completed both surveys, with 82% of respondents giving this answer (n = 194). When we completed the quartile analysis for this question, an interesting relationship appeared between the EI subsection and the CC subsection. 78.85% of those who scored in the lower quartile on the EI subsection indicated the presence of information submission regret, compared to 92% of those in the highest quartile. In contrast, 92% of those who scored in the lower quartile on the CC section experienced this regret, compared to 77.36% of those in the highest quartile. This implies that consumers who are highly

informed about cloud computing technologies may be more accepting of demands for information than those who are less informed, while knowing more about how online companies make money may have the opposite effect.

We also compared knowledge scores and opinion responses for the online behavior question concerning how often the participant reads TOS agreements. Participants were presented with five options: Always, Usually, Sometimes, Rarely, and Never. For the purpose of analyzing the smaller sample, we grouped these responses into three categories: Always/Usually, Sometimes, and Rarely/Never. In none of the knowledge subsections did the difference between the quartiles for Always/Usually pass our 10% threshold. The only notable difference between the quartiles for the Rarely/Never options was in the EI section, where 76.92% of those who scored in the lower quartile reported that they rarely or never read TOS agreements, while 56.86% of those in the higher quartile reported that they rarely or never read TOS agreements. Most of the visible differences between quartiles were among those who reported that they sometimes read TOS agreements, as shown in Table 14.

Table 14. Quartile analysis: "Sometimes" readership of TOS agreements

	I sometimes read TOS agreements	
Comparing to knowledge scores in:	Percentage (%) of bottom quartile	Percentage (%) of top quartile
Economics of the Internet	11.5	27.5
Cloud Computing	11.5	26.4
Privacy Law	7.8	27.1
Overall Score	*8.0*	*32.1*

After evaluating the quartile differences for "Sometimes" responses, we conclude that those who have higher levels of knowledge are more likely to recognize the importance of these agreements and at least consider reading them.

Overall, analysis of the online behavior responses alongside knowledge scores led to some interesting results. The most intriguing pattern to come out of the Online Behavior section is the potential predictive value of cloud computing knowledge for identifying consumers who are more willing to use more services, share more information, and have fewer regrets about it. We suspect that those who are more informed of cloud computing services are more aware of the benefits of these services and thus potentially more likely to consider the benefits to be more worthwhile when being asked to exchange their information for services.

b. Personal Privacy

The second online behavior section that we analyzed alongside the knowledge results is the Personal Privacy section.

Legal scholars are typically conscious of the dichotomy between rights and privileges. The first two questions of this section concerned this dichotomy, but the results indicate that the general population often does not make a strict differentiation between privacy as a right and privacy as a privilege. This is consistent with our

findings in the larger sample as shown in Figure 3. The responses to these questions are summarized in Figures 18 and 19.

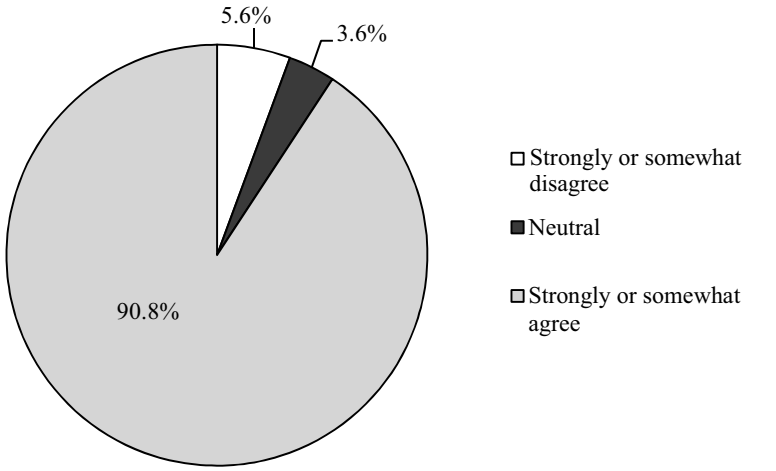

Figure 18. Personal privacy is a right to which everyone is entitled (n = 195)

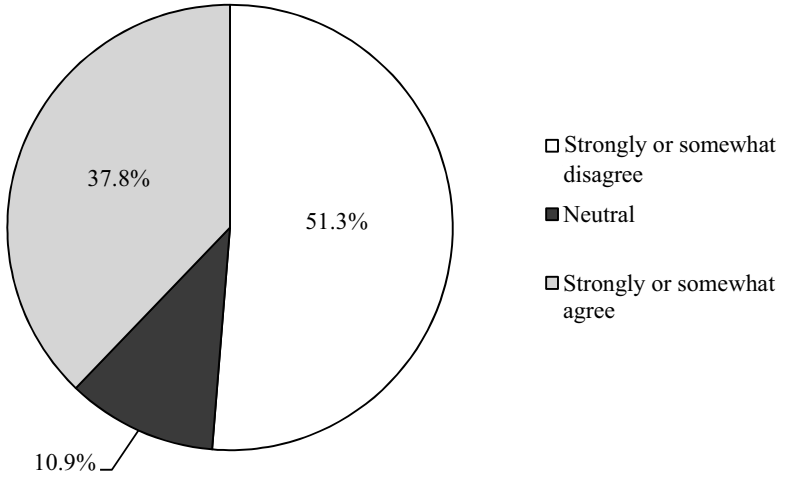

Figure 19. Personal privacy is a privilege to which people are entitled only if they follow the law (n = 193)

In the quartile analysis for the question about personal privacy as a right, there were no quartile differences greater than twenty points. On the other hand, scoring in the upper quartile in the sections about the economics of the Internet (EI), cloud computing (CC), or online security (OS) was associated with disagreeing with the statement that privacy is a privilege. Furthermore, as Table 15 shows, neutral

responses tended to be more concentrated among the lower quartile scorers, with fewer higher quartile scorers indicating neutrality on this question.

Table 15. Quartile analysis: Privacy is a privilege

	Privacy is a privilege Percentages (%)					
	Disagree		Neutral		Agree	
Comparing to knowledge scores in:	Bottom quartile	Top quartile	Bottom quartile	Top quartile	Bottom quartile	Top quartile
Economics of the Internet	32.7	59.2	19.2	4.1	48.1	36.7
Cloud Computing	37.3	59.6	23.5	3.9	39.2	36.5
Online Security	32.0	66.1	18.0	3.4	50.0	30.5
Education Records	42.6	60.0	13.1	8.0	44.3	32.0
Privacy Law	42.0	54.2	20.0	10.2	38.0	35.6
Overall Score	*34.7*	*64.8*	*20.4*	*3.7*	*44.9*	*31.5*

As Table 14 shows, the effects are less for some knowledge sections, like the ER section and the PL section, but greater in others. The largest percentage point difference between the quartiles was for the OS section, where 32.05% of scorers in the lower quartile did not believe that privacy was a privilege, compared to 66.10% of scorers in the upper quartile. Other than the selection of the disagreement options, the clearest difference between the quartiles is often that the lower performing respondents were more likely to be neutral on the issue of privacy as a privilege.

The third statement, "Personal privacy is only important to people who have something to hide," had responses that were almost an inverse of the first question concerning privacy as a right. As shown in Figure 20, 87.5% of participants in the smaller sample (n = 193) disagreed somewhat or strongly with this statement. In terms of the quartile analysis, the percentage of respondents who disagreed with this statement tended to be smaller by about twenty points for the lower quartile scorers for all knowledge sections except the ER section. Because the percentage of respondents who agreed with or were neutral toward this statement represented a small portion of the total, Table 16 only provides the percentages for the Disagree responses.

Table 16. Quartile analysis: Something to hide

	Only important if you have "something to hide"	
	Disagree	
Comparing to knowledge scores in:	Percentage (%) of bottom quartile	Percentage (%) of top quartile
Economics of the Internet	74.5	94.0
Cloud Computing	78.0	98.1
Online Security	76.0	94.9
Privacy Law	75.5	94.8
Overall Score	*72.9*	*94.4*

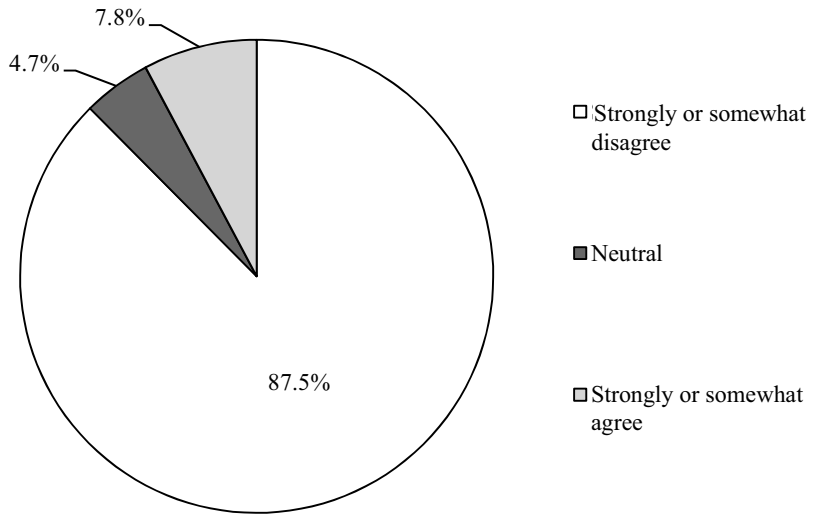

Figure 20. Personal privacy is only important to people who have something to hide (n = 193)

The responses to these three questions indicate a relationship between personal conceptions of privacy and knowledge in related areas. For instance, while most participants view privacy as a right, our analysis indicates that being familiar with the dichotomy between rights and privileges was likely to indicate higher knowledge in other areas relating to online activities and privacy.

Another statement from the Likert scale portion of this section that revealed interesting results was "If given a choice, I would prefer that my health care records were not stored online." We asked this question because it is becoming more common for patients to be able to access test results and appointment information through a service like Epic instead of relying on a phone call from a nurse or an expensive visit with a doctor. As shown in Table 17, our quartile analysis revealed that in all of the knowledge sections other than the ER section, a higher percentage of upper quartile scorers disagreed with this statement.

Table 17. Quartile Analysis: I would prefer my health records not be stored online

	"If given a choice, I would prefer that my health care records were not stored online" Percentages (%)					
	Disagree		Neutral		Agree	
Comparing to knowledge scores in:	Bottom quartile	Top quartile	Bottom quartile	Top quartile	Bottom quartile	Top quartile
Economics of the Internet	17.3	31.4	40.4	19.6	42.3	49.0
Cloud Computing	15.7	48.1	43.1	23.1	41.2	28.9
Online Security	21.3	41.7	33.8	26.7	45.0	31.7
Education Records	31.2	26.7	34.4	26.7	34.4	41.3
Privacy Law	20.0	33.9	42.0	22.0	38.0	48.1
Overall Score	*20.4*	*40.0*	*36.7*	*12.8*	*42.9*	*47.3*

The largest percentage point difference between the lower and upper quartile for this statement was found when comparing the responses to CC scores. In the above section, we noted that high scores on the CC knowledge section seem to be correlated with more permissive approaches to online data collection. The responses to this statement about health records continue this pattern, as almost half of those who scored in the upper quartile for the CC section disagreed. It is also noteworthy that for both the CC and OS sections, people who scored in the lower quartile also seemed to be more likely to not want their health care records stored online.

This statement about health care record storage also has implications for the idea of harms relating to privacy. If an individual would prefer that their health records not be stored online, this suggests that they see the potential for harm if these records were compromised. This is not to say that people who are in favor of storing this information online do not see the potential for harm. If those in favor of online storage for electronic health records (EHR) know more about cloud computing technologies or online security issues, they might reasonably believe that the risk of a breach is small and the potential benefits are significant. This raises the question of what is the harm that could follow from the online storage of EHR? If a person's EHR are released without authorization, the person's privacy has been violated. The harm that might result if one's EHR became public knowledge may be reputational, or the harm may be derived from the increased risk of identity theft. The harm might also be the chilling effect the disclosure might have on the individual's willingness to seek help for serious medical problems.

As Table 17 illustrates, the quartile differences tended to be small for those who agreed with the statement about not wanting their EHR stored online. The differences were instead more notable between the "Disagree" and "Neutral" options. Neutral responses to some other questions were also associated with lower scores in the knowledge survey. Figure 21 shows the distribution of scores in the EI section compared to responses to the privacy-as-privilege question. Figure 22 shows the distribution of the total knowledge scores compared to responses to the "something to hide" question. In the below figures, the line indicates the median, and the diamond indicates the mean. The scores in each are scaled to 100 for visualization purposes.

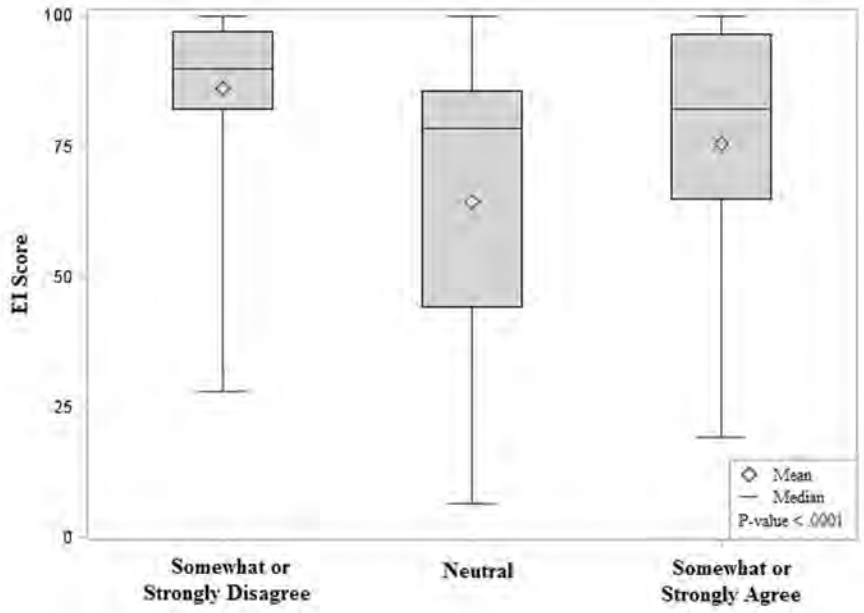

Figure 21. Distribution of EI score for "privacy is a privilege" statement

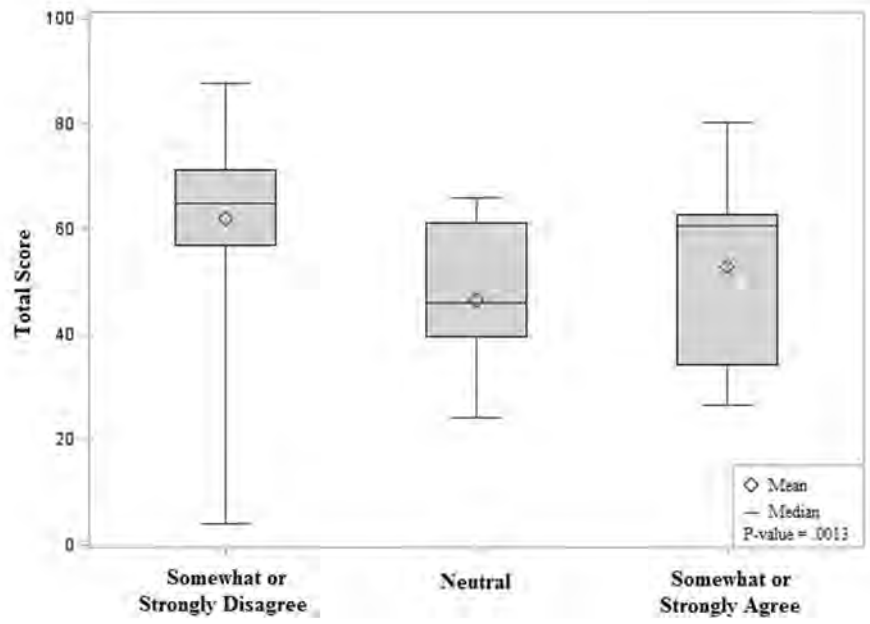

Figure 22. Distribution of total score for "something to hide" statement

Questions 11 through 14 of the personal privacy section asked participants about whether online marketing and advertising companies should be allowed to

engage in certain behaviors. Table 18 presents the results for the smaller sample of responses, which resembles the distribution shown for the larger sample in Table 1.

Table 18. Response distribution for online marketing and advertising questions

Activities of online marketing and advertising companies	Percentage (%) responding should be allowed to do	Percentage (%) responding should not be allowed to do
Track consumers' online activity without asking for permission (n = 195)	12.8	87.2
Participate in a data trade in which consumers' online activity on multiple websites is combined to create demographic profiles (n = 191)	33.0	67.0
Use demographic profiles to display advertisements that are relevant to individual consumers (n = 192)	53.6	46.4
Track consumers' interactions with advertisements to determine which ones are the most relevant to individual consumers (n = 192)	55.7	44.3

When we compared the responses to these questions with performance on the knowledge survey, we did not find any notable observations for the first two questions. In contrast, higher scores on both the CC and OS knowledge sections were associated with greater support of the business practice of using demographic profiles to display relevant advertisements. For the CC section, the quartile analysis revealed that 40.82% of those in the lower quartile agreed that companies should be allowed to use demographic profiles to display relevant advertisements, compared to 66.67% of those in the upper quartile. The quartile analysis of the OS section revealed a similar pattern, with 39.24% of those in the lower quartile supporting this practice, compared to 66.67% of those in the upper quartile who support this practice. The distributions of scores in this section were also statistically significant, as indicated by Figures 22 and 23.

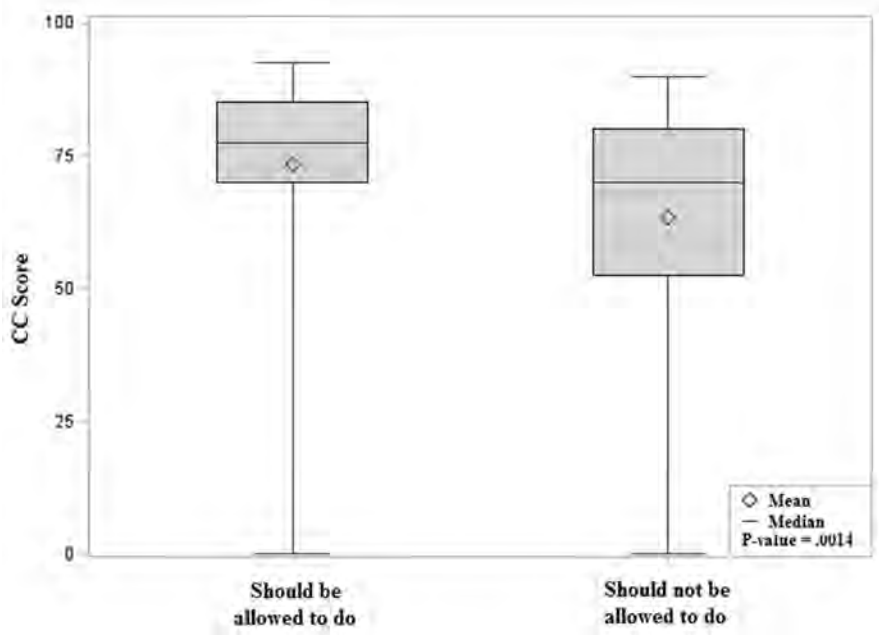

Figure 23. Distribution of CC score for the display of relevant advertisements

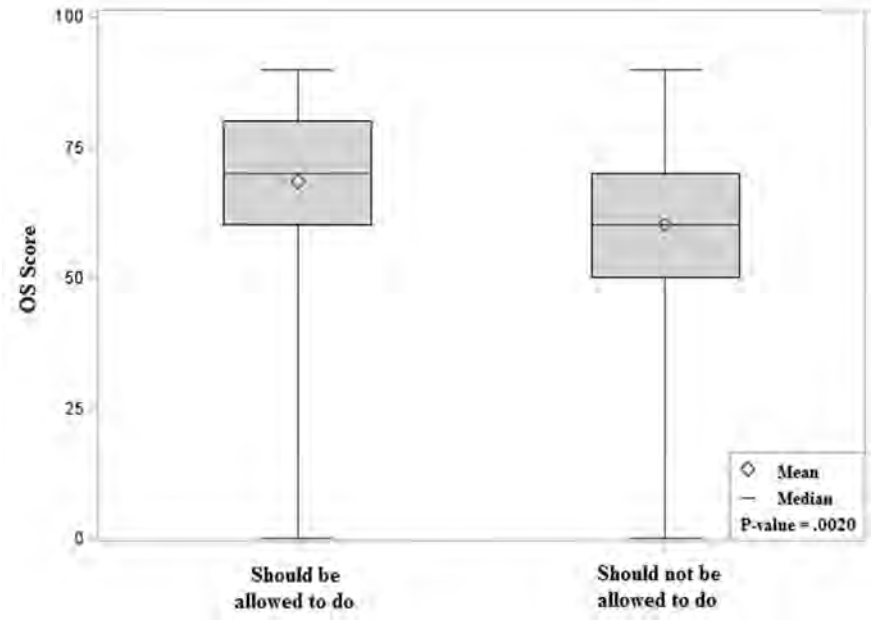

Figure 24. Distribution of OS score for the display of relevant advertisements

The final question concerning prohibiting or allowing marketing practices focuses on the use of technology to track how consumers interact with those advertisements. For the quartile analysis, the upper quartile of scorers for each knowledge section

except ER and PL were notably more likely to approve of this practice than those who scored in the lower quartile.

All of the questions in Table 14 reference the types of practices that lead to more relevant advertising. These practices often involve data brokers and go to the heart of what scholars think of when they hear the phrase "big data." It is interesting that these practices were often met with disapproval by our survey participants, even though the participants were more likely to approve of actually receiving relevant advertisements. It is also interesting that the approval of the use of profiles and technologies to display and track the effectiveness of relevant advertising tended to increase with users' understanding of topics like cloud computing and online security.

c. Cloud Service Providers

Next, we examine the opinion section that focuses on opinions about cloud service providers. The first set of questions that we will analyze alongside the knowledge survey results concerns participant knowledge of certain types of cloud provider practices and whether the participants approve or disapprove of these practices. These questions focus on the users' intellectual property and the content of their accounts. The responses for this smaller sample roughly tracked the responses for the larger sample as summarized in Table 2. Table 19 lists the awareness rates and approval rates of the larger sample and smaller sample.

Table 19. Comparing awareness and approval between samples

	Aware		Approve	
	Percentage (%) of larger sample	Percentage (%) of smaller sample	Percentage (%) of larger sample	Percentage (%) of smaller sample
Some cloud service providers may publish, publicly display, or distribute content uploaded by users	46	51	12	17
Some cloud service providers may reproduce or modify works or content uploaded by users	29	28	6	7

We conducted a quartile analysis for these responses. Those who indicated that they were previously aware of these practices were more concentrated in the upper quartile of scores for many of the knowledge sections. Among those who scored in the lower quartile of the CC knowledge section, 31.91% were previously aware that some providers might publish, display, or distribute user-uploaded content, compared to 59.18% of those in the upper quartile. Similarly, 19.18% of those in the lower quartile of the OS knowledge section were previously aware that providers may reproduce or modify content uploaded by users, compared to 39.29% of those in the upper quartile of the OS knowledge section. When comparing these responses to the quartile analysis for the total scores in the knowledge section, 42.86% of those

in the lower quartile for the knowledge survey as a whole were previously aware that user content may be searched for the purpose of displaying targeted advertisements, compared to 78.85% of those who scored in the upper quartile for the whole survey.

While knowledge levels did occasionally correlate with self-reported awareness, knowledge levels did not appear to have much connection to whether or not the participants approved of practices concerning the use of users' intellectual property. The exception to this is that quartile analysis indicated that 41.07% of those who scored in the upper quartile in the OS section and 43.14% of those who scored in the upper quartile in the PL section approve of the practice of searching user content for advertising purposes, compared to 22.58% and 26.32% of those in the lower quartile for those two sections respectively.

The next part of this section of the survey introduced the concept of trust and invited participants to indicate whether they thought that certain potential recipients could be trusted to protect confidential information. The reported rates of trust listed in Table 20 are very similar to the rates of trust found in the larger sample as shown in Table 4.

Table 20. Trusting different groups to protect confidential information

Entity	Percentage (%) responding that the entity could be trusted
Law enforcement agencies (*n* = 189)	50.8
Government, non-law enforcement (*n* = 188)	42.6
Internet service providers (*n* = 189)	16.9
Advertising agencies (*n* = 189)	3.2
Search engines (*n* = 188)	7.4
Cloud service providers (*n* = 187)	16.6
Your close friends (*n* = 192)	65.6
Your employer (*n* = 188)	55.3
Your family (*n* = 193)	80.3
Your health care provider (*n* = 188)	64.9
Your school/university (*n* = 188)	64.4

The topic of trust is the target of a lot of study in the social sciences, and the results from our survey indicate an interesting overlap between trust and knowledge, which is ripe for examination by policy professionals. For these questions, we focused on quartile analysis. We found that the level of trust of law enforcement professionals was greater for those who scored in the lower quartile of the EI section than it was for those who scored in the higher quartile of the same section. In the EI section, 40% of those in the lower quartile indicated a belief that law enforcement could not be trusted to protect confidential information, compared to 62.75% of those in the upper quartile. This suggests that there is some factor connecting high levels of comprehension of how the online market system works with low levels of trust for law enforcement. When we conducted the quartile analysis for performance on the OS section, we found a similar connection to a lesser degree. 53.16% of those who scored in the lower quartile of the OS section indicated that law enforcement could be trusted with confidential information, compared to 42.86% of those who scored in the upper quartile who believed that law enforcement could be trusted.

Results were similar for the quartile analysis of the trustworthiness of other government actors. For the EI section, 56.25% of those in the lower quartile indicated that these government actors could be trusted, compared to 31.37% of those in the upper quartile. On the other hand, those who scored in the upper quartile on the EI and CC sections were more likely to say that their friends could be trusted with confidential information than those who scored in the lower quartile on the same sections. 58.82% of those in the lower quartile in the EI section and 54.9% of those in the lower quartile on the CC section would trust their friends with confidential information, compared to 78.43% of those in the upper quartile in the EI section and 76% of those in the upper quartile on the CC section. There was not a substantial difference between upper quartile and lower quartile scorers for trusting healthcare providers and educational institutions. This indicates that participants' willingness to trust providers of these types of services is not affected by relative knowledge levels. Thus, trust of government actors may be lower for more informed participants, and trust of private businesses like Internet service providers (ISPs) and advertising companies is lower than the levels of trust that participants have in the government. Nonetheless, some segments of the private sector are consistently trusted more than the government and Internet-oriented businesses.

We also asked participants about how they would react if a cloud service provider (CSP) shared their personal information with several of the above categories of actors, either by giving away the information or by selling the information, without the subject of the information being notified first. Our findings are provided in Table 21.

Table 21. Giving away or selling information to . . .

	Approval rate percentage (%)	
Entity	Give away	Sell
Law enforcement	28.1	3.6
Other government agency	18.8	4.2
ISPs	4.7	3.1
Advertisers	5.7	7.3
Search engines	5.8	6.2
Your employer	6.8	1.6

The biggest difference seen in Table 21 is the substantial drop in approval for selling information to law enforcement agencies. This is consistent with the data in Table 5 for the larger sample.

In the quartile analysis, approval of this information being given to law enforcement echoes the pattern seen above. 30% of those who scored in the lower quartile of the EI section would approve of the provider giving the information to law enforcement, while only 15.69% of those in the upper quartile would approve. On a separate question, those who scored in the upper quartile of the EI section also indicated more awareness that cloud providers do sometimes sell information than those in the lower quartile (98% compared to 80.43%).

We also asked participants if they believed that privacy policies should be standardized. Of 186 respondents, 42.47% answered that allowing providers to use different privacy policy templates is acceptable, while 57.53% of respondents

indicated that there should be a law that requires standardized privacy policy templates. The quartile analysis for this question is presented in Table 22.

Table 22. Quartile analysis: Should there be a law?

	Responding yes, there should be a law requiring standardized privacy templates	
Comparing to knowledge scores in:	Percentage (%) of bottom quartile	Percentage (%) of top quartile
Economics of the Internet	60.0	46.9
Cloud Computing	56.5	47.1
Online Security	68.1	5.8
Education Records	63.8	59.5
Privacy Law	71.7	51.7
Overall Score	*65.1*	*44.4*

We also asked participants to consider whether there should be a law to prohibit cloud service providers from using customer information for targeted advertising purposes. 58.58% of respondents in the smaller sample think that permitting the use of information for these purposes is acceptable, while 41.42% of respondents indicated that there should be a law to prohibit this practice. As shown in Table 23, quartile analysis revealed that those in the lower quartile in the knowledge subsections were more likely to support such a law than those in the upper quartile.

Table 23. Quartile Analysis: Should there be a law?

	Responding yes, there should be a law to prevent CSPs from using customer information for targeted advertising	
Comparing to knowledge scores in:	Percentage (%) of bottom quartile	Percentage (%) of top quartile
Economics of the Internet	54.8	35.6
Cloud Computing	52.4	27.1
Online Security	57.1	22.5
Education Records	50.0	42.2
Privacy Law	52.3	35.2
Overall Score	*56.1*	*31.4*

The data in these tables are somewhat surprising because they suggest that those who scored lower were more likely to be in favor of government regulation of privacy policies and online advertising. The mean scores for those in favor of a new law were also lower, as shown by Figure 25.

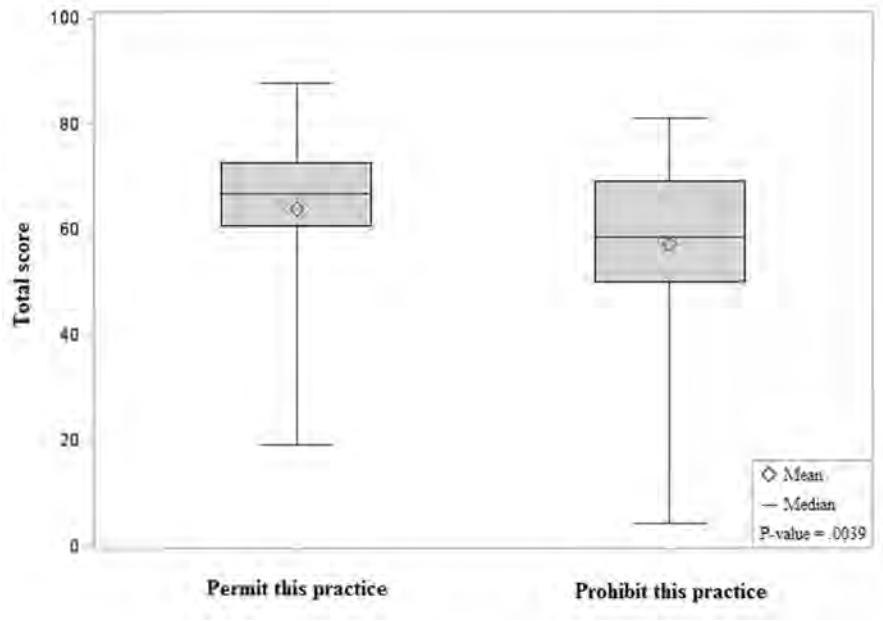

Figure 25. Should companies be allowed to use customer information for targeted advertising?

Some of these results suggest that people who are more informed about issues relating to data online may be less trusting of the government and government intervention on these topics. Policymakers are therefore faced with a dilemma. Our survey results indicate that consumers are often displeased with the options available in the private market, but they feel helpless to change the system. The primary actors in the online economy have precious few incentives to change, because the consumers are willing to use the services anyway, even though consumers do not trust the service providers, and the services are deficient in terms of what the consumers want. Because reasonable alternatives do not exist, the companies can therefore continue to benefit even in the face of this market failure. Policy intervention, on the other hand, would likely be met with skepticism and distrust, especially from those who are already more informed about issues relating to online activities.

d. Government Surveillance and Privacy Law

The final section that we analyze alongside the knowledge results is the section focusing on privacy issues relating to law and government actions.

The first question that we asked in this section concerns potential targets of government investigations and about whom the government should have the ability to collect and analyze private information. Figure 26 presents the pie chart for the results of the smaller sample.

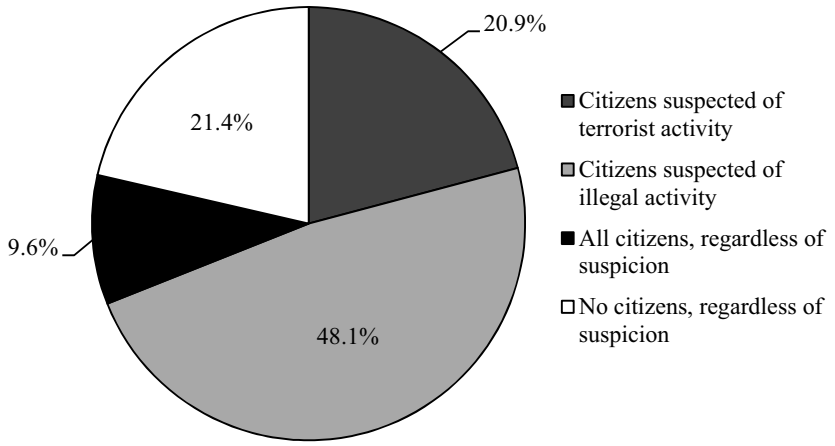

Figure 26. The government should be able to collect and analyze private information about: (*n* = 187)

In terms of the broader distribution of knowledge scores, the greatest significance for this question was found in the distribution of scores for the CC section, as Figure 27 shows.

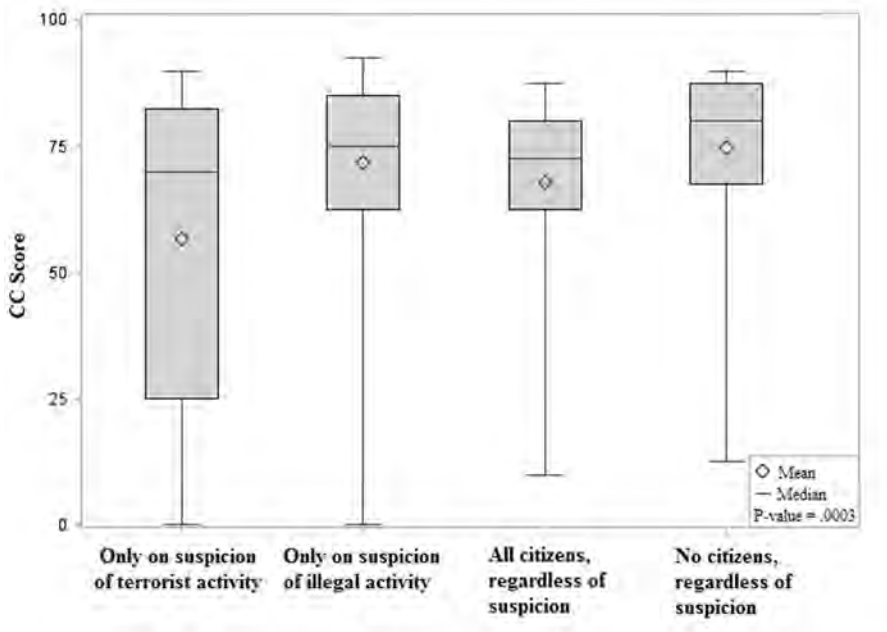

Figure 27. Distribution of CC score for question about government power to collect information

We also asked participants about the amount of power that the U.S. government currently has to collect and analyze citizens' private information in the name of public safety. Only 10 of the 181 respondents indicated that they believed the government had too little power to engage in these activities, and the mean score of those 10 participants in the EI section was significantly lower than the mean EI score for participants who selected the options stating that the government either has "sufficient" or "too much" power for this purpose.

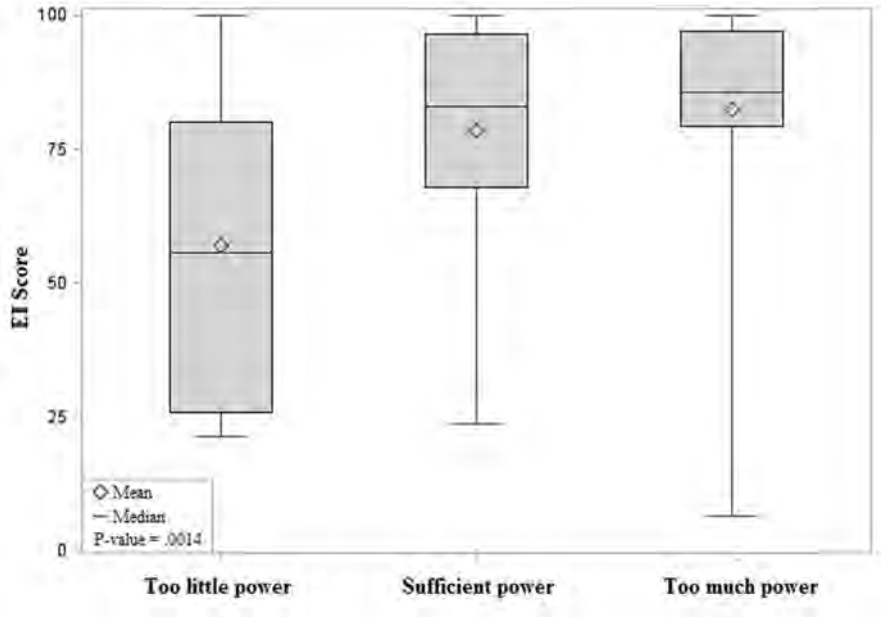

Figure 28. Distribution of EI score for question about government power

In Table 6, we summarized our findings from a series of questions exploring the effect of the presence or absence of a warrant on participants' opinions concerning whether a target's consent should be obtained before a search and whether the target should be notified after a search. The responses to these questions in the smaller sample resemble the patterns found in the larger sample. We identified notable differences in the quartile analysis for the question concerning notification after a warrantless search, as shown in Table 24.

Table 24. Quartile analysis: Notification after warrantless search

	No warrant			
	Google should notify Bill		Government should notify Bill	
Comparing to knowledge scores in:	Percentage (%) of bottom quartile	Percentage (%) of top quartile	Percentage (%) of bottom quartile	Percentage (%) of top quartile
Economics of the Internet	52.1	67.4	39.6	24.5
Online Security	48.7	60.3	38.5	25.9

Table 24 indicates that those who have a higher level of knowledge in some areas may be more willing to place the obligation of notification on the private company than to place that same obligation on the government. This may be related to the findings from the Cloud Service Provider section above, where we found that participants who performed better on the EI and OS knowledge sections were less inclined to trust law enforcement with their confidential information.

On the other hand, when the government presents Google with a warrant, survey participants were more inclined to conclude that Bill's consent was not necessary, and a greater number of participants who thought that Bill should be notified believed that the government should be responsible for the notification. Our analysis also indicated a connection between participants' views about the search with a warrant and their knowledge levels.

As Figure 29 shows, the mean score in the Cloud Computing (CC) section for those who thought Google should obtain Bill's consent before complying with the warrant was lower than the mean CC score for the other two potential responses.

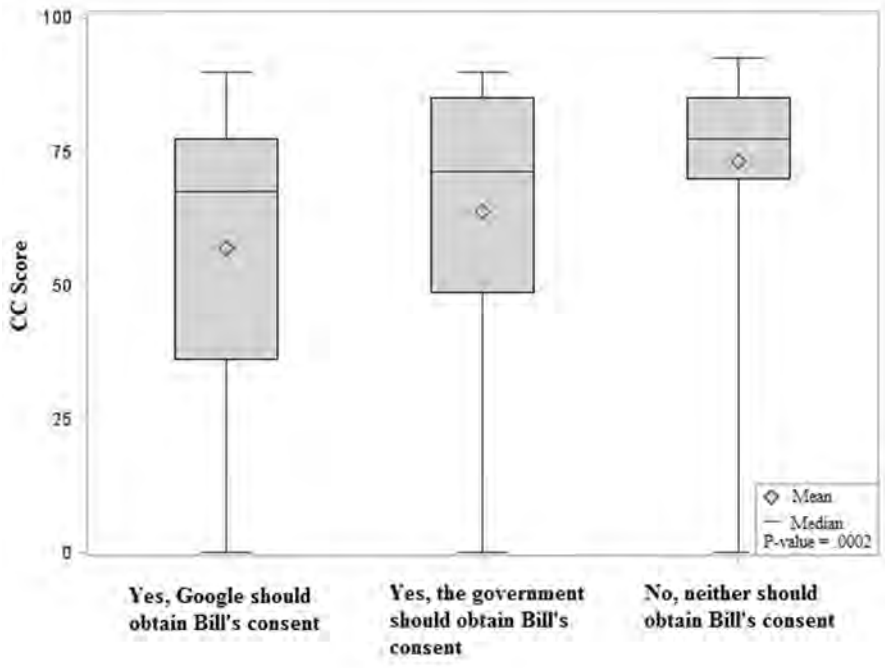

Figure 29. Distribution of CC score for question about obtaining consent before complying with or executing a warrant

As shown in Figure 30, the mean total score for participants who did not think Bill's consent was necessary when there was a warrant was also higher than the mean total score for participants who thought that either Google or the government should obtain Bill's consent.

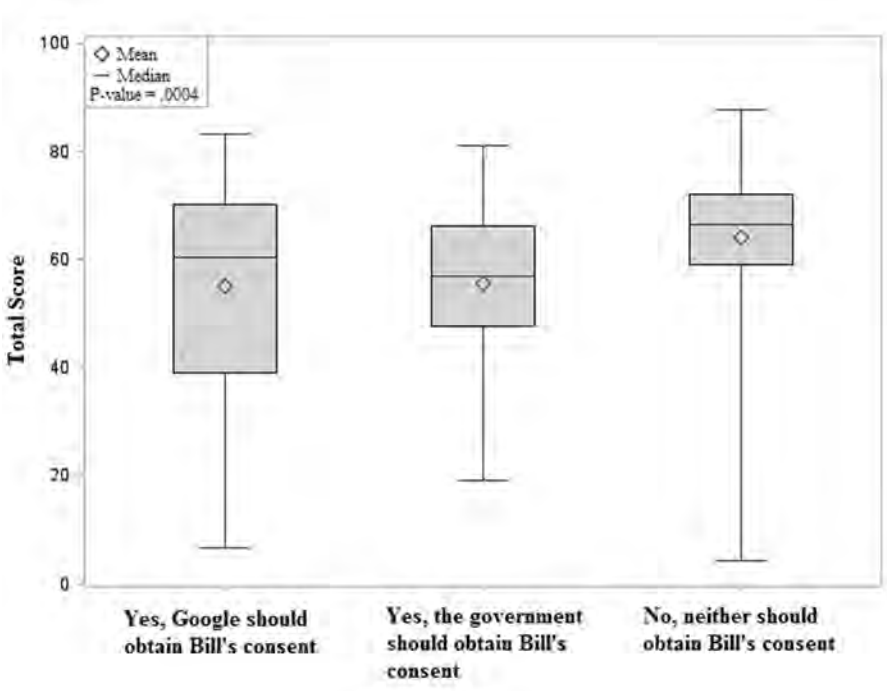

Figure 30. Distribution of total score for question about obtaining consent before complying with or executing a warrant

Following the questions about surveillance of Bill's Gmail account, we turned the topic on the participants, who were asked to imagine that their e-mails and phone calls were being monitored in the same way that Bill's Gmail account is being monitored. We further built on the question by informing the participants that a court has approved of the monitoring program, so the program may continue to operate without a warrant and without the knowledge or consent of those whose information is being monitored. The results are summarized in Table 25.

Table 25. Response to hypothetical court decision

Percentage (%) responses to the effect of a hypothetical court decision that the program is legal: (n = 185)	
Makes you feel more comfortable with the monitoring program	9.7
Makes you feel less comfortable with the monitoring program	36.2
Has no effect on your feelings about the monitoring program	13.5
Has no effect on your feelings about the monitoring program, and makes the court seem less trustworthy	40.5

Upon completing the quartile analysis, we found several notable results, which are listed in Table 26.

Table 26. Quartile analysis: Response to hypothetical court decision

	Court held surveillance program to be constitutional							
	More comfortable with the program		Less comfortable with the program		No effect		Less comfortable with the program and court less trustworthy	
Comparing to knowledge scores in:	Percentage (%) of bottom quartile	Percentage (%) of top quartile	Percentage (%) of bottom quartile	Percentage (%) of top quartile	Percentage (%) of bottom quartile	Percentage (%) of top quartile	Percentage (%) of bottom quartile	Percentage (%) of top quartile
Economics of the Internet	10.2	10.4	42.9	31.3	16.3	2.1	30.6	56.3
Cloud Computing	8.3	14.0	43.8	24.0	12.5	18.0	35.4	44.0
Online Security	10.8	7.0	40.5	28.1	14.9	10.5	33.8	54.4
Education Records	10.5	4.2	38.6	40.3	15.8	15.3	35.1	40.3
Privacy Law	6.4	8.8	40.4	26.3	17.0	17.5	36.2	47.4
Overall Score	*13.0*	*11.3*	*39.1*	*30.2*	*13.0*	*13.2*	*34.8*	*45.3*

Some of the findings in the above table are again illustrative of the idea that those with higher knowledge levels may have a lower level of trust of the government, though this is not seen in all tested knowledge areas.

Because a lot of analysis in the context of the Fourth Amendment and the SCA turns on whether the information sought is content or noncontent information, we also asked participants to indicate whether they approved or disapproved of a variety of information being collected through warrantless searches in the interest of identifying terrorist activity. The results for this sample are summarized in Table 27.

Table 27. Government monitoring of different types of information

Type of information	Percentage (%) approving	Percentage (%) disapproving
The date and time text messages are sent or received (n = 183)	50.3	49.7
The "to" and "from" fields of text messages (n = 186)	41.4	58.6
The content of text messages (n = 186)	10.8	89.2
The date and time phone calls are sent or received (n = 183)	51.4	48.6
The "to" and "from" fields of phone calls (n = 185)	38.9	61.1
The content of voicemails (n = 188)	11.2	88.8
The date and time e-mails are sent or received (n = 184)	51.1	48.9
The "to" and "from" fields of e-mails (n = 187)	41.2	58.8
The content of e-mails (n = 188)	12.8	87.2

Somewhat predictably, fewer participants indicated that they would approve of the warrantless collection of content information. For the questions concerning the date and time of communications, only one quartile comparison resulted in greater than 20 percentage points difference between the lower and upper quartile. That comparison is for the EI group, where 55.32% of the lower quartile approved of warrantless collection of the date and time that phone calls were sent or received, compared to 34.69% of those in the upper quartile for the same group. Similarly, only one quartile comparison of responses to the "to" and "from" fields questions exceeded 20 percentage points. That comparison is again for the EI group, where 45.83% of scorers in the lower quartile approved of the warrantless collection of information from these fields in text messages, compared to 24.49% of those in the upper quartile. No other quartile differences greater than 20 percentage points were found for options concerning the contents of communications.

Two of the final questions that we asked in this section focused on the possibility that the Supreme Court might conclude that there is no reasonable expectation of privacy in information that consumers allow to be used for advertising purposes. Figure 31 lists the results for this question.

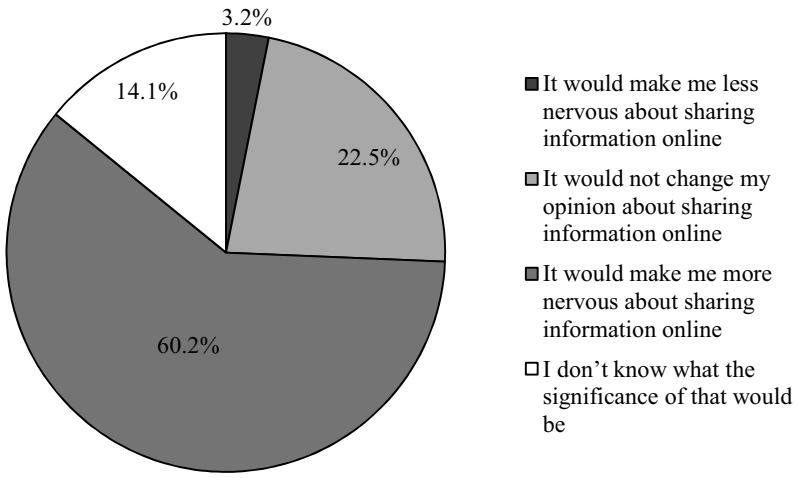

Figure 31. "Reasonable expectation of privacy" change hypothetical (*n* = 191)

First, we asked participants to indicate how such a decision would affect their feelings about sharing information online. The results from the quartile analysis are provided in Table 28.

Table 28. Quartile analysis: Reasonable expectation of privacy

	Court finds no reasonable expectation of privacy in online information also used for ad purposes							
	Less nervous about sharing		No change		More nervous about sharing		Don't know the significance	
Comparing to knowledge scores in:	Percentage (%) of bottom quartile	Percentage (%) of top quartile	Percentage (%) of bottom quartile	Percentage (%) of top quartile	Percentage (%) of bottom quartile	Percentage (%) of top quartile	Percentage (%) of bottom quartile	Percentage (%) of top quartile
Economics of the Internet	4.1	-	26.5	12.2	55.1	77.6	14.3	10.2
Cloud Computing	8.0	2.0	26.0	27.5	44.0	68.6	22.0	2.0
Online Security	5.2	-	24.7	22.0	51.9	67.8	18.2	10.2
Education Records	4.9	2.7	32.8	16.2	50.8	62.2	11.5	18.9
Privacy Law	6.1	-	26.5	20.7	51.0	67.2	16.3	12.1
Overall Score	*4.2*	-	*33.3*	*18.5*	*47.9*	*70.4*	*14.6*	*11.1*

As Table 28 shows, scoring higher on the knowledge survey was typically associated with being more wary of this sort of change in the law.

The second part of this question examined what the effect of this increased nervousness might be. Specifically, we asked about whether the hypothetical Supreme Court decision would change participants' behavior towards online agreements that might address the use of personal information for advertising purposes. The results are presented in Figure 32.

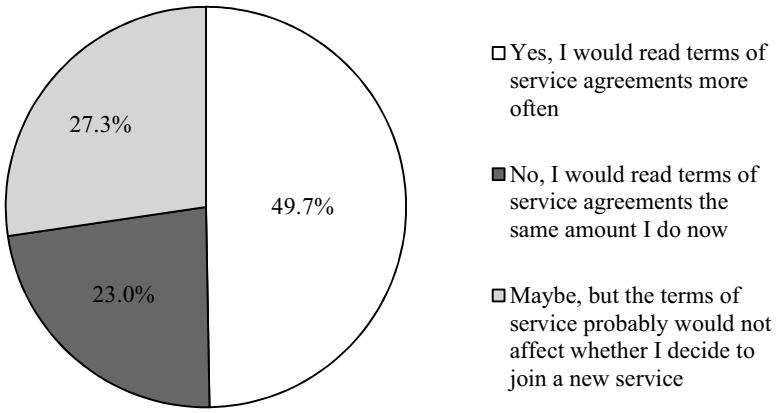

Figure 32. Would such a decision affect TOS readership? (n = 183)

When we conducted the quartile analysis, only the EI group had quartile differences greater than 20 percentage points. There, 40% of those in the lower quartile indicated that they would read TOS agreements more often, compared to 62.5% of those in the upper quartile. Furthermore, 42.22% of those in the lower quartile of the EI section selected the "Maybe" answer, compared to only 14.58% of those in the upper quartile. There were also some quartile differences that were greater than 10 percentage points but less than 20. For example, 44.68% of those who scored in the lower quartile on the CC section indicated that they would read these agreements more often, compared to 58.82% of those who scored in the upper quartile. When considering total knowledge scores, 43.18% of those in the lower quartile said they would read TOS agreements more often under these circumstances, compared to 60.38% of those in the upper quartile. In other words, that sort of hypothetical change in the law could result in increased TOS readership for some participants who already have higher knowledge levels. However, we would not recommend that option as a path to increasing consumer engagement.

E. Implications

The overarching goal of our study is to evaluate the interactions between consumer knowledge, behavior, and opinions regarding online privacy. The results of our survey indicate that potentially millions of consumers have inadequate knowledge to make meaningful choices about how their data is used online. Table 7 shows that 75% of participants correctly identified the description of cloud computing when prompted, but that means that a quarter of the participants in our

survey were confused about cloud computing. About 40% of participants had never heard of FERPA even when they were told the full name of the law and the acronym. As Table 11 shows, 32% of participants thought that they could amend their agreement with Facebook by posting the "amendment" to their Facebook wall.

Ultimately, we reject the simplicity of a "privacy paradox" for explaining away why privacy actions do not always line up with privacy preferences. The "privacy paradox" is only paradoxical if analysis is limited to the disconnect between what people do and what people want. There are many more factors in play. Our findings suggest that consumers often go along with data practices that they disagree with. For example, in Figure 1, we note that 81% of respondents had previously submitted information online when they wished that they did not have to do so. There are many possible explanations for this. First, they may do this because they believe that the benefits of using the service outweigh the downsides of using the service. Or they may believe that the benefits from using the service are greater than the benefits from *not* using the service. Another possible explanation is that the consumers do not know enough to make meaningful decisions about their online privacy. In the alternative, maybe consumers know enough but feel helpless to make a decision that differs from what companies are willing to offer.

Introducing new business models can, at least partially, address the problem of consumers feeling helpless when it comes to changing how their data is used online. As Figure 5 illustrates, almost one in five respondents expressed a willingness to pay a small amount for a service in exchange for greater control over their information. This demonstrates that there is a demand for more meaningful choices regarding digital privacy. While a majority of respondents would still prefer to have fewer features but more control for free, roughly twice as many respondents would rather pay for full access to a service with money, compared to those who would rather pay for the same access with their personal information. This finding is supported by a recent study conducted in Brazil[175] and indicates that there may be a considerable demand for alternative business models. Results like these provide valuable insights into respondents' interactions with Internet issues, and application of these lessons could lead to new and lucrative market-development opportunities. Further research should explore the attractiveness of paid services when users' control over their personal information is the primary selling point.

Our results also indicate that consumers are likely to think that privacy laws are more protective than they are in reality. When we designed the knowledge questions about legal aspects of privacy, we provided answer choices that referred to different levels of regulation. The most popular incorrect answers were often those that represented stricter privacy laws. With both FERPA and general privacy law questions, the trend appears to be towards believing that the law is more protective than it actually is. There are some exceptions, however. In the Economics of the Internet section of the survey, one of the true/false questions included the statement "Online advertising companies are required by law to ask for permission before tracking user activities." 14.1% of respondents said that was true and 67.3% of

175. Kellyton dos Santos Brito, Frederico Araujo Durao, Vinicius Cardoso Garcia & Silvio Romero de Lemos Meira, *How People Care About Their Personal Data Released on Social Media*, 2013 INT'L CONF. ON PRIVACY, SECURITY & TR. 111.

respondents said that was false. 17.6% of respondents marked the option "Don't know." This suggests that most respondents are at least aware that online advertising practices are not closely regulated.

This raises the question of addressing the discrepancy between expectations and reality for privacy law. Consider the SCA, for example. 55.6% of respondents incorrectly answered that a private citizen who wants information about a subscriber to a service would need to file a police report and have the police request that information. Companies may have a variety of different policies about when they will share certain types of information, but that is something that the company decides and is not something required by law. One possible way to address this discrepancy is to work on educating people about the limitations of privacy law and explaining that such laws often do not cover everything that people might think they should. This could lead to more public support for attempts to amend or enact privacy legislation.

One of the recurring themes of privacy literature concerns the type of harm that results from a privacy violation. In Part I, we considered several possible types of harm, including dignitary harms; a chilling effect from law enforcement having too much control over individual expression; and circumstances that interfere with an individual's ability to exercise freedoms or develop a sense of self-determination. Some of the results from our survey suggest that multiple harms may be part of the justifications that participants use for why privacy should be protected. For example, the results in Table 2 and Figure 4 show that participants placed a high value on privacy for protecting secrets and also for preventing law enforcement from targeting individuals with unreasonable searches. Survey participants soundly rejected the idea that privacy is only important to those with something to hide, so the reasoning for wanting to protect secrets and be protected from government overreach must be something broader. The desire to keep some things secret, for example, may be because participants want to avoid reputational harm, while the preference against law enforcement having excessively broad investigative powers may be related to the aforementioned chilling effect rationale.

The chilling effect is not equally persuasive to everyone, but other factors may reduce support for the potentially privacy-infringing behaviors of law enforcement. As shown in Table 5, approximately 31% of participants indicated that they would approve of cloud service providers giving their personal information to law enforcement, but that approval rate dropped to below 8% when the question involved service providers *selling* personal information to law enforcement. One interpretation of this substantial drop in approval is that even though both actions would have the same end result, the idea of commodifying consumers by selling their information is more offensive to participants' desire for self-determination. If correct, this provides additional support for Cohen's argument that preventing harms to self-determination is a major goal of privacy.

Table 2 and Figure 4 also explored the nature of privacy. Two of the questions concerned privacy as data control and privacy as secrecy. Some academics have spoken against the idea of privacy as focusing on secrecy. Solove warns that too much emphasis on this "'secrecy paradigm'" leads to the misguided conclusion that someone's privacy has not been violated unless something which was secret has been

exposed.[176] Solove refers to the harm caused by being deprived of control over data as a dignitary harm.[177] However, while a majority of participants indicated that data control is highly important to them, more participants expressed that secrecy was of the highest importance. An examination of the table and figure shows that participants were slightly more likely to choose an ambivalent response like a score of five for the importance of data control. This may indicate that people are less worried about the privacy of nonsensitive information. We should note, however, that these questions were presented prior to the questions that explored whether or not participants knew about and approved of various data broker practices involving nonsensitive information. It is possible that participants would have indicated a stronger desire for data control after being informed of several common permissions included in privacy policies.

We next turn to the second half of our analysis, where we identified over 200 participants who completed both parts of the survey. While the larger samples are useful for observing patterns of responses, comparing the outcomes of both surveys allows us to see how knowledge and opinions interact in the context of online privacy.

Above, we noted that there could be many factors involved in the so-called privacy paradox. When we began examining the interactions between the knowledge and opinion surveys, we initially expected to find that lower scores would be associated with behavior that is inconsistent with placing a high value on data privacy. In other words, we expected that the paradox might be explained because people knew too little. In some situations, we found the opposite effect. In conducting quartile analyses, for example, we found that participants with high Cloud Computing scores were sharing more types of information and using more services than their lower scoring counterparts. These findings are presented in Tables 12 and 13. This finding largely supports our "balancing of interests" alternative explanation. When people are more informed about cloud services, this may also mean that they are more aware of the benefits of these services. They then might conclude that the benefits outweigh the downsides of the service or that the benefits of using the service outweigh the benefits of not using the service. Future research could examine this interaction in more detail.

Responses in these surveys indicated both a low level of approval of common online data practices and a low level of trust of the government as a data holder. Relevant response levels are presented in Tables 19, 20, and 21. Our findings concerning trust are further supported by the work of Xu's research team, who found that study participants who had a higher trust of Facebook were more likely to accept broad privacy permissions from Facebook applications by default.[178] As we note in Part I, transparency is a goal of many reform suggestions, and we posit that the lack of transparency contributes to the low level of trust of public and private holders of data. Consumers are stuck between a rock and a hard place. Those who are more knowledgeable about online data issues seem to realize that being more knowledgeable does not currently give them any advantages because they do not

176. Solove, *supra* note 48, at 497.
177. *Id.* at 522.
178. Xu et al., s*upra* note 139, at 4.

have any meaningful alternatives. The results of these surveys are sometimes frustrating because, in examining the responses, we heard an echoing refrain of "Why bother?" Why bother reading a company's privacy policy if the company is just going to do what it wants anyway? Why bother avoiding services because of unattractive data practices when a company with 500 million users would hardly notice the loss of one? Why bother trusting the government when mass surveillance practices suggest that they don't trust you?

As shown in Figure 32, almost half of the participants in our smaller sample said that they would start reading TOS agreements more if the Supreme Court decided that there was no reasonable expectation of privacy in information that consumers allow to be used for advertising purposes. Short of the fear of losing essential liberties, it is difficult to determine what else would prompt people to demand more involvement in decisions about their data online. Americans deserve better than to be this jaded about data privacy. A major overhaul of how digital information is handled is necessary. A modest proposal to help consumers educate themselves more would just be a drop in the ocean. What is needed is a total reversal of course in terms of how information is handled online. Transparency and notice must be included in this, but more importantly, we must rid our online culture of this refrain of "Why bother?"

IV. RECOMMENDATIONS

In this Part, we offer recommendations that will make privacy more accessible to those who desire it. Survey responses indicated a low level of trust of the government as a data holder, especially among more knowledgeable survey participants, and also revealed that trust levels for many private data holders are even lower. However, the trust deficit does not exist equally in all sectors, as many participants indicated that they would trust healthcare providers and educational institutions to protect their personal data. To address this trust deficit, urge transparency, and empower consumers to make the privacy decisions that they want to make, rather than the decisions they feel like they have to make, a trusted third party should be introduced to the equation. We begin this section by discussing our most ambitious recommendation: a complete overhaul of data privacy law and the creation of profile repositories that we call Profile Information Reporting Agencies (PIRAs). As a market-based clearinghouse solution, PIRAs will allow consumers to manage their profile information as used for marketing purposes in a way similar to how credit report agencies allow consumers to monitor activities that affect their credit.

A. PIRAs and Privacy Law Reform

Modern privacy law in the United States often focuses on mandatory notice requirements, relying on the fiction that if customers are told about the uses of their information, they will vote with their feet if they do not like the terms. But consumers of services do not have the time or energy to read all of the mandatory disclosures that are shared with them.[179] Even if they did, they are left with the strong impression

179. Ben-Shahar & Schneider, *supra* note 20.

that companies and the government will continue to use their information in whatever ways they want, leaving consumers without a meaningful way to exercise control. President Obama has recognized the promises and pitfalls of the world of big data in terms of both the private sector and the government,[180] and the time is ripe for the enactment of comprehensive federal data privacy legislation.

Our previous research called for baseline regulations that emphasize protection for personally identifiable information and ensure that consumers are able to control their data.[181] After conducting a detailed survey to examine knowledge, opinions, and behaviors regarding online privacy, it appears likely that privacy regulations will be met with skepticism if adequate controls are not assured. Participants in our survey seem to want to be involved in the choices made about their data, but the current paradigm makes this difficult. To this end, data privacy legislation should emphasize a two-part approach. Consumers should be assured that their information will be shared only with their consent through an opt-in system, and they should also have the ability to view, challenge, and remove this information. This increase in transparency is likely to lead to an increase of trust and thus move the consumer's relationship with data holders from complacency to consent.

Because of the current trust deficit, however, any legal reforms should focus on consumer protection and trusted third parties, with the government playing an oversight role. Such reforms could follow the model provided by the Fair Credit Reporting Act (FCRA).[182] Many of the provisions of FCRA apply to credit reporting agencies, which are private companies that compile consumer credit information. FCRA gives consumers the right to access their credit reports annually and also to challenge incomplete or incorrect information in the report.[183] FCRA also limits the disclosure of credit reports by using an opt-in model.[184] The oversight provided by FCRA arguably helps ensure that the private companies providing credit reporting services can build and maintain the public's trust. Enabling consumers to stay informed about credit activity associated with their identities is extremely valuable. We argue that it is also valuable to allow consumers to stay informed about the arguably less sensitive information that shapes their experiences in the modern economy. Trusted third parties and a legal oversight regime would ensure that control over this type of information can be used consistently and meaningfully.

It is for this reason that we recommend PIRAs, which we envision as independent organizations that serve as a place of cooperation between the individual, the private sector, and the government. Comprehensive privacy legislation could recommend the initial organization of PIRAs, establish government subsidies for PIRAs, and provide guidelines concerning the situations when a consumer's profile information can be shared with entities other than the consumer. This legislation should also

---

180. *See* EXEC. OFFICE OF THE PRESIDENT, BIG DATA: SEIZING OPPORTUNITIES, PRESERVING VALUES (2014), *available at* http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf [http://perma.cc/2QQ5-RQ5F].

181. Kesan et al., *supra* note 18, at 462–71.

182. 15 U.S.C. § 1681 (2012).

183. CONSUMER FIN. PROT. BUREAU, A SUMMARY OF YOUR RIGHTS UNDER THE FAIR CREDIT REPORTING ACT 1, *available at* http://files.consumerfinance.gov/f/201410_cfpb_summary_your-rights-under-fcra.pdf [http://perma.cc/8K4S-2ZK5].

184. *See* 15 U.S.C. § 1681b (2012).

address acceptable data practices for both the government and the private sector and should also include amendments to existing laws pertaining to online data. For example, this legislation should amend the Stored Communications Act to bring it up to date. Thirty-year-old distinctions between providers of electronic communication services and providers of remote computing services are not helpful when dealing with the Internet today, and the provisions concerning e-mail should be updated to consider today's ubiquitous e-mail access and perpetual web-based e-mail storage.

As proposed, PIRAs would provide a location where an individual consumer could register and view aggregated profile information about herself. A consumer who wants to challenge or remove certain information from her profile should be given the option to do so, either permanently or temporarily. Consumers would likely benefit from the increased transparency and the ability to be informed about how their online decisions affect their experiences. The FTC has called for congressional action to make the practices of data brokers more transparent. Our recommendation goes further than the data broker industry and would also require the same level of transparency and control for most nonsensitive information that the government holds about citizens. PIRAs should be a place of cooperation between the private sector and the government, but they should act independently of either.

One of our arguments for giving consumers the ability to view, challenge, and remove data from their profile is an ethical one. Consider, for example, a woman who desires to become a mother, but who has had many miscarriages. One of the most profitable markets is pregnant women and mothers, and she likely has been considered part of this market ever since her first purchase of prenatal vitamins. For this individual, an endless stream of advertisements about pregnancy and children is not only ineffective for the advertiser, it is cruel to the target. PIRAs would allow her to remove aspects of her profile that could trigger emotional distress, giving her the freedom to choose how she represents herself.

Part of the inspiration for this recommendation comes from the website AboutTheData.com, which the data broker Acxiom recently launched to allow consumers to view and correct the profile information that Acxiom has accrued about them.[185] However, we strongly urge against placing the control of such a massive database with data brokers themselves or with the government. The results of our survey indicate that many people strongly disapprove of the current practices of data brokers and marketing companies, and that there is a high level of distrust of the government as a keeper of data, especially among consumers who are more educated about the relevant technology and economic activities. If PIRAs are to be effective, people must trust PIRAs to protect their data, and our survey results show that data brokers and the government both fail in this respect.

PIRAs could have benefits for businesses as well as consumers. PIRAs would allow the consolidation of hundreds of millions of profiles. Marketing companies could benefit by purchasing access to profiles fully vetted by the consumers themselves. Consumers could receive more targeted marketing materials that are actually relevant to their needs, taking full advantage of the personalization benefits

185. ABOUTTHEDATA.COM (2014), https://www.aboutthedata.com/ [https://perma.cc/XF8N-SPXA].

that cloud services provide. The results of our survey suggest that people who know more about cloud computing are more willing to use more of these services and share more of their information. We argue that this is because they recognize the benefits provided by these services. We believe that if you give consumers more control, and centralize this control so that they do not have to track their profile through multiple services, they will take advantage of it in a way that benefits themselves and businesses.

A caveat is warranted, however. A repository like this also has the potential to exacerbate information insecurity issues because it creates a single point of failure. For this reason, profile information should be accessible solely by the person to whom it applies. Furthermore, no sensitive information should be held by this clearinghouse. There is no need for specific account numbers, addresses, or passwords to be maintained in these profiles. If the consumer has a magazine subscription, that could be in the profile, but not the address to which the magazine is mailed. Some information, like date of birth, could be known by the clearinghouse, but any lists provided to marketing firms would only include the relevant demographic age range for the profile. There are many more rules and frameworks that should be established in order to make this data privacy clearinghouse a reality, and future research should explore the options. Future research should also be conducted to evaluate the interaction of this proposal with European privacy law, which is much more protective of data privacy than current laws in the United States.

B. Educational Programs

Our survey results demonstrate the variation in knowledge levels across topics, as well as how these knowledge levels intersect with various positions a consumer might take concerning online privacy. In Part III.E, we noted that knowledge of privacy law tends to be fairly low and that our survey respondents often seemed to think that privacy laws are more protective of consumers than they really are. We argue that increased education on these topics would bring benefits to the marketplace and consumers alike and would also increase support for a substantial revision of data privacy law as we proposed in Part IV.A.

There are a number of possible benefits for increasing knowledge about each of the categories that we tested in our knowledge survey. For the cloud computing section, we found that people who scored high on this section were more likely to share more information and use more cloud-based services. It is thus possible that educating the general public about these services and the benefits of these services would increase revenue for online advertising and other online businesses. Educational efforts that focus on online security and the economics of the Internet, on the other hand, would make consumers more informed about the full context and risks of online services. This would provide a good complement to the benefits of increased cloud computing knowledge because consumers would be more informed about companies' business practices and the possible risks of using these services. This would allow consumers to more effectively balance the risks and benefits of their possible choices. Finally, ensuring that consumers are more informed about privacy laws and what these laws permit and prohibit would give consumers a more realistic idea of the regulatory environment in which online businesses operate. If

consumers are more informed about the current state of the law, attempts to amend privacy laws or enact new privacy laws could elicit more public support.

We recommend that any steps to improve public education on these topics emphasize one category at a time. Online privacy is a broad topic, and knowledge is not transitive between these categories. Some people may be more knowledgeable about the technical aspects of services but not the financial or legal aspects. Someone who is familiar with privacy law might not understand the technical aspects of online services. Therefore, efforts to increase knowledge in the general population thus should be adaptable to the consumer in order to obtain the most benefit.

C. Personalization of Privacy Plans Based on Consumer Knowledge

As our results show, there are many situations where knowledge and privacy behaviors interact to varying degrees. As part of the transition to a comprehensive privacy regime that includes PIRAs, current approaches to privacy must become more adaptable to the consumer's current level of knowledge. It is with this in mind that we propose a new privacy agent which we have termed Knowledge-based Individualized Privacy Plans (KIPPs).

The excessive length, complexity, and technical nature of many privacy notices make the notices appear almost impenetrable to the average person. By addressing certain crucial gaps in preexisting knowledge, KIPPs could enhance the role of privacy notices in informing consumer decision making. Furthermore, KIPPs could include elements to inform users about the benefits of a service in addition to its risks, thereby providing an environment where decisions are made meaningfully with a more complete understanding of relevant risks and benefits.

The widespread diversity of people who use information technology across the United States and elsewhere makes a "one-size-fits-all" approach impractical. Personalizing the presentation and structure of information for diverse individuals has been successful in other domains. For example, Individualized Education Programs (IEPs) are widely used in schools across the United States to help students with disabilities improve their educational outcomes. In developing IEPs, teachers team with parents and other individuals to identify the unique needs of a student through an evaluation process. Once a student's needs are identified, IEP team members construct an individualized plan for the student and subsequently reevaluate the plan based on the student's progress.[186]

The overall goal of KIPPs would be quite similar. Individuals' privacy knowledge would be assessed, and a personalized privacy plan could be generated based on the results. The evaluation process could either be subjective (i.e., asking the user to rate his or her privacy knowledge) or objective (i.e., having the user answer a few knowledge-based questions), depending on a user's preferences and the practicality of each approach in a given context. Regardless of form, however, the assessment would need to be brief and carefully constructed in order to be useful on a large scale. Effective questions could assess knowledge regarding the data trade for personal

186. *See generally* U.S. Dep't of Educ., *Topic: Individualized Education Program (IEP)*, ED.GOV, http://idea.ed.gov/explore/view/p/%2Croot%2Cdynamic%2CTopicalBrief%2C10%2C [http://perma.cc/VT3M-RQ8Q].

information online, the alternatives to using a service, the capabilities of similar services, and the privacy protections available to consumers.

To increase the usefulness of KIPPs, this approach should be combined with an opt-in model of information sharing. Businesses often have unfavorable views of opt-in models because consumers often share less when the default is to not share information. However, with the use of KIPPs, the benefits and risks of a service can be presented in a balanced manner, and consumers will have the opportunity to better evaluate whether they believe the benefits outweigh the less optimal privacy terms for the service. As our analysis shows, consumers who are more informed about the benefits of cloud computing services may be willing to use more services and share more information with these services. We anticipate that this will continue to be true under an opt-in model where consumers are informed about the benefits alongside the potential risks. KIPPs also provide an incentive to keep service providers accountable for their privacy practices in order to successfully persuade users that the benefits from the service outweigh any privacy compromises.

Once the evaluation process is completed, the results would be used to form the basis of a KIPP. The degree of specificity versus generality for KIPPs could vary for users depending on the context. KIPPs could be as simple as highlighting the importance of privacy notices for users who are unaware of the techniques commonly used by websites to collect, process, and distribute personal information. They could also be more complex and aid users to more efficiently skim privacy notices for key terms and provisions or help users compare the privacy practices of different companies. These details and other implementation issues would likely depend on the goals of individual users and the entities that provide KIPPs. At this time, we are not proposing specific content that would allow KIPPs to function in an optimal manner, but these overarching themes and values provide guidelines for future research.

In order to be effective in practice, KIPPs would need to address the incentives of consumers, businesses, and relevant third parties. Thus, as a starting place, one might ask, "Why would companies or third parties want to provide users with KIPPs?" Our response depends on whether KIPPs are offered on a company-by-company basis or on a broader scale, encompassing the privacy notices of multiple companies within a given industry. If the former approach is used, companies would likely provide KIPPs to help users interpret their own privacy notices. Companies might be interested in using KIPPs in order to improve customer relations, increase transparency, and compete with market alternatives. If KIPPs are implemented using a broader approach, they would likely need to be provided by a third party, such as a privacy advocacy group, industry-wide committee, or our proposed PIRAs. The third-party organization could set forth general guidelines for structuring privacy notices to accommodate consumers with varying degrees of preexisting knowledge. If the guidelines proved effective and gained popularity, individual companies would have an incentive to adhere to the guidelines in order to attract consumers.

Further research should be conducted to determine the optimal implementation for KIPPs. We view KIPPs as being a promising complement that could increase consumer engagement and involvement with data privacy issues while PIRAs become more established.

CONCLUSION

Online privacy is only going to grow in importance as consumers become more aware of the problems surrounding information insecurity. It seems that a new security breach is announced every week. Consumers are starting to pay more attention to how much of their information is out there and who is holding on to it. Unfortunately, they remain helpless to make any meaningful changes in how their data is handled.

We conducted a survey to explore consumer knowledge and opinions about how digital data is collected and used by a variety of actors. We then compared the results of the separate knowledge and opinion surveys to identify areas where higher or lower knowledge was associated with different privacy opinions or behaviors. Our survey results highlight the current state of affairs when it comes to online data. Consumers often want more options than the market gives them. Some of our results suggest that part of the reason that there is a disconnect between what consumers want and what they do when it comes to privacy is because they do not have any meaningful alternatives. The private sector currently has very little incentive to provide these alternatives because consumers have been responding with complacency. Thus far, the response from policymakers has largely been to require companies to give more notices to their customers, thereby giving individuals who are likely already overwhelmed with service information even more things to read and understand. Another explanation for the seeming discrepancy lies in the individual's level of knowledge of various topics. For example, a person may value privacy very highly, but if they also know a lot about cloud computing services and the benefits of using them, they might conclude that the observable benefits of a particular service outweigh its potential deleterious effect on privacy.

In this work, we urge further development of statutory privacy regimes and also recommend creating a way for consumers to become more involved in decisions concerning their personal information. We propose centralizing consumer profiles into Profile Information Reporting Agencies (PIRAs) and giving consumers the ability to not only learn who has access to their information, but also to control what information is in their consumer profile. PIRAs are a largely market-based mechanism for addressing this need, though we also urge legal reforms to provide adequate oversight for the programs. We also suggest that private companies allow users to adopt privacy agents that are adaptable based on each user's level of knowledge. Our proposal for Knowledge-based Individualized Privacy Plans (KIPPs) could provide complementary support while the cogs are put in place for a total overhaul of data privacy law that includes oversight for PIRAs. Educational programs to increase knowledge of cloud computing, online security, and privacy law would support both KIPPs and the establishment of PIRAs.

The growth of the online marketplace will slow prematurely if businesses do not address the desires of consumers. Right now, consumer complacency is leading to favorable outcomes for businesses, but this is not sustainable. Privacy activists and policymakers recognize that many of these practices are unfair to consumers. These practices must eventually change dramatically, or they will undermine the very economy that allowed e-commerce to thrive.