A Comparative Study of Online Privacy Policies and Formats

Aleecia M. McDonald[1], Robert W. Reeder[2], Patrick Kelley[1], Lorrie Faith Cranor[1]

[1] Carnegie Mellon, Pittsburgh, PA
[2] Microsoft, Redmond, WA

Abstract. Online privacy policies are difficult to understand. Most privacy policies require a college reading level and an ability to decode legalistic, confusing, or jargon-laden phrases. Privacy researchers and industry groups have devised several standardized privacy policy formats to address these issues and help people compare policies. We evaluated three formats in this paper: layered policies, which present a short form with standardized components in addition to a full policy; the Privacy Finder privacy report, which standardizes the text descriptions of privacy practices in a brief bulleted format; and conventional non-standardized human-readable policies. We contrasted six companies' policies, deliberately selected to span the range from unusually readable to challenging. Based on the results of our online study of 749 Internet users, we found participants were not able to reliably understand company's privacy practices with any of the formats. Compared to natural language, participants were faster with standardized formats but at the expense of accuracy for layered policies. Privacy finder formats supported accuracy more than natural language for harder questions. Improved readability scores did not translate to improved performance. All formats and policies were similarly disliked. We discuss our findings as well as public policy implications.

1 Introduction

The United States relies on a self-regulation approach to Internet privacy. There are some Internet privacy laws, for example the Children's Online Privacy Protection Act of 1998 (COPPA), which protects children's privacy[5], and the Gramm-Leach-Bliley Act (GLB), which applies to financial data [9]. But by and large the theory of Internet privacy hinges on two assumptions:

- Consumers will choose companies with acceptable privacy policies.
- Companies will not violate their privacy policies because the Federal Trade Commission (FTC) can bring action for unfair and deceptive practices.

* Thanks to Robert McGuire and Keisha How for programming assistance

In both cases privacy policies play a vital role in Internet privacy. Self-reports show three quarters of Internet users take active measures to protect their privacy, ranging from installing privacy protective technology to providing false information to web sites [1]. Yet only 26% read privacy policies during a recent study and readership outside of laboratory conditions is believed to be far lower [12]. Free market mechanisms based in consumer choice will fail to protect privacy if consumers do not understand the choices available to them.

To study the effectiveness of various approaches to improving the usability of privacy policies, we investigated the performance of three different formats for privacy policies and compared policies from six different companies.

In section two we describe related work and the formats we contrasted. We describe our methods in section three. We present accuracy and time to answer results in section four, and psychological acceptability results in section five. We discuss implications from these results and conclude in section six.

2 Related Work

Several studies frame willingness to read privacy policies as an economic proposition and conclude that asymmetric information is one reason why people find it not worth their time to read privacy policies [25, 1]. Other studies show that privacy policies require a college reading level to understand [10, 21, 8, 2]. A study of ambiguities in privacy policies shows they contain language that downplays privacy issues [17]. The 2006 Kleimann report on GLB financial privacy notices found that subheadings and standard formats dramatically improved readability [19]. In response to these issues, privacy researchers and industry groups devised several standardized formats for privacy policies based on the expectation that standardized formats would improve comprehension. Our study is a comparative analysis to analyze how well standardized policies work in practice.

Prior work investigated data visualizations to promote understanding of online privacy policies. Results to date are mixed. When study participants searched for products to purchase and saw a single icon view that evaluated the privacy practices for each site, they were willing to pay a small premium for more privacy-protective sites [24, 7]. On the other hand, translating an entire privacy policy into a grid that conveyed information by icons and colors did not improve comprehension [18]. Attempts at visualizing privacy are ongoing, including a recent set of icons modeled after Creative Commons [3]. This study, in contrast, examines three text-based formats as described below.

2.1 Privacy Finder

Privacy Finder (PF) was developed by AT&T and refined at the Cylab Usable Privacy and Security (CUPS) laboratory. Privacy Finder is a privacy-enhanced front end to Yahoo! and Google search. Privacy Finder has many components including a privacy report, which is the component we tested. Privacy Finder's privacy report was designed to avoid many of the problems that stem from

free-form natural language policies by generating standardized text from P3P policies. The Platform for Privacy Preferences (P3P) is a standardized format for privacy policies, and is formally recommended by the World Wide Web Consortium (W3C) [26]. P3P provides a taxonomy to express privacy practices in XML (eXtended Markup Language), which is computer readable and thus allows software tools to help people manage their privacy preferences.

Because Privacy Finder generates text from P3P tags, the Privacy Finder report avoids "weasel words" and ensures uniform presentation. However, Privacy Finder reports allow a free-form text description of the highest level of policy statements. This can improve readability by providing context for readers, but also means that companies with identical practices may have different Privacy Finder reports.

2.2 Layered Notices

The law firm Hunton & Williams popularized the notion of layered notices [22] with the goal of a short notice that is brief, standardized, and easy to compare directly which then links to the full policy. The first layer provides a short overview with required standardized headings. Although the text within each section is free form, layered policies are typically only one screen of text. As a result of this brevity the first layer omits many details and links to the second layer, which is a full natural language policy.

By 2005, several large companies deployed layered policies including Microsoft (MSN), Procter & Gamble, IBM, and JP Morgan [14]. European Union Information Commissioner Richard Thomas called for the use of layered policies in response to research showing nearly 75% of participants said they would read privacy policies if they were better designed [16]. Article 29 of European Union Directive created the "Working Party on the Protection of Individuals with regard to the processing of Personal Data," which issued guidance on how to create layered policies [27]. Privacy commissioners in EU countries supported layered policies. In Australia, the Privacy Commissioner released a layered policy for their own office, intending it "as a model for other agencies and organisations" [23].

2.3 Natural language

Most privacy policies are in natural language format: companies explain their practices in prose. One noted disadvantage to current natural language policies is that companies can choose which information to present, which does not necessarily solve the problem of information asymmetry between companies and consumers. Further, companies use what have been termed "weasel words" — legalistic, ambiguous, or slanted phrases — to describe their practices. Natural language policies are long, require college-level reading skills, and information is not in a standard place or described using consistent language.

3 Methods

We conducted an online study from August to December, 2008 in which we presented a privacy policy to participants and asked them to answer questions about it. We posted advertisements on craigslist and used personal networks to recruit participants. We offered a lottery for a chance to win one of several $75 Amazon gift certificates as incentive for participating in the study.

We ran a between-group design and assigned each participant to one of fifteen privacy policy representations. We used a between-group design rather than within group design because in this context it is unrealistic to eliminate learning effects simply by reordering policies. Reading the questions could affect how participants read subsequent policies. Second, it is unrealistic to expect participants to spend more than 20 minutes completing an online survey. Questions remained constant over all conditions; only the policy differed.

3.1 Study Conditions

We contrasted six different companies' conventional natural language (NL) policies and their corresponding Privacy Finder privacy report format (PF) plus three layered policies. We refer to these companies as A through F. We analyzed 749 participants across 15 conditions, for an average of 50 participants per condition. Note that we did not study layered policies for companies A, C, and E. The study conditions are listed in table 1.

Table 1. Participants per Condition

Company	NL	PF	Layered
A	41	50	N/A
B	47	46	52
C	46	41	N/A
D	47	47	49
E	52	51	N/A
F	62	55	63

For natural language polices we used black text on white backgrounds regardless of the original graphic design. We left other formatting that might aide comprehension (for example, bulleted lists) intact. We replaced all companies' names with "Acme" to avoid bias from brand effects.

Of the six companies, only B and D had layered policies. We followed the directions from the Center for Information Policy Leadership to create a third layered policy for company F [4]. We created this policy prior to designing the study questions.

As deployed in practice, Privacy Finder highlights the most important information at the top of the report and provides links to expand details. We discovered in earlier testing that people rarely expanded the Privacy Finder report.

We were interested in testing how well people are able to use the information in the Privacy Finder report, not how well they are able to navigate the user interface so in our research we presented all information in a single flat file.

We selected privacy policies from six popular websites that engage in e-commerce, and thus must collect a variety of personal information as part of their business. We chose what we believe to be a comparatively easy to read and a comparatively difficult to read policy with several typical policies.

We selected policies guided by several measurements of readability summarized in table 2. For each company, we noted the length of the natural language policy. We calculated the Flesch-Kincaide Reading Ease Score, which ranges from a low of 1 to a high of 100 based on syllable count and line lengths. High Flesch-Kincaide scores are more readable than low scores. In general, experts suggest a score of at least 60—70, which is considered easily understandable by 8th and 9th graders [15]. Reader's Digest has a readability index in the mid 60s, Time is in the low 50s, and Harvard Law Review in the low 30s [11]. Note that while the policies we selected span a range from 32 to 46, even the most readable policy is more challenging than is normally recommended for a general audience.

We calculated the percentage of sentences written in the passive voice, which is both more difficult for readers to understand and an indicator the company may not be comfortable taking full responsibility for their privacy practices. We counted the number of cross references within each policy; the more times readers are asked to refer to other parts of the document the more difficult it is to understand. Finally, we note that the standardized Privacy Finder format also has a range of lengths due to differing numbers of statements, how much information they collect, and how much text the policy authors elected to supply.

Table 2. Attributes of six company's privacy policies

Company	NL Words	NL Pages	Reading Ease	% Passive	Cross references	PF Words
A	6329	13	31.8	11%	27	880
B	3725	7	35.5	22%	0	1964
C	2920	6	36.3 1	7%	7	2011
D	2586	8	42.8	18%	2	554
E	2550	8	44.9	11%	0	1373
F	928	3	46.3	9%	1	1843

3.2 Study Questions

Study questions comprised several groups:

– *Comprehension.* Participants answered a series of multiple choice questions to determine how well they were able to understand the policy. These questions are realistic information retrieval tasks based on typical privacy concerns, and are similar to questions used in an earlier study by Cranor et al

[6]. We conducted three rounds of pilot tests with over two dozen people to ensure the questions were well-worded and understandable. We randomized the order of these questions to mitigate learning effects and captured both accuracy and time to respond. We also included a warm-up task which we did not score.

- *Psychological Acceptability.* Saltzer coined the term psychological acceptability to convey that if people do not like a system they will not use it. He wrote, "It is essential that the human interface be designed for ease of use, so that users routinely and automatically apply the protection mechanisms correctly" [20]. Participants answered subjective questions on a seven-point Likert scale.
- *Demographics.* We collected basic information like gender, educational attainment, and income so we could understand how closely our study population resembles Internet users as a whole.

We also measured the time it took for participants to answer each one of the comprehension questions.

3.3 Analysis

We performed a comparative analysis across all three formats (Natural Language, Privacy Finder, and Layered) and from all six companies to see if there were statistically significant differences in the mean scores for accuracy, time to completion, and psychological acceptability questions.

After we removed outliers[3] we performed ANOVA analysis for both time data and psychological acceptability, which we recorded on a seven point Likert scale and treated as continuous variables. Accuracy questions were categorical data (either accurate or false) so we used Chi Squared tests. We performed all tests of statistical significance at the $\alpha = 95\%$ confidence level.

[3] We only included results from participants who completed all of the accuracy questions. Because this was an online study to enter a drawing for a gift certificate, a few people just "clicked through" answers without engaging with the material. We picked a fixed lower threshold of 1.5 seconds per question and removed participants entirely if they had two or more questions they answered in under 1.5 seconds (7 participants removed out of an original 756 for a total of 749.) For participants with only one time under 1.5 seconds, it is possible they accidently double-clicked once but answered other questions properly. We removed the time and accuracy data for just the affected question (3 question/time pairs out of 3000.) At the other extreme, sometimes people were diverted by other tasks while answering questions and we recorded unduly long times to answer. We discarded question times in excess of 2.5 times the mean for their condition along with their corresponding answers. This resulted in $N = 723$ for cookies, 728 for opt out, 726 for share email, and 723 for the telemarketing questions.

4 Accuracy and Speed Results

Accuracy scores are all reported as the percentage of people who answered the question correctly. Answers are always either Yes, No, or the policy Does Not Say. Higher percentages are always better. In some cases, participants may have been confused about when to use Does Not Say, so we also report the combined percentage of participants who answered either Yes or Does Not Say. As compared to natural language, we found that layered policies led to lower accuracy scores for topics not in the short layer. Privacy finder was indistinguishable from natural language until questions became harder, at which point privacy finder was superior to natural language.

Accuracy spanned a wide range. An average of 91% of participants answered correctly when asked about cookies, 61% answered correctly about opt out links, 60% understood when their email address would be "shared" with a third party, and only 46% answered correctly regarding telemarketing (arguably, some of these scores are actually higher: see the full discussion of Does Not Say below). Recall that with only three possible answers, if participants guessed randomly we would expect 33% accuracy.

Times are in minutes. All other things being equal, lower times are better. Participants answered more quickly with both layered and privacy finder formats. Times to answer increased with question difficulty, with an average of 2.3 minutes to answer the question about cookies, 4.7 minutes to answer about opt out links, 5.3 minutes for email sharing, and 6.7 minutes for telemarketing.

Table 3. Percentage correct and minutes to answer, cookies question.

Policy	% correct	Time
A NL	87%	3.6
A PF	96%	1.5
B NL	96%	2.0
B PF	98%	1.6
B Layered	86%	2.3
C NL	93%	2.4
C PF	98%	3.5
D NL	86%	2.6
D PF	91%	1.9
D Layered	69%	2.2
E NL	96%	2.6
E PF	96%	1.8
F NL	100%	2.3
F PF	94%	2.7
F Layered	80%	2.3

4.1 Cookies

We asked: Does the Acme website use cookies?

Answer: Yes for all policies.

Most participants got the cookie question right (91%). This was an easy question to answer because our question is phrased with the same term the policies use. All policies, in all formats, call out cookies use explitictly. For example, one policy has a heading of "Cookies and Other Computer Information" with a paragraph that begins: "When you visit Acme.com, you will be assigned a permanent 'cookie' (a small text file) to be stored on your computer's hard drive." There is no ambiguity. Even someone who has no idea what a cookie is, or what the implications for privacy are, can skim through any of the natural language policies to find the word "cookie" and answer correctly.

We found significant differences in accuracy for company[4] and format.[5] For the six companies, the span between the worst performance (D, 82%) and best performance (E, 96%). See table 3 for a summary of results.

Layered policies gave participants a little more trouble (78%) than other formats. Cookie information was under the heading "Personal Information" in F Layered (80%,) which may not be where people expected to look. In D Layered (69%,) the policy mentions in passing that "You may also turn off cookies in your browser," without explicitly saying they use cookies. People must deduce that information or go to the full policy for a direct statement that the site uses cookies. This highlights two results we will see again: first, when participants needed to think about an answer rather than just perform a search for information, accuracy dropped. Second, it appears few people ventured beyond the first page of the layered policies.

Table 4. Percentage correct, correct including "Does not say" responses, and minutes to answer for the opt out question. Bold policies have correct answer of No.

Policy	% correct	with DNS	Time
A NL	33%	45%	5.7
A PF	85%	96%	3.7
B NL	33%	65%	9.3
B PF	91%	98%	4.6
B Layered	18%	38%	4.8
C NL	80%	86%	3.2
C PF	73%	90%	5.1
D NL	29%	67%	6.1
D PF	71%	89%	3.8
D Layered	19%	46%	5.5
E NL	55%	65%	5.4
E PF	51%	73%	4.6
F NL	93%	95%	3.4
F PF	79%	91%	3.7
F Layered	92%	93%	2.2

In another sign that this was an easy question for most participants, times to answer were shorter than the other questions (2.3 minutes.) We found no significance for time based on company[6] but format was significant.[7] Privacy Finder (2.1 minutes) and Layered (2.3 minutes) supported faster responses, but the Layered condition was also more likely to result in incorrect answers.

4.2 Opt Out Link

We asked: Does the company provide a link to a webform that allows you to remove yourself from Acme's email marketing list?

Answer: Yes for all policies except: B NL, D NL, D Layered, E NL, which are No.[8]

[4] χ^2(d.f. 5)=12.16, $p = .033$
[5] χ^2(d.f. 2)=28.95, $p < .001$
[6] $F(5)$=1.18, $p = .320$
[7] $F(2)$=4.50, $p = .012$
[8] Answers are not the same across a given company because the companies elected to provide different information in different formats. P3P requires an opt out link, which is then included in Privacy Finder.

This question is a little more difficult than the question about cookies. Policies refer to this concept as "opting out." For example, one policy phrases it as "To opt out of receiving all other Acme mailings after you have registered, click here or click the appropriate unsubscribe link contained within the email that you receive." Participants need to map the concept of removing themselves from an email marketing list to the technical jargon of opting out. However, this question is again fairly straight forward. Either there is an opt out link or there is not. See table 4 for a summary of results.

Interpreting results is complicated by potential confusion of how participants answered when there is no opt out link. The straight-forward answer we envisioned is "No." However, participants may also have replied that the policy "Does Not Say", intending to convey the same information since there is no opt out link within the policy. Arguably, the correct way to score responses is to combine No and Does Not Say. However, where policies do have an opt out link, answering Does Not Say (DNS) is flatly incorrect: the link was there, but the study participants missed it.

We found significant differences for company both without DNS[9] and with DNS.[10] Format is also significant without DNS[11] and with DNS.[12] Natural language policy accuracy rates are dissimilar, with averages ranging from 93% (F) to 33% (A). Finding the opt out link in the A NL policy was looking for a needle in a haystack: there is one link halfway through the policy in the middle of a paragraph without any headings or other cues—and the policy runs to 13 pages when printed.

It would seem Privacy Finder should have consistent results across all six policies, since an opt out link is a standard part of Privacy Finder reports. However, companies with an opt out default have additional links for each category of opt out data. As a result, policies with worse practices fared better, ranging from 85% correct for a A PF with less privacy protective practices that had many prominent opt out links, to 51% correct for E PF which required opt out for all data collection. While 51% is a low accuracy score, this is not a bad direction for the outcome: opt out links are easiest to find when they are most valuable. Interestingly, the F PF policy (79%) has identical practices as E PF (51%) yet different accuracy scores. The author of the F PF policy included an additional opt out link in the text at the very end of the policy, which is prime real estate for readers' attention. Policy authors choices affect outcomes, even within the PF standardized presentation.

Since there is no requirement to discuss opt out choices within the layered format, once again we see dissimilar results across a standardized format. B layered policy (18%) required clicking the opt out link to see what it did, phrased as "For more information about our privacy practices, go to the full Acme Online Privacy Statement. Or use our Web form," with a link from "Web form" to the

[9] χ^2(d.f. 5)=108.31, $p < .001$
[10] χ^2(d.f. 5)=53.44, $p < .001$
[11] χ^2(d.f. 2)=40.80, $p < .001$
[12] χ^2(d.f. 2)=53.44, $p < .001$

opt out page. In contrast, results were quite good with F layered (92%), which contained the same opt out text as at the end of the F PF (79%) policy.

We found significance differences in time to answer for company[13] as well as format.[14] We would expect longer times for longer policies since this is in many ways an information search task. Instead, time appears to be based on the underlying practices: policies without opt out links took longer. Since some of the policies with opt out links mentioned them at the end, it is unlikely the difference in times is based on needing to read through the entire policy to determine the absence of a link. Instead, participants likely re-read to satisfy themselves that they had not missed anything. Once again participants completed the task more quickly with layered (4.0 minutes) and Privacy Finder (4.2 minutes) than Natural Language (5.4 minutes,) but the wide variance and sometimes poor performance for standardized policies reduces the strength of this result.

4.3 Share Email

We asked: Does this privacy policy allow Acme to share your email address with a company that might put you on their email marketing list (with or without your consent)?

Answer Yes for all policies except: companies E and F (all formats) which are No.

We tested the wording of this question in multiple pilot studies to ensure people understood it without asking something pejorative or jargon-laden like "will Acme sell your email address to spamers." This question requires participants understand the question, read the policy carefully, and make inferences for most policies. For example, C NL reads: "We may provide your contact information and other personal data to trusted third parties to provide information on products and services that may be of interest to you." Participants need to understand that "contact information" includes email, that "trusted third parties" are companies other than Acme,

Table 5. Percentage correct, correct including "Does not say" responses, and minutes to answer for the email sharing question. Bold policies have correct answer of No.

Policy	% correct	with DNS	Time
A NL	76%	88%	3.2
A PF	53%	71%	5.4
B NL	49%	51%	5.9
B PF	64%	69%	5.9
B Layered	52%	52%	4.8
C NL	80%	86%	4.7
C PF	72%	79%	6.9
D NL	67%	85%	4.6
D PF	78%	87%	4.0
D Layered	56%	73%	4.7
E NL	53%	63%	6.9
E PF	44%	56%	6.2
F NL	50%	57%	6.0
F PF	54%	59%	4.4
F Layered	62%	78%	5.0

[13] $F(5)=5.58$, $p < .001$
[14] $F(2)=3.59$, $p = .028$

and that "provide information on products and services" means marketing messages, in order to correctly answer "Yes". Arguably in this case, "Does Not Say" is also a correct answer, since the policy is so indirect. See table 5 for a summary of results.

Overall accuracy was only 60%. We found significant differences for company without DNS[15] and for company with DNS.[16] Format was not significant without DNS[17] and not significant with DNS.[18] Times to answer averaged 5.3 minutes, which indicates people had a harder time completing this task. We found no significant results for time based on company[19] or format.[20]

As the answers to our questions become more nuanced we would expect the more readable policies to shine, yet that is not the case. Company A, with the hardest to read policy, had a higher accuracy score (64%) than F (55%) with the most readable policy and there was no overall discernible pattern based on readability. Similarly, we would expect standardized policies to convey information better, especially the Privacy Finder format which avoids the emotion-rich wording of "trusted third parties" and "valuable offers," yet we did not find significant differences between formats. Privacy Finder summarizes "With whom this site may share your information" as "Companies that have privacy policies similar to this site's" which again requires participants to refer to a separate section to determine if the parent company may engage in email marketing.

4.4 Telemarketing

We asked: Does this privacy policy allow Acme to use your phone number for telemarketing?

Answer Yes for all policies except: companies A, E and F (all formats) which are No.

Participants struggled with this question as shown in table 6. Except in the Privacy Finder version where companies are required to provide information about their telemarketing practices, policies typically do not highlight telemarketing practices. The way to answer this question correctly was typically to read through the entire policy for all mentions of when the company collects phone numbers, then see what policies they have around that data. For example, B NL discloses telemarketing as: "You may also have the option of proactively making choices about the receipt of promotional e-mail, telephone calls, and postal mail from particular Acme sites or services." Sometimes policies were even more vague, for example D NL, "The information you provide to Acme on certain Acme Web sites may also be used by Acme and selected third parties for marketing purposes. Before we use it, however, we will offer you the opportunity to

[15] χ^2(d.f. 5)=22.43, $p < .001$
[16] χ^2(d.f. 5)=37.05, $p < .001$
[17] χ^2(d.f. 2)=1.90, $p = .387$
[18] χ^2(d.f. 2)=0.20, $p = .903$
[19] $F(5)$=1.81, $p = .109$
[20] $F(2)$=0.15, $p = .864$

choose whether or not to have your information used in this way." Not only is telemarketing swept under the phrase "marketing purposes," telephone numbers are not mentioned explicitly either. It was necessary to deduce practices from a very careful and nuanced reading, frequently referring to multiple sections of the policy and then putting pieces together like a jigsaw puzzle. One could even make the case that answering "The policy does not say" is correct in cases as above where "information you provide" may be used for "marketing purposes" is by no means an explicit statement about telemarketing. However, we think it is important to note that the company likely does believe they have conveyed their practices: privacy policies are vetted by lawyers and are generally expected to be able to withstand a court (or FTC) challenge. If necessary, companies can point to the language in their policy and show that they did not violate the text by telemarketing.

We found significant differences for company without DNS[21] and company with DNS.[22] Format was significant without DNS[23] but not significant with DNS.[24] We found no significant results for time based on company[25] but format does have significant differences.[26] Once again layered (5.7 minutes) and Privacy Finder (5.5 minutes) are an improvement over natural language (8.2 minutes) but with the caveat that layered does not do as well for accuracy.

Even though we called out D NL as particularly indirect, it falls solidly in the middle of the accuracy scores (42%, or 87% if we count Does Not Say as an accurate response.) Once again, more readable policies do not seem to fare particularly better.

When participants cannot find information in layered policies, by design they should continue to the full policy for more details. In practice this appears not to happen, with a very low accuracy of 28% omitting Does Not

Table 6. Percentage correct, correct including "Does not say" responses, and minutes to answer for the telemarketing question. Bold policies have correct answer of No.

Policy	% correct	with DNS	Time
A NL	23%	48%	8.7
A PF	43%	65%	5.9
B NL	41%	68%	6.7
B PF	67%	72%	5.9
B Layered	16%	68%	6.2
C NL	42%	69%	9.2
C PF	68%	75%	5.5
D NL	42%	87%	7.6
D PF	82%	87%	3.2
D Layered	33%	77%	5.5
E NL	65%	78%	10.2
E PF	56%	78%	5.4
F NL	26%	87%	7.1
F PF	55%	87%	7.4
F Layered	34%	95%	5.9

[21] χ^2(d.f. 5)=24.99, $p < .001$
[22] χ^2(d.f. 5)=44.34, $p < .001$
[23] χ^2(d.f. 2)=50.08, $p < .001$
[24] χ^2(d.f. 2)=0.20, $p = .217$
[25] $F(5)$=1.75, $p = .122$
[26] $F(2)$=8.59, $p < .001$

Say, jumping to 81% when Does Not Say is counted as a correct answer. This is why format loses significance with Does Not Say included: participants appear not to seek information beyond the initial screen of the layered policy.

Privacy Finder does support more accurate answers (61%) even in contrast to natural language (39%) when Does Not Say is omitted. Privacy Finder is the only format that requires a company to disclose, yes or no, if they telemarket. For example, under the heading "The ways your information may be used:" D PF includes "To contact you by telephone to market services or products – unless you opt-out." Privacy Finder does seem to enable more accurate answers, but again there is a lot of variation between companies based on the supplemental text they provide. For example B PF, is particularly confusing by stating in free form text "While Acme does not currently support telemarketing, it is possible that in the future Acme properties may contact you by voice telephone," directly above an automatically generated statement that they may use information for telemarketing.

5 Psychological Acceptability Results

After completing the initial information search tasks, participants answered a series of questions designed to elicit their emotional reactions. Participants responded on a from 1 = strongly disagree to 7 = strongly agree. Most answers hovered right around 4, which is a neutral reaction. Higher numbers are better.

5.1 Ease of Finding Information

We asked four questions about how easy it was to find information. We expected responses to these questions to reflect how well participants were able to understand a particular policy, and thus be related to the accuracy questions and times. However, we found few significant results and participants found layered easier to understand even though they were less likely to answer questions accurately.

- "I feel that Acme's privacy practices are explained thoroughly in the privacy policy I read" (M = 4.7, s.d. = 1.5). We found significant effects for company.[27] but not format[28]
- "I feel confident in my understanding of what I read of Acme's privacy policy" (M = 4.7, s.d. = 1.6). We found no significant differences between companies[29] or formats.[30]
- "This privacy policy was easier to understand than most policies" (M = 4.5, s.d. = 1.5). We found no significant differences between companies[31] but did

[27] $F(5)=2.4$, $p = .038$
[28] $F(2)=1.6$, $p = .203$
[29] $F(5)=1.9$, $p = .099$
[30] $F(2)=.33$, $p = .722$
[31] $F(5)=1.6$, $p = .148$

find significant results for formats.[32] Layered (M=4.8) scored better than natural language (M=4.4) or privacy finder (M=4.4).
– "It was hard to find information in Acme's policy" (M = 3.8, s.d. = 1.6). We found no significant differences between companies[33] or formats.[34] (Note that based on the wording for this question we had to report the inverse of responses to keep higher numbers as better.)

5.2 Trust

If a format conveys information well but results in lack of trust of the company, it is unlikely that corporations will adopt the format. Participants trusted Privacy Finder formats slightly more than other formats.

– "I feel secure about sharing my personal information with Acme after viewing their privacy practices" (M = 4.0, s.d = 1.7). We found significant effects for both company[35] and format.[36]
– "I believe Acme will protect my personal information more than other companies." (M = 4.0, s.d = 1.6). We found significant effects for both company[37] and format.[38]

5.3 Enjoyment

We asked two questions to gauge how much participants liked reading the privacy policy. If people are unwilling to read policies then improving them does not provide much benefit. We found no significant differences between formats.

– "Finding information in Acme's privacy policy was a pleasurable experience" (M = 3.7, s.d. = 1.7). We found no significant differences between companies[39] or formats.[40] This was the lowest score of all eight psychological acceptability questions.
– "If all privacy policies looked just like this I would be more likely to read them" (M = 4.2, s.d. = 1.7). We found significant effects for company[41] but not format.[42]

[32] $F(2)$=2.98, $p = .051$
[33] $F(5)$=.75, $p = .589$
[34] $F(2)$=.60, $p = .549$
[35] $F(2) = 14.4$, $p < 0.001$
[36] $F(5) = 7.0$, $p < 0.001$
[37] $F(2) = 8.0$, $p < 0.001$
[38] $F(5) = 3.9$, $p = 0.020$
[39] $F(5) = 1.7$, $p = .135$
[40] $F(2) = .62$, $p = .539$
[41] $F(2) = 2.4$, $p = 0.032$
[42] $F(5) = 2.4$, $p = .096$

6 Discussion

Many researchers start from the observation that privacy policies are not usable in their current format, and suggest ways to fix the problem. All of the formats were tested were unsatisfactory with a low rate of comprehension on questions that required synthesis of information. Participants did not like privacy policies of any type, and the highest mean score on the psychological acceptability questions was barely above neutral.

The standardized formats we studied still offer policy authors quite a bit of leeway. Companies with identical practices conveyed different information. Layered policies and the Privacy Finder report format supported faster decision making than natural language. However, layered policies were no better than natural language for accuracy and in many cases were worse. Participants appeared not to go beyond the initial layer which frequently left them with incorrect impressions of the company's privacy practices. Privacy Finder was not an improvement over natural language for easy questions, but did support more accurate answers for complex questions. While the accuracy scores for Privacy Finder were low in some cases, the format does represent a step forward from the status quo.

Privacy researchers tend to talk about policies as being uniformly bad. We expected that more readable natural language policies would have higher accuracy scores, lower times, and improved psychological acceptability than less readable policies, but that was not the case. These results could suggest that readability metrics are not a good way to differentiate between policies. This seems unlikely because the Flesch index has proven robust in many contexts and we do not immediately see any reason why privacy policies should be dramatically different from other types of textual analysis. It could instead be the case that the range from 32 to 46 on the Flesch index is too similar to see major variations in outcome: even the most readable policies are too difficult for most people to understand them and even the best policies are confusing.

Our study used a between subjects rather than within subjects structure. We expect that we would see larger differences, particularly in psychological acceptability, if we were to place policies side-by-side. Prior work[6] found that when participants have both the natural language and the Privacy Finder versions available, Privacy Finder fares well. By only showing one policy, our study did not capture one of the potential advantages to standardized formats. Standardized formats should become more useful once readers understand where to find information.

Early results testing a new format for privacy policies based around a nutrition label concept are encouraging [13]. Ideally, future formats will identify problems with existing approaches and attempt to improve upon what has come before.

References

1. Acquisti, A., and Grossklags, J. Privacy and rationality in individual decision

making. *Security & Privacy Magazine, IEEE 3*, 1 (January-February 2005), 26–33.
2. Anton, A., Earp, J. B., Qingfeng, H., Stufflebeam, W., Bolchini, D., and Jensen, C. Financial privacy policies and the need for standardization. *IEEE Security & Privacy 2*, 2 (Mar-Apr 2004), 36–45.
3. Bendrath, R. Icons of privacy, May 2007. http://bendrath.blogspot.com/2007/05/icons-of-privacy.html.
4. Center for Information Policy Leadership. Ten steps to develop a multi-layered privacy policy, 2007. http://www.hunton.com/.
5. Children's Online Privacy Protection Act of 1998 (COPPA), Public Law No. 104–191, October 1998.
6. Cranor, L. F., Guduru, P., and Arjula, M. User interfaces for privacy agents. *ACM Transactions on Computer-Human Interaction (TOCHI)* (2006).
7. Egelman, S., Tsai, J., Cranor, L. F., and Acquisti, A. Timing is everything? the effects of timing and placement of online privacy indicators. In *CHI 2009* (Boston, MA, USA, April 2009).
8. Graber, M. A., D'Alessandro, D. M., and Johnson-West, J. Reading level of privacy policies on internet health web sites. *Journal of Family Practice* (July 2002).
9. U.S. Gramm-Leach-Bliley Financial Modernization Act of 1999, Public Law no. 106–102, November 1999.
10. Hochhauser, M. Lost in the fine print: Readability of financial privacy notices, July 2001. http://www.privacyrights.org/ar/GLB-Reading.htm.
11. Huang, H.-J. Language-focus instruction in EFL writing : Constructing relative clauses in definition paragraphs. In *2008 International Conference on English Instruction and Assessment* (2008).
12. Jensen, C., Potts, C., and Jensen, C. Privacy practices of Internet users: Self-reports versus observed behavior. *International Journal of Human-Computer Studies 63* (July 2005), 203–227.
13. Kelley, P., Bresee, J., and Cranor, L. F. A "nutrition label" for privacy, In Preparation.
14. Lemos, R. MSN sites get easy-to-read privacy label. *CNET News.com* (2005). http://news.com.com/2100-1038$_3$ − 5611894.*html*.
15. My Byline Media. The Flesch reading ease readability formula. http://www.readabilityformulas.com/flesch-reading-ease-readability-formula.php.
16. OUT-LAW News. Drop the jargon from privacy policies, says privacy chief, September 2005. http://www.out-law.com/page-5791.
17. Pollach, I. What's wrong with online privacy policies? *Communications of the ACM 30*, 5 (September 2007), 103–108.
18. Reeder, R. W., Kelley, P. G., McDonald, A. M., and Cranor, L. F. A user study of the expandable grid applied to P3P privacy policy visualization. In *WPES '08: Proceedings of the 7th ACM workshop on Privacy in the electronic society* (2008), ACM, pp. 45–54.
19. Report by Kleimann Communication Group for the FTC. Evolution of a prototype financial privacy notice, 2006.
20. Saltzer, J. H., and Schroeder, M. D. The protection of information in computer systems. *Proceedings of the IEEE 63* (September 1975), 1278–1308.
21. Sheng, X., and Cranor, L. F. An evaluation of the effect of US financial privacy legislation through the analysis of privacy policies. *I/S - A Journal of Law and Policy for the Information Society 2*, 3 (Fall 2006), 943–980.
22. The Center for Information Policy Leadership, H. . W. L. Multi-layered notices.

23. The Office of the Privacy Commissioner. Release of privacy impact assessment guide and layered privacy policy, August 2006. http://www.privacy.gov.au/news/06$_1$7.*html*.
24. Tsai, J., Egelman, S., Cranor, L. F., and Acquisti, A. The effect of online privacy information on purchasing behavior: An experimental study. In *The 6th Workshop on the Economics of Information Security (WEIS)* (2008).
25. Vila, T., Greenstadt, R., and Molnar, D. Why we can't be bothered to read privacy policies: models of privacy economics as a lemons market. *ACM International Conference Proceeding Series 5* (2003), 403–407.
26. W3C Working Group. The platform for privacy preferences 1.1 (P3P1.1) specification, November 2006. http://www.w3.org/TR/P3P11/.
27. Wire, B. European union issues guidance on privacy notices; