Is Privacy Policy Language Irrelevant to Consumers?

Lior Jacob Strahilevitz and Matthew B. Kugler

ABSTRACT

This article reports the results of two experiments in which large, census-weighted samples of Americans read short excerpts from Facebook's, Yahoo's, and Google's privacy policies, which are at issue in high-stakes privacy class-action lawsuits. Subjects were randomly assigned to read language from either vague policies, some of which had been adjudicated insufficient to notify consumers about the companies' practices, or explicit policies. Though many experimental subjects read these privacy policy excerpts closely, subjects who saw the explicit policies did not differ from those who saw vague policies in their assessment of whether their assent to the policies would permit the corporate practices at issue. Subjects generally stated that agreement to either vague or explicit language authorized companies to collect or use their personal information, even though consumers regarded these corporate practices as intrusive. These experiments show that courts and laypeople can understand the same privacy policy language quite differently.

1. INTRODUCTION

Privacy class actions have become a major financial liability for technology companies. Both Yahoo and Google have been sued over their practices of scanning the contents of users' e-mails to serve them with personalized advertisements, with plaintiffs alleging that their actions violated the Wiretap Act (Order Granting in Part and Denying in Part Defendant's

LIOR JACOB STRAHILEVITZ is the Sidley Austin Professor of Law at the University of Chicago. MATTHEW B. KUGLER is an Assistant Professor at Northwestern University's Pritzker School of Law. The authors owe thanks to Omri Ben-Shahar, Yun-chien Chang, David Hoffman, William Hubbard, Randy Picker, Heather Whitney, workshop participants at the University of Chicago Law School, and participants in the Coase-Sandor Institute for Law and Economics conference Contracting over Privacy for helpful comments. Adam Woffinden and Taylor Coles provided terrific research assistance. The Russell J. Parsons, Bernard Sang, and Carl S. Lloyd Faculty Research Funds and the Coase-Sandor Institute for Law and Economics provided generous research support.

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

Motion to Dismiss, *In re Google Inc. Gmail Litig.,* No. 13-MD-02430-LHK, 2013 WL 5423918 [N.D. Cal. September 26, 2013]; *In re Yahoo Mail Litig.,* 7 F. Supp.3d 1016 [N.D. Cal. 2014]; 18 U.S.C. 2511). In each case the potential liability would have been staggering. Plaintiffs would have been entitled to a minimum of $100 per day of the violation, easily leading to total damages in the trillions for a company with as many users as Google (18 U.S.C. 2520[2][B]).[1] Facebook and Shutterfly have been sued for similar privacy violations under the Illinois Biometric Information Privacy Act (740 Ill. Comp. Stat. 14) for their use of biometrics to identify people in uploaded photos. Here, too, liability could be enormous.[2]

In the Google and Yahoo cases, the defenses have turned on the content of the privacy policies active during the relevant period. Yahoo's policy was deemed sufficiently explicit about the e-mail monitoring that a judge ruled that its users had consented to the monitoring, which defeated the wiretap claim. That same judge held that Google's policy was not sufficiently clear, however, so its users had not consented. Though the plaintiffs' motion for class certification there was ultimately defeated, which limited Google's exposure, both the original case and related litigation are ongoing (Stempel 2014; *Corley v. Google Inc.,* No. 5-16-cv-00473, Complaint [N.D. Cal. January 27, 2016]).

The lawsuit against Facebook is still in the early stages, but it appears that it too could turn on whether its users consented to the alleged conduct by agreeing to its privacy policies and whether they were sufficiently informed about how the data would be used (*Pezen v. Facebook,* No. 1:15-cv-03484, Class Action Complaint [N.D. Ill. April 21, 2015];

[1] Most Gmail users send or receive some e-mail every day, and Gmail has approximately 500 million users worldwide. If Gmail had an average of 50 million American users during the 5-year period of alleged violations, then its liability under the lawsuit could be $9 trillion (50 million users × 365 days per year × 5 years × $100 per user per day).

[2] The stakes in the Facebook suit are again high because of a $5,000 minimum statutory damages provision in the Illinois law (740 Ill. Comp. Stat. 14/20[2]; Welinder 2012). A back-of-the-envelope calculation reveals that even if no Illinois Facebook user could sue for multiple violations of the law, Facebook's potential exposure is still approximately $37.5 billion. Approximately 58 percent of Americans had Facebook accounts as of 2015, and Illinois had about 12.9 million residents at that time. Assuming that Illinois residents use Facebook at national average rates, that means there were about 7.5 million Facebook users in the state. Multiplying that figure by $5,000 yields $37.5 billion. But if each instance of unauthorized tagging is a separate violation, then Facebook's potential liability could quickly escalate from there. The statutory text seems ambiguous on the question (740 Ill. Comp. Stat. 14/1–99).

Licata v. Facebook, Inc., No. 3:15-cv-03747-JD, Consolidated Class Action Complaint [N.D. Cal. August 28, 2015]). A copycat lawsuit against Shutterfly for its violations of the same Illinois statute, based on similar underlying conduct, has already withstood the defendant's motion to dismiss (*Norberg v. Shutterfly,* Class Action Complaint [N.D. Ill. June 17, 2015]; Order Denying Defendant's Motion to Dismiss, *Norberg v. Shutterfly* [N.D. Ill. December 29, 2015]).

In each of these cases, courts have been tasked with interpreting consumer privacy policies. In ruling on Google's motion to dismiss the initial Wiretap Act suit, the district court assumed that Gmail users read the privacy policies in question and then found that agreeing explicitly to the terms of those policies would not have amounted to consent to the automated e-mail content analysis as a matter of law. As the district court knew, and as scholars have long argued, consumers do not typically read privacy policies and other online disclosures, even for products like Gmail that they use every day (McDonald and Cranor 2008; Marotta-Wurgler 2011; Ben-Shahar and Schneider 2014; Ayres and Schwartz 2014; Porat and Strahilevitz 2014). But the duty to read is nevertheless very well established in contract doctrine (Knapp 2015). Courts know that most consumers do not read privacy policies but pretend otherwise for the purposes of contract law and then ask how a reasonable consumer would have interpreted the contract.

Suppose that consumers actually read consumer contracts and privacy policies. What would they understand from them? Would consumers draw the same distinctions between, say, the Yahoo and Gmail privacy policies that the district court did? This article addresses that question through an experimental approach, and the results are surprising. After reading policy language from Gmail, Yahoo, and Facebook, American users of e-mail and social networking websites largely believe that by using those products they have consented to automated content analysis of their e-mails and the use of facial recognition biometrics to suggest photograph tags. That is true regardless of whether consumers read versions of those privacy policies (like Yahoo's) that are extremely explicit or whether consumers read companies' older privacy policies, which (at least in the Gmail litigation) a court deemed inadequate to obtain users' consent. In short, even when consumers do read privacy policies, their beliefs about the nature of their bargains with technology companies seem to depend more on their preexisting expectations than on the terms of the policies (Wilkinson-Ryan 2014).

Interestingly, it does not appear that Americans' views that they have consented to such privacy intrusions stem from normative approval of Google's and Facebook's practices. When asked about the intrusiveness of Google's and Facebook's practices, respondents rate these practices as highly intrusive. In light of these reactions, the most plausible interpretation of the data presented here is that e-mail and social networking users believe these practices are part of the bundle associated with Gmail and Facebook and believe themselves to have accepted that bundle, all the while preferring that the bundle included greater privacy protections.

2. PRIOR LITERATURE

There is a slowly growing experimental literature on consumer contracts. Some of it, like the present study, employs random-assignment techniques to determine what effects changes in contract language or structure have on consumers' behavior. For example, Eigen (2012) randomly assigned online survey participants to conditions that mimicked standard contract boilerplate, a compelled choice between two terms, and notice plus choice. He found that respondents assigned to the boilerplate condition were less likely to read contractual terms and also devoted less energy to performing the task the experiment asked them to do. Mitts (2014) randomly assigned a mix of real and fictitious contract terms to respondents and identified surprising and unexpected terms. Such terms were then highlighted with warnings for consumers. He found that the more times warnings about unexpected terms were given to consumers, the less effective each warning was in helping consumers understand the terms of the agreement. And Hoffman (forthcoming) finds that consumers, particularly younger ones, generally regard written contracts to be more binding than oral contracts.

Other experimental research identifies the role that consumer contract language can have in shaping consumers' expectations about the nature of the bargain. Mamonov and Benbunan-Fich (2015) find that consumers regard privacy breaches to be more disturbing when they are told that the party storing the data has rights to use it than when told that the party storing the data lacks such use rights. This research suggests that respondents do care about what is in privacy policies and that such content can shape their understanding of a counterparty's obligations. Other experimental research suggests that attributes like contract length affect consumers' likelihood of accepting or rejecting a written contract (Plaut and Bartlett 2012). Similarly, the existence of liquidated-damages provisions in mortgage contracts affects the extent to which experimental subjects regard contractual breaches to be immoral (Seiler, forthcoming).

A separate literature examines the psychology of consumer contracts. This literature indicates that only a minuscule percentage of consumers read boilerplate contractual language (Marotta-Wurgler 2012) and that parties' expectations about the contents of a contract are driven not only by written terms of the deal but also by moral and legal norms (Wilkinson-Ryan 2012). We see that the formalization of a contractual arrangement looms large in the lay understanding of what it means to be bound by promises, and contract formation is less of a binary on-off switch than a gradual process in which parties feel increasingly bound as the relationship becomes more formalized over time (Wilkinson-Ryan and Hoffman 2015). Finally, and most relevant for present purposes, consumers who have signed contracts often feel morally bound to those terms, even when they regard the terms to be substantively unfair and when their agreement to those terms causes them to suffer harm (Wilkinson-Ryan 2014).

Another relevant experimental literature explores the existence of a privacy paradox. Privacy paradoxes arise because Americans often say they care a great deal about privacy and yet are willing to permit third parties to obtain sensitive information about them in exchange for relatively inexpensive goods and services or in exchange for longshot odds to win a prize in a random drawing (Acquisti, Brandimarte, and Loewenstein 2015; Holland 2010; Swire 1999). The diminished value placed on privacy may stem in part from framing effects (Acquisti, John, and Loewenstein 2013).

3. DATA AND EMPIRICAL APPROACH

3.1. The Sample

Toluna, a professional survey research firm with an established panel, administered a survey to a weighted sample of 1,441 adult US citizens between May 26, 2015, and June 2, 2015. Data from some of these respondents were discarded because of abnormally fast survey completion times and failed attention checks, which left a final sample of 1,382. The median age of respondents was 47 (range = 18–89, mean = 46.62, SD = 16.37). Females composed 49.8 percent of the sample. Compared with the population in the US census, a higher percentage of the panel had completed high school or at least some college course work, but the educational attainment of the respondents was otherwise similar to that of the adult census population. A total of 79.9 percent of the sample self-identified as white, 13.0 percent as black, and 4.1 percent as South Asian or East Asian. On a separate question, 16.2 percent of the sample reported that they were Latino or Hispanic. Respondents were asked their political orientation on a scale of 1 (very liberal) to 7 (very conservative), with a mean response of 4.16 (SD = 1.78), indicating an ideologically moderate sample. The Gmail and Facebook questions were administered at the end of a 10–15-minute survey that included questions for other studies on topics such as Fourth Amendment privacy expectations and trademark questions designed to assess attributions of product sponsorship.[3]

Participants were screened on the basis of whether they reported having e-mail accounts for the Gmail questions and whether they said they had Facebook accounts for the Facebook questions. That left 1,377 potential respondents to the e-mail questions and 1,052 potential respondents for the Facebook questions.[4] Approximately 76.1 percent of the respondents were therefore Facebook users. This utilization rate is close to the one produced by a Pew Research study conducted a few months earlier, which found that 72 percent of American adults with Internet access use Facebook (Duggan 2015).[5] In each instance, eligible respondents were randomly assigned one of three privacy policies for both the Gmail

[3] These survey results are discussed in Kugler and Strahilevitz (2016b) and Kugler (forthcoming), respectively.

[4] Facebook users were, on average, slightly younger than Facebook nonusers (users mean = 45.06, SD = 16.17; nonusers mean = 51.58, SD = 16.09). The Facebook user population was also more female (52.4 percent) than the general sample. The racial breakdown was roughly equivalent, however (79.0 percent white, 13.6 percent black, 4.0 percent South Asian or East Asian). Note that 28 respondents indicated that they had Facebook accounts but did not answer any of the other Facebook-related questions, so they were dropped from this experiment.

[5] The Pew Research study (Duggan 2015) reports that 62 percent of all US adults are Facebook users. Although our Toluna sample is census weighted, Americans without Internet access were necessarily excluded from the online survey. This exclusion does not seem problematic given our interest in learning how consumers of privacy policies and online apps understand those policies. The exclusion of those without Internet access (13 percent of the adult population) largely explains the disparity in education levels between our sample and the adult population (Perrin and Duggan 2015). The (declining) American population of Internet nonusers is older, lower income, less educated, and more rural than the population of Internet users (Anderson and Perrin 2016).

and Facebook questions. In each instance the privacy policy language subjects read was taken from actual language that Google or Facebook employed at some point in time.[6] The policy language varied in terms of how explicit it was about Google's and Facebook's data practices. Not surprisingly, the current policy language (posted after the main lawsuits at issue here were filed) is more explicit about company practices than the prelawsuit language.

3.2. The Survey Instrument

The randomization strategy in the experiment allows for a clean test of what effect differing policy language has on consumers' views of what they have agreed to. The difference in the new language and old language was (to these lawyers' eyes, at least) dramatic enough to warrant the following pre-experiment hypothesis: lay understandings of privacy policies will depend significantly on the policy language chosen. Given the prominent display of just the relevant language to respondents, enough consumers would read the privacy policy excerpts closely to render the substantial differences between the old and new privacy policies significant.

All respondents were asked to assume that when they signed up for e-mail they agreed to permit advertisements to be shown next to their inboxes in exchange for a free account, and they were also asked to assume that they had read the terms and conditions when signing up for the account. They were then shown randomly assigned privacy policy language that concerned whether these advertisements could be personalized. For example, some saw Gmail's current language, which is quite explicit: "[E-mail provider's] automated systems analyze your content (including e-mails) to provide you personally relevant product features, such as customized search results, tailored advertising, and spam and malware detection." Others saw much vaguer language that Gmail used to post: "[E-mail provider] reserves the right to pre-screen, review, flag, filter, modify, refuse, or remove any or all content from any service. For some services, [e-mail provider] may provide tools to filter out explicit sexual content." Gmail argued unsuccessfully in court that its users' agreement to even that very vague language granted Google's consent to show personalized advertisements to Gmail users.

[6] The Gmail questions used both Google's current language from 2015 and the circa 2011 language quoted in the Gmail litigation. The Facebook questions used Facebook's current language and earlier versions of related privacy policies obtained via the Internet Archive Wayback Machine.

Respondents were also asked other questions targeting issues beyond whether they had consented to the legally relevant conduct by Google and Facebook. Respondents to the Gmail survey were asked, "On a scale of 1 to 10, how intrusive is the e-mail provider's automated e-mail scanning and ad personalization practice?" After answering this question, they were asked, "If there were an option to keep the same e-mail account but pay some amount of money to avoid having the automated systems analyze e-mail content for the purposes of showing you personalized advertisements, how would you respond? (1) I would keep the free e-mail account with the automated e-mail analysis and personalized advertisements. (2) I would be willing to pay some amount of money to avoid the automated analysis." Respondents who selected option 2 were asked how much money they would be willing to pay per year for a more privacy-protective e-mail product.

Respondents to the Facebook question, all of whom had Facebook accounts, were randomly shown various Facebook privacy policies and then asked four questions designed to elicit responses that would shed light on whether Facebook had complied with its obligations under Illinois law. Again, as can be seen from perusing the policy language in the online appendix, differences in the privacy policies seemed stark at first blush. All Facebook respondents were then asked, "Did Facebook's language (above) inform you that information about your facial features was being collected and stored?" "Did Facebook's language (above) inform you of the reason why information about your facial features was being collected, stored, and used?" "Did Facebook's language (above) inform you of the length of time for which information about your facial features would be stored?" (740 Ill. Comp. Stat. 14–XX). And then finally, they were asked the consent question: "Would your decision not to adjust your Timeline and Tagging settings allow Facebook to collect, store, and use information about your facial features?" Respondents to the Facebook questions were then asked to rate on a scale of 1–10 the intrusiveness of Facebook's use of facial recognition software to suggest tags for people whose faces appear in uploaded photos.

4. RESULTS

Given the substantial differences between the privacy policies that e-mail and social networking site users were shown—all this language is reproduced in the online appendix—we predicted that our respondents who saw the highly explicit disclosures from Google and Facebook would be more likely to say that their decision to leave their privacy preferences unchanged after reading the relevant privacy policies allowed Google and Facebook to engage in the content analysis and facial recognition practices at issue. Surprisingly, the data did not support that hypothesis. Regardless of what language respondents were shown, they had statistically indistinguishable views about what practices their inertia would have authorized.

4.1. Experiment 1: Gmail Results

In the Gmail experiment, random assignment to one of three conditions—Google's very explicit current privacy policy, Google's moderately explicit historic section 17 language, or its least explicit historic section 8 language—had no significant effect on consumers' judgment about what they had authorized Google to do to their e-mails. Nor did the privacy policy language have any significant effect on the perceived intrusiveness of Google's automated content analysis of its customers' e-mails. Differences in language that lawyers and judges would deem critical made no evident difference to a representative sample of adult American e-mail users (compare Reidenberg et al. 2015). The differences in means are even in the wrong direction for the comparison between the most explicit policy and the moderately explicit one.

Moreover, in every condition, most respondents said that if they read the short privacy language at issue and then did not change their privacy settings to prohibit content analysis, Google would be authorized to engage in the automated content analysis. Roughly two-thirds of the sample expressed this view in all three conditions.

One possible interpretation of this result is that e-mail users like receiving personalized advertisements and do not mind the automated content analysis of their e-mail that facilitates this personalization. On this interpretation of the data in Table 1, consumers' normative views would be driving their answers to the question of what Google can do. But this interpretation is not supported by the intrusiveness data, shown in Table 2.

The mean intrusiveness response for Google's conduct is 7.63 on a 10-point scale. Consumers are saying that they regard the automated content analysis to be rather creepy, but nevertheless authorized, even when presented with language that few lawyers would regard as consenting to the practice at issue. Intrusiveness ratings, predictably, were not

Table 1. Responses to Gmail Consent Question by Privacy Policy

	Most Explicit	Moderately Explicit	Least Explicit	Overall
Definitely allowed	28.1	28.1	23.5	26.6
Probably allowed	35.7	40.2	39.6	38.5
Probably not allowed	13.9	10.6	15.6	13.4
Definitely not allowed	22.2	21.2	21.3	21.3
Mean	2.30	2.25	2.35	2.30
	(1.10)	(1.08)	(1.06)	(1.08)

Note. Subjects were asked, “Would your agreement to this provision allow the e-mail provider to direct its automated systems to scan the contents of the e-mails you send and receive and show you personalized advertisements?” Gmail’s current privacy policy language is the most explicit, the section 17 language is moderately explicit, and the section 8 language is the least explicit. The frequency differences across conditions are not significant; $\chi^2(2, N = 1{,}363) = 8.38$, $p = .21$. Values in parentheses are standard deviations, and means (allowed = 1, not allowed = 4) do not differ across conditions.

significantly affected by whether respondents saw more explicit or less explicit privacy policies. Those who believe that Google is less authorized to scan e-mails view the practice to be slightly more intrusive ($r(1{,}363) = .192$, $p < .001$), but the effect size is very small.

Our results are consistent with the privacy paradox as well. Although the mean respondent rated automated content analysis of e-mails as 7.63 out of 10 on an intrusiveness scale, just 35.4 percent of the respondents expressed a willingness to pay any amount of money to receive a version of their e-mail service that did not use automated e-mail content analysis to deliver personalized ads. Among the roughly one-third of the respondents who were willing to pay some amount of money, the median willingness to pay was $15 per year. Just 3 percent of the sample expressed a willingness to pay more than $120 per year for such an e-mail service.

Perhaps these data indicate that the intrusiveness ratings offered by our respondents are not to be taken seriously. Maybe the 7.63 figure for intrusiveness is just cheap talk. On this reading of the data, automated e-mail content analysis is not a serious concern for most Americans, which explains why they feel that Google is allowed to engage in the practice even without explicit ex ante warnings. Another possibility is that users of the Internet have grown accustomed to free e-mail, news, weather, media content, and so forth, such that putting e-mail behind a paywall prompts significant resistance even when doing so would create a

Table 2. Responses to Gmail Intrusiveness Practices

E-mail Condition	Intrusiveness Mean	SD	*N*
Most explicit	7.60	2.47	445
Moderately explicit	7.62	2.34	463
Least explicit	7.65	2.43	455
Total	7.63	2.41	1,363

Note. Respondents were asked, "On a scale of 1 to 10, how intrusive is the email provider's automated email scanning and ad personalization practice?" For the dependent variable, intrusiveness, $F(2, 1{,}360) = .06$, $p = .95$; $\eta^2 = .000$.

substantially more privacy-protective product (Dou 2004; Acquisti, John, and Loewenstein 2013). Alternatively, perhaps consumers say that they are reluctant to pay any dollar amount for a privacy-protective e-mail account precisely because they know that other e-mail services (for example, Hotmail) exist, and respondents correctly surmise they do not engage in automated content analysis.[7] Finally, it may be that by purchasing a right to be free of automated content analysis, consumers would be acquiring just a tiny privacy enhancement that would make little difference given other invasive practices. Perhaps if consumers could bundle together a lack of content analysis with limits on behavioral marketing, the commercial use of geolocation, facial recognition software, and sharing of personal information across websites, they would be willing to fork over a more meaningful amount of money.

In any event, the shortage of consumers willing to pay meaningful sums for more privacy-protective e-mail services suggests there may be a limited market for premium products that protect users' privacy. Recent estimates indicate that a year's worth of data is worth $50 to $5,000 per consumer to Google and $45 to $190 per consumer to Facebook (Howe 2015). The sorts of fees they would be able to obtain from users for greater privacy protection are relatively small potatoes, though it is conceivable that enhanced data security would prompt a more robust response from consumers. In any event, the shortage of consumers willing to pay meaningful amounts for more privacy-protective e-mail accounts

[7] Microsoft's privacy policy states, "[W]e do not use what you say in email, chat, video calls or voice mail, or your documents, photos or other personal files to target ads to you" (Microsoft 2016).

plausibly explains why Google has not offered privacy-differentiated e-mail products.

Our data provide information that permits some inferences to be drawn about the dynamics at play. It does not appear that differential views about the intrusiveness of automated e-mail content analysis are driving users to one e-mail provider or another. Mean intrusiveness ratings were not significantly different among Gmail, Yahoo, AOL, and Hotmail users ($F(3, 1{,}136) = 1.44$, $p = .23$; $\eta^2 = .004$). Nor do consumers appear to be choosing their e-mail providers on the basis of their privacy preferences and companies' policies more broadly. When we analyzed responses to questions about the intrusiveness of Facebook's facial recognition software (discussed below) on the basis of what e-mail provider respondents used, there were no significant differences.

It is less clear if awareness of different company practices affects respondents' assessments of whether automated content analysis is permitted. Gmail users were significantly more likely than AOL users to believe that e-mail content analysis was permitted, but so were Hotmail users, and the effect sizes were small in any event.[8] (AOL and Hotmail evidently do not perform automated content analysis of their customers' e-mails.) Given that respondents were asked about whether their own e-mail providers were allowed to engage in automated content analysis, it seems that at most a small portion of the population is attentive to the differences between Google's content analysis and Hotmail's lack thereof.

Our study also generated mixed evidence on the question of willingness to pay. On the one hand, respondents willing to pay some amount of money to avoid content analysis rated the intrusiveness of the content analysis at 8.65 (SD = 1.83), whereas those unwilling to pay any amount rated it at 7.06 (SD = 2.50) ($F(1, 1{,}136) = 149.49$, $p < .001$; $\eta^2 = .10$). On the other hand, the amount people were willing to pay (above zero) bore no relationship to either the perceived intrusiveness or the authorization of automated content analysis.

4.2. Experiment 2: Facebook Results

Under the Illinois Biometric Information Privacy Act, the pertinent legal questions are whether Facebook informed its users about the fact that in-

[8] The results are Gmail 2.23 (SD = 1.07), Yahoo 2.39 (SD = 1.09), AOL 2.45 (SD = 1.06), Hotmail 2.19 (SD = 1.03), total 2.31 (SD = 1.07); $F(3, 1{,}136) = 2.93$, $p = .033$; $\eta^2 = .008$. Results for Gmail and Hotmail were significantly lower than those for AOL and Yahoo ($p < .05$) and did not differ significantly from each other.

formation about their facial features was being collected and stored, the reason why information about their facial features was being collected, and the length of time for which information about their facial features would be stored. In addition, the law renders germane the question of whether Facebook had its users' permission to collect and store information about their facial features. Each of these four questions depends on Facebook users' understanding of Facebook's terms of service. This experiment was designed to test whether, if Facebook users had read the relevant information, they would feel that Facebook had adequately informed them of its practices and obtained their authorization to engage in them.

There was a clear consensus among respondents on all four questions, and the consensus is particularly interesting on the third of the four questions. Despite a sample size of 1,052 respondents, on none of the questions presented in Table 3 does the language from Facebook's privacy policies have any significant effect. More than two-thirds of the sample regarded themselves as having been informed of Facebook's collection and storage of their biometric information after having read any of Facebook's current or historic policy language. And an only slightly lower percentage of Facebook users viewed Facebook's language as informing them of the purpose of Facebook's use and collection. Again, the wording of the policy language at issue made no significant difference, even though in one condition the language was very explicit about the purposes of Facebook's collection of information and in the other it was not. Similarly high percentages of respondents said that users' inaction with respect to privacy settings authorizes Facebook's facial recognition practices.

Viewed in context, the most striking responses in Table 3 are those to the third question, which asks about the length of time for which Facebook retains its information. In none of these conditions did the privacy policy language provided to respondents address the duration of storage explicitly. Up to 67 percent of the respondents seem to have noticed this. This reversal of the usual ratios across all three conditions suggests several possible implications. First, it seems that at the very least a third of the sample is reading the lengthy privacy policy language in the prompt carefully. These are the respondents who flip from a pro-Facebook stance on the other questions to an anti-Facebook stance on the third question. Second, it is possible (though unlikely) that whereas Facebook users have intuitions about the fact that a facial recognition algorithm is being used

Table 3. Positive Responses to Facebook Questions

	Policy			
	We Collect	We Use	When Someone Uploads	Total
Did Facebook inform you about collection and storage?	67.8	65.9	70.7	68.2
Did Facebook inform you of the reason for collection, use, and storage?	59.0	61.8	67.1	62.5
Did Facebook inform you about the length of time information would be stored?	35.1	34.0	30.7	33.3
Does leaving settings unchanged allow Facebook to collect, use, and store information?	63.5	62.3	61.0	62.3

Note. Values are the percentages of respondents who answered yes to the questions. The "we collect" policy is the least explicit about Facebook's actions and purposes. The "we use" policy is the least explicit about Facebook's actions but has more disclosures about its purposes. The "when someone uploads" policy is the most explicit about Facebook's actions but has fewer disclosures about its purposes.

and the reason why it is being used (perhaps based on their use of the feature on Facebook), they lack a strong prior about the length of time for which facial recognition information should be retained, so the privacy policy language may play a larger role than context in shaping their understanding.[9] Third, unless there is other privacy policy language that Facebook can cite,[10] it appears plausible that, according to consumers, Facebook's facial recognition feature has been violating one provision—though only one provision—of the Illinois law. To confirm this hypothesis we would need to test the effects of Facebook's data-retention-duration language on consumers.

[9] Data from Pew Research suggests that Americans generally do have articulated prior beliefs about the length of time for which their personal data should be retained. Just 4 percent of respondents said that social media or online video sites should be able to retain their data for "as long as they need it" (Rainie 2016).

[10] As of January 14, 2016, Facebook's data policy provided in pertinent part, "We store data for as long as it is necessary to provide products and services to you and others" and equivocated on how much data would be eliminated if the account was deleted (Facebook 2016). We did not show this clause to our experimental subjects because this language is plausibly too vague and indefinite about the duration of data retention to satisfy the Illinois statute, and the permanent retention of biometric information gleaned from photos uploaded by other users who do not delete their accounts could well violate the Illinois statute. More broadly, the failure to disclose the duration of data retention appears to be quite commonplace in the United States (Marotta-Wurgler 2016).

Respondents were also asked about the intrusiveness of Facebook's practice of using facial recognition software to suggest tags for people whose faces appear in uploaded photos. Mean responses were a little lower than in the Gmail question (mean = 7.29, SD = 2.36) and did not differ by condition ($F(2, 1{,}044) = 1.30$, $p = .27$; $\eta^2 = .002$). Thus, it does not appear that exposure to different policy language affected consumers' underlying beliefs about how problematic Facebook's practices are. Once again, majorities of consumers appear to regard Facebook's practice as troubling yet authorized. Comparing across conditions between authorization and perceived-intrusiveness responses did not yield significant results ($p = .396$).

5. REPLICATION

Whenever a null result is observed in this type of vignette experiment, one possibility is that participants simply did not attend to the materials. As we know from prior research, boilerplate policy language may encourage consumers to tune out fine details (Eigen 2012). We therefore conducted a replication study that aimed to assess how carefully participants read the provided materials and whether more attentive participants differed from less attentive ones. The new features in this replication were a measure of how long participants spent on the main Facebook and e-mail scenario pages, a manipulation for half of the participants in the e-mail portion of the study that asked them to explain why they thought monitoring was or was not allowed (to encourage deeper thought), a manipulation in the Facebook portion of the study that either provided or omitted information about how long the information would be retained, self-report questions in both the Facebook and e-mail scenarios asking participants how well they felt they understood the materials (10-point scale), and the imposition of a more cognitively demanding attention check that permits us to test our first study's results on a subsample of the most attentive respondents.

The replication experiment also introduced a new e-mail condition that included privacy policy language from Yahoo. The judge who held the less explicit Gmail language to be inadequate was satisfied by this alternative. Adding a condition with this language therefore addresses concerns about our determination that Gmail's current privacy policy—unlike the earlier policy language that is at issue in the Gmail litigation—

would be deemed sufficiently clear and precise to secure consent from consumers who read it.

The procedure was otherwise as before, though the Facebook and e-mail questions came much earlier in the survey, immediately after the background questions, and an effort was made to recruit more participants who had not completed high school or attempted college course work so as to have a more representative mix of education levels. A total of 1,300 participants were recruited, 1,283 of whom completed the e-mail questions and 1,045 of whom completed the Facebook questions.[11]

In general, the main results of the first study were replicated, participants appeared to be paying attention, and more attentive participants did not draw greater distinctions between scenarios than less attentive ones (see Table 4). Three different e-mail scenarios were used: the least explicit from study 1, the most explicit, and a version from Yahoo's terms of service that was even more explicit than any employed by Gmail: "automated systems scan and analyze all incoming and outgoing communications . . . to match and serve targeted advertising." Results again showed no significant differences in whether the policies allowed the described monitoring or in perceived intrusiveness. There was a slight difference across conditions in whether the participants felt they understood the policy: participants were slightly less confident that they understood the least explicit Gmail policy than either of the other two ($p < .06$).

Further, asking participants to explain why they thought the policies they were given did or did not allow monitoring had no effect on whether they thought the policies permit such monitoring ($F = 1.37$) and did not interact with condition to predict whether permission was imputed ($F = .47$). In fact, the only effect of requiring explanations was to make participants take longer on the page ($F = 123.09$, $p < .001$).[12]

A series of linear regressions was conducted in an attempt to predict

[11] The median age of respondents was 44 (range = 18–90; mean = 45.72; SD = 16.08). Females comprised 51.0 percent of the sample; 81.5 percent of the sample self-identified as white, 10.5 percent indicated they were black, and 3.5 percent self-identified as South Asian or East Asian. On a separate question, 15.5 percent of the sample reported that they were Latino or Hispanic. Respondents were asked their political orientation on a scale of 1 (very liberal) to 7 (very conservative), with a mean response of 4.11 (SD = 1.75); 11.38 percent of the sample had not finished high school, 30.38 percent had high school diplomas, 29.08 percent had some college experience, 19.00 percent had college degrees, and 10.15 percent had some kind of graduate degree.

[12] Since time spent on the page was not normally distributed (some participants were on the page for a long time), the variable was capped at 250 seconds for this and all subsequent analyses.

Table 4. E-mail Questions from the Replication Study

	Gmail Policy					
	Least Explicit	Most Explicit	Yahoo Policy	Total	*F*-Statistic	*p*-Value
Allow monitoring	2.39 (1.10)	2.34 (1.12)	2.26 (1.12)	2.33 (1.11)	1.68	.19
Intrusive	7.37 (2.39)	7.57 (2.31)	7.70 (2.22)	7.55 (2.31)	2.33	.10
Understand policy	7.74 (2.07)	8.01 (2.01)	8.08 (1.99)	7.95 (2.03)	3.34	.04

Note. Values are mean responses with significance tests. Standard deviations are in parentheses.

the score on the allow-monitoring dependent measure from condition and its interactions with either self-reported understanding or time spent on the web page. Neither variable interacted with condition, which indicates that people who spent longer with one policy or another, or felt that they better understood one policy or another, did not differ from other participants in whether they thought the policy allowed monitoring.[13]

For the Facebook scenarios, the most ("when someone uploads") and least ("we collect") explicit policies reprised their roles from study 1, but a new version of the most explicit policy was created that included the line "We automatically delete all facial recognition information once it has been stored in our system for three years." This changes the correct response for the length of time the information is stored. As can be seen in Table 5, participants are sensitive to this change: the majority of the respondents in that condition recognized that they had received this information, significantly more than in the other conditions. That said, about 37 percent of the sample answered this question incorrectly, which indicates that they did not read the policy closely or that their prior beliefs overwhelmed the policy language. The other questions, assessing what the policy means for users, did not produce different answers across conditions, which replicates study 1. Perceived understanding also did not differ across conditions ($F = .67$, overall mean = 7.44, SD = 2.26).

[13] Interestingly, there were two main effects. Those who spent longer on each page (regardless of condition) were less likely to say they believed the policies allowed the monitoring (standardized $\beta = .116$, $p < .001$), and those who felt they better understood the policies were more likely to believe that monitoring was allowed (standardized $\beta = -.149$, $p < .001$).

Table 5. Facebook Question Replication Responses

	When Someone Uploads	When Someone Uploads: Limited	We Collect	χ^2	p-Value
Did Facebook inform you about collection and storage?	73.68	77.30	71.74	2.81	.24
Did Facebook inform you of the reason for collection, use, and storage?	62.57	61.10	65.73	1.59	.45
Did Facebook inform you about the length of time information would be stored?	39.59	62.64	34.17	62.45	<.001
Does leaving settings unchanged allow Facebook to collect, use, and store information?	64.12	59.83	65.63	2.62	.27

Note. Values are the percentages of respondents who answered yes to the question. The limited policy includes language indicating that facial recognition information is deleted after 3 years.

Several regressions were conducted to see whether the effects of perceived understanding or length of time on the page affected responses to these questions differently depending on condition. For the understanding question, there were no significant interactions, which means that those who thought they understood the prompt well did not come to different answers depending on which prompt they read.[14] For time spent on the page, the only interactions were on the question about the duration of data retention.[15] Those who spent longer on the page were more likely to indicate the correct answer on that question, which means that they answered yes for the duration-limit condition and no for the other two conditions.

Finally, the introduction of the new 3-year-time-limit condition in the

[14] There were significant main effects of self-reported understanding on the first three questions: odds ratios of .53, .69, and .66, respectively ($p < .001$ for each).

[15] This effect is easier understood in terms of an analysis of variance. There was a significant interaction between Facebook condition and the answer to the duration-limit question on time spent on the page: $F(1, 998) = 23.51$, $p < .001$; $\eta^2 = .05$; "when someone uploads," yes = 46.99 (49.89), no = 100.75 (66.09); with 3-year limit, yes = 95.11 (69.09), no = 80.72 (60.26); and "we collect," yes = 70.94 (76.72), no = 112.67 (74.30). In the duration-limited condition, those saying yes took significantly longer. In the others, where this was the wrong answer, those saying yes took significantly less time ($p < .05$ for each). The binary logistic regression version of this analysis is available from the authors.

replication experiment permitted us to apply a relatively demanding new attention check to our sample. After our first experiment, we hypothesized that respondents' assessment of whether e-mail content analysis was permitted would not differ across conditions even among our most attentive readers. In the replication study we tested this hypothesis by examining whether our headline results would change if we omitted the Yahoo and Gmail responses of subjects who answered the Facebook data-retention questions incorrectly. Even respondents who read the Facebook questions closely enough to notice the presence or absence of a single sentence buried in a paragraph from a privacy policy did not differentiate between Yahoo's legally adequate and Gmail's legally inadequate privacy policies in terms of whether content analysis was authorized ($F < 1$).

6. DISCUSSION

The key lesson from both the Facebook and e-mail data is that users of e-mail and social networking sites appear to regard even highly ambiguous privacy policy language as authorizing controversial company practices that implicate their personal privacy. Wilkinson-Ryan (2014) finds a similar result in the context of other boilerplate consumer contracts. Though federal courts have determined that Yahoo's privacy policy informed e-mail users of the company's automated content analysis and that Gmail's privacy policy did not, American e-mail users did not differentiate between the purportedly adequate and inadequate policies. To the contrary, they thought that agreeing to either policy would establish their consent to automated content analysis.

What explains the divergence between lawyerly judgments and lay consumers' judgments about what constitutes consent? One possible explanation is that consumers had formed strong prior beliefs about the sort of privacy-related conduct that companies are permitted to engage in, and these prior beliefs inform their understanding about what they agree to when they use Gmail or Facebook without changing their privacy settings (Martin 2015). Even when consumers are familiar with the formal law and written policy language, expectations are also driven by social norms. When consumers interpret contracts, they bring in these prior beliefs and integrate them with the policy language to produce an understanding of the bargain to which they are agreeing (Wilkinson-Ryan 2012). Consumers may not like the bargain in all material respects—and their intrusiveness scores suggest discomfort with automated e-mail content analysis and the automated use of facial recognition software—but they seem to believe that the privacy sacrifices inherent in their use of e-mail and social networking sites outweigh those costs.

When faced with data like these and a consent defense by a defendant who invokes this sort of empirical evidence, what should a court do? In our view, data such as these, collected using rigorous survey techniques and analyzed by academics with no stake in the outcome, ought to play a large role in litigation over privacy policies in particular and consumer contracts in general.[16] Under such an approach, interpretation of consumer contracts would become a question of fact rather than a question of law. Where a consensus emerges among consumers as to the contours of a deal, this consensus understanding would become the contract's meaning, even among those consumers who had subjective views of the contract that placed them in the minority. This survey-driven approach would represent a break with American law's dominant paradigm for contract interpretation (see, for example, *Antilles Steamship Co. Ltd. v. Members of Am. Hull Ins. Syndicate,* 733 F2d 195, 204 n2 [2d Cir. 1984], J. Newman concurring).[17] Under a survey-driven approach, the interpretation of consumer contracts would more closely resemble what courts do in trademark litigation, where consumer surveys are dispositive (Diamond and Franklyn 2014).

Though at first blush this change in the law would make the law more hostile to business interests (by making it harder to win a motion to dismiss in a contract suit) and friendlier to plaintiffs' interests, this result is hardly inevitable. Battling over the legal meaning of contractual terms is not cheap. A consumer-survey-driven approach to contract interpretation would resolve cases at a later stage, but that does not mean that more money would be spent before resolution. Legal research that takes place early in litigation now could be replaced with survey research. To the extent that dominant survey methodologies emerged quickly, then the parties could promptly settle in the shadow of their experts' survey results. Indeed, a lot of current contract claims might never be brought in the first instance because plaintiffs' attorneys would have a relatively inexpensive

[16] Ben-Shahar and Strahilevitz (2016) develop this argument in much more depth.

[17] One paper comes close to advocating such an approach but uses it to resolve a hypothetical question about the likelihood of scarce goods being available in the future, as opposed to a question of what the contract language means (Olazábal, Marmorstein, and Sarel 2014).

way to test whether a breach-of-consumer-contracts claim would be viable. It is plausible, though by no means certain, that prompting the law to focus on ordinary consumers' understandings of contractual provisions would be more efficient than the current approach.

The goal of companies designing privacy policies and consumer contract language should be to inform consumers about what the companies are doing and why they are doing it. Companies already field-test their products extensively. For similar reasons, they should field-test their policy language on consumers and avoid presuming that the only information consumers have is what is disclosed in the policy language. It is precisely because lawyers are trying to cram so much information into policies that policies become unduly lengthy, and the result is that they go unread entirely by rational consumers (Ben-Shahar and Schneider 2014). The meaning of a consumer contract is a product of consumers' expectations and the contract language, with the former seemingly looming larger than the latter in some contexts. The product is readily measurable, even if teasing out what work the expectations are doing and what work the language is doing is more complex. At least in the instance of Gmail, privacy policy language chosen by Google and the other information that consumers are receiving or intuiting from various sources do adequately inform most consumers about the nature of the bargain.[18]

Several important caveats remain. First, we know both that consumers very rarely read privacy policies and that courts adjudicating class-action cases nearly always impose a duty to read on consumers. There may be sensible reasons for the courts to proceed on that basis, particularly at the motion-to-dismiss stage or the summary judgment stage. But if they do assume that consumers read these contracts, it seems highly problematic to assume an interpretation of those contracts that relatively few lay readers of those contracts would share. The duty to read cannot possibly mean a "duty to hire a lawyer to read in a lawyerly way." Can it?

Second, in assessing the generalizability of these results, it is important to recall that our respondents were asked to read only a short excerpt of a much lengthier privacy policy. Respondents were not charged with scanning a dense policy and finding the relevant provision. Had we asked

[18] If this approach were adopted, it is possible that there would be certain contract provisions that survey respondents would find so surprising and unbelievable that firms could never successfully integrate them into a bargain, no matter how explicit the contractual language employed. We think that this data-driven approach to contract unconscionability might be more appealing than existing approaches. In any event, in this study we have not identified any such terms.

respondents to read a lengthier policy carefully, few would have been incentivized to comply. However, the Gmail and Facebook questions in our first experiment were presented to our respondents toward the end of a 10–15-minute online survey that also asked them a number of questions about Fourth Amendment privacy questions and trademark issues. That was in part the rationale for our replication study, which placed the privacy questions much earlier in the survey. In any event, the results here should be conceived of as relevant to the question, What would happen if consumers actually read the pertinent parts of privacy policies?—an inquiry that, though hypothetical, winds up being outcome determinative in a great many litigated cases.

Third, there is an adaptive-preferences problem built into our survey methodology that could affect the interpretation of the results. The Facebook experiment was limited to respondents from a nationally representative sample who said that they have Facebook accounts. The respondents therefore had already been exposed to Facebook's tagging suggestions, and many may have already realized that Facebook employs facial recognition software to suggest tags. This previous exposure had benefits and drawbacks. One benefit is that many consumers already understood a technological feature that might have been difficult to explain otherwise. (For reasons related to the complexity of the technology, we did not ask Facebook nonusers to answer the questions.) But a drawback is that by the time Facebook was sued and we presented respondents with our survey, Facebook had been employing facial recognition technology for nearly 5 years (Ducklin 2010). Facebook users' initial understanding of Facebook's practices is arguably as relevant as their contemporary understanding of Facebook's practices. The problem is present too in the Gmail survey, because the firm's practice was again longstanding by the time the survey launched. To be sure, the lack of large differences in the responses of Gmail users and demographically similar Hotmail users alleviates some concerns about conditioned responses. Still, as a result of these issues, our study lacks a clear "before" to go with its "after" result. Because it takes time to get a survey developed, approved, funded, and launched, it is unlikely that third-party researchers will ever be able to test consumers' understandings of companies' new practices before those practices have been implemented. But firms themselves might hire reputable academic researchers to obtain data that predate consumers' adaptation to a new feature. That said, in both the Facebook and Gmail litigation, plaintiffs are seeking continuing damages over a period of several years. Even if we cannot identify precise consumer sentiment at the time a controversial practice began, understanding contemporary responses may help place an upward bound on the damages that are appropriate in any given case.

Finally, there is a hard question of what to do with respondents' heterogeneity. When presented with language in the We Use condition that (to these lawyers' eyes anyway) very clearly informs Facebook users of the reasons why Facebook is collecting facial recognition data, 38 percent of our respondents said that Facebook did not inform them of the reasons for the data collection. A similar percentage of respondents (between 30 percent and 40 percent depending on condition) in the initial and replication surveys provided an objectively incorrect answer to the question of whether Facebook had informed its users about the length of time for which it would be retaining biometric information. And when presented with language in the When Someone Uploads condition that (again, in our judgment) unambiguously informs readers what Facebook is doing, 23–29 percent of our respondents said the language failed to do so. With any survey instrument, there are going to be some people who do not read very carefully or answer most questions at random but nevertheless answer standard attention-check questions correctly, and there will be others who have sufficiently strong views about the facts or morality of an issue to not be swayed by any exculpatory contract language. It appears that in our experiment, those groups combined to form somewhere between 25 and 40 percent of the overall sample. In a world where lawyers have determined that contract or policy language should have some efficacy in shaping consumers' expectations, the fact that 30 or 35 percent of a sample articulates the view that particular policy language with which they were presented is inadequate should not sway a court unduly.

As one examines pleadings in cases such as *In re Google Inc. Gmail Litigation*, *In re Yahoo Mail Litigation*, and *In re Facebook Biometric Information Privacy Litigation* (Defendant Facebook, Inc.'s Motion to Dismiss, No. 3:15-cv-03747-JD [N.D. Cal. October 9, 2015]), the absence of empirics about how consumers respond to terms-of-service language is striking. This is information that litigants (or, better yet, social scientists) ought to be producing and that courts ought to be evaluating (Martin 2015). The survey results presented here were neither particularly difficult nor costly to gather. The total costs for our first survey sample were $4,550, but this sample was used to provide the data for this project as well as three other research papers dealing with disparate topics. Compared with a few billable hours of a good lawyer's time, such experimental research is a bargain. And for a corporation that is trying to limit its exposure to class-action suits, making nationally representative consumer survey results legally dispositive could be a blessing. Instead of engaging in guesswork about which boilerplate language courts would regard to be adequate or inadequate for the purposes of securing consumers' consent, corporations could make an ex ante determination that is presumably likely to remain stable down the road (Kugler and Strahilevitz 2016a).

7. CONCLUSION

It is well understood that consumers typically do not read boilerplate privacy policies and that, for the purposes of determining whether consumers have consented to particular companies' privacy practices, courts nevertheless assume that consumers do read those policies. Our experiments suggest that even if a large number of consumers did read controversial privacy policies, their interpretations of those policies and of what conduct they had authorized would differ from conventional legal interpretations of those policies' meaning. More precisely, consumers seem to regard themselves as having authorized several controversial privacy-related practices by Google, Yahoo, and Facebook regardless of whether they were randomly assigned to read vague language that does not seem to explain the corporate practices in any meaningful detail or precise language that describes the corporate practices at issue with admirable clarity and specificity.

These experimental findings suggest that differences in policy language that are quite salient to lawyers are essentially irrelevant to consumers. Context, experience, and norms, rather than privacy policy language, seem to provide a benchmark for consumers' understandings about what conduct they are authorizing, and that is the case even in those instances in which one can be confident that consumers have read the relevant policy language rather carefully. Moreover, the experiments reported herein suggest that normative prior beliefs about what corporate practices are more or less invasive do not significantly affect most consumers' understandings about what companies like Facebook, Yahoo, and Google are authorized to do. Even though consumers think that uses of facial image recognition and automated e-mail content analysis are invasive, they still regard even vague and imprecise policy language as authorizing Facebook, Yahoo, and Google to engage in those practices. Finally, our experiments provide significant reason to doubt that market forces will significantly incentivize firms to offer privacy-protective alternatives to free services that enhance e-mail privacy. Although consumers dislike automated content analysis, their willingness to pay for a version of Gmail that does not perform content analysis is quite limited, and there is no evidence to indicate that concerns about e-mail content analysis are presently driving consumers to choose substitute e-mail services that eschew e-mail content analysis.

REFERENCES

Acquisti, Alessandro, Laura Brandimarte, and George Loewenstein. 2015. Privacy and Human Behavior in the Age of Information. *Science,* January 30, pp. 509–14.

Acquisti, Alessandro, Leslie K. John, and George Loewenstein. 2013. What Is Privacy Worth? *Journal of Legal Studies* 42:249–74.

Anderson, Monica, and Andrew Perrin. 2016. 13% of Americans Don't Use the Internet: Who Are They? *Pew Research Center Fact Tank*. September 7. http://www.pewresearch.org/fact-tank/2016/09/07/some-americans-dont-use-the-internet-who-are-they/.

Ayres, Ian, and Alan Schwartz. 2014. The No-Reading Problem in Consumer Contract Law. *Stanford Law Review* 66:545–609.

Ben-Shahar, Omri, and Carl E. Schneider. 2014. *More Than You Wanted to Know: The Failure of Mandated Disclosure*. Princeton, NJ: Princeton University Press.

Ben-Shahar, Omri, and Lior Jacob Strahilevitz. 2016. Interpreting Contracts via Surveys and Experiments. Unpublished manuscript. University of Chicago Law School, Chicago.

Brandom, Russell. 2016. Someone's Trying to Gut America's Strongest Biometric Privacy Law. *The Verge,* May 27. http://www.theverge.com/2016/5/27/11794512/facial-recognition-law-illinois-facebook-google-snapchat.

Diamond, Shari Seidman, and David J. Franklyn. 2014. Trademark Surveys: An Undulating Path. *Texas Law Review* 92:2029–73.

Dou, Wenyu. 2004. Will Internet Users Pay for Online Content? *Journal of Advertising Research* 44:349–59.

Ducklin, Paul. 2010. Automatic Photo Tagging: Facebook Friendships Get Creepier. *Naked Security,* December 17. https://nakedsecurity.sophos.com/2010/12/17/facebook-friendships-get-creepier/.

Duggan, Maeve. 2015. The Demographics of Social Media Users. *Pew Research Center: Internet, Science, and Tech,* August 19. http://www.pewinternet.org

/2015/08/19/the-demographics-of-social-media-users/.

Eigen, Zev. 2012. Experimental Evidence of the Relationship between Reading the Fine Print and Performance of Form-Contract Terms. *Journal of Institutional and Theoretical Economics* 168:124–41.

Facebook. 2016. Data Policy. January 14. https://www.facebook.com/full_data_use_policy.

Hoffman, David A. Forthcoming. From Promise to Form: How Contracting Online Changes Consumers. *New York University Law Review*.

Holland, H. Brian. 2010. Privacy Paradox 2.0. *Widener Law Journal* 19:893–932.

Howe, Jared. 2015. How Much Is Your Personal Data Worth? *Private WiFi*, June 9. http://blog.privatewifi.com/how-much-is-your-personal-data-worth/.

Knapp, Charles L. 2015. Is There a "Duty to Read?" *Hastings Law Journal* 66:1083–1112.

Kugler, Matthew B. Forthcoming. Measuring Sponsorship Materiality. *U.C. Davis Law Review*.

Kugler, Matthew B., and Lior Jacob Strahilevitz. 2016a. The Myth of Fourth Amendment Circularity. Unpublished manuscript. University of Chicago Law School, Chicago.

———. 2016b. Surveillance Duration Doesn't Affect Privacy Expectations: An Empirical Test of the Mosaic Theory. *Supreme Court Review* 2015:205–63.

Mamonov, Stanislav, and Raquel Benbunan-Fich. 2015. An Empirical Investigation of Privacy Breach Perceptions among Smartphone Application Users. *Computers in Human Behavior* 49:427–36.

Marotta-Wurgler, Florencia. 2011. Some Realities of Online Contracting. *Supreme Court Economic Review* 19:11–23.

———. 2012. Does Contract Disclosure Matter? *Journal of Institutional and Theoretical Economics* 168:94–119.

———. 2016. Self-Regulation and Competition in Privacy Policies. *Journal of Legal Studies* 45:S13–S39.

Martin, Kirsten. 2015. Privacy Notices as Tabula Rasa: An Empirical Investigation into How Complying with a Privacy Notice Is Related to Meeting Privacy Expectations Online. *Journal of Public Policy and Marketing* 34:210–27.

McDonald, Aleecia M., and Lorrie Faith Cranor. 2008. The Cost of Reading Privacy Policies. *I/S: A Journal of Law and Privacy for the Information Society* 4:543–65.

Microsoft. 2016. Microsoft Privacy Statement. September. https://privacy.microsoft.com/en-us/privacystatement.

Mitts, Joshua. 2014. How Much Mandatory Disclosure Is Effective. Unpublished manuscript. Columbia University Law School, New York.

Olazábal, Ann Morales, Howard Marmorstein, and Dan Sarel. 2014. Frequent Flyer Programs: Empirically Assessing Consumers' Reasonable Expectations. *American Business Law Journal* 51:175–250.

Perrin, Andrew, and Maeve Duggan. 2015. Americans' Internet Access: 2000–2015. *Pew Research Center: Internet, Science, and Tech,* June 26. http://www.pewinternet.org/2015/06/26/americans-internet-access-2000-2015/.

Plaut, Victoria C., and Robert P. Bartlett III. 2012. Blind Consent? A Social Psychological Investigation of Non-readership of Click-through Agreements. *Law and Human Behavior* 36:293–311.

Porat, Ariel, and Lior Jacob Strahilevitz. 2014. Personalizing Default Rules and Disclosure with Big Data. *Michigan Law Review* 112:1417–78.

Rainie, Lee. 2016. The State of Privacy in America: What We Learned. *Pew Research Center Fact Tank,* September 21. http://www.pewresearch.org/fact-tank/2016/09/21/the-state-of-privacy-in-america/.

Reidenberg, Joel R., Travis Breaux, Lorrie Faith Cranor, and Brian French. 2015. Disagreeable Privacy Policies: Mismatches between Meaning and Users' Understanding. *Berkeley Technology Law Journal* 30:39–68.

Seiler, Michael J. Forthcoming. Do Liquidated Damages Clauses Affect Strategic Mortgage Default Morality? A Test of the Disjunctive Thesis. *Real Estate Economics*.

Stempel, Jonathan. 2014. Google Won't Face Email Privacy Class Action. *Reuters.com,* March 19. http://www.reuters.com/article/us-google-gmail-lawsuit-idUSBREA2I13G20140319.

Swire, Peter P. 1999. Financial Privacy and the Theory of High-Tech Government Surveillance. *Washington University Law Quarterly* 77:461–512.

Welinder, Yana. 2012. A Face Tells More than a Thousand Posts: Developing Face Recognition Privacy in Social Networks. *Harvard Journal of Law and Technology* 26:165–239.

Wilkinson-Ryan, Tess. 2012. Legal Promise and Psychological Contract. *Wake Forest Law Review* 47:843–73.

———. 2014. A Psychological Account of Consent to Fine Print. *Iowa Law Review* 99:1745–84.

Wilkinson-Ryan, Tess, and David A. Hoffman. 2015. The Common Sense of Contract Formation. *Stanford Law Review* 67:1269–1301.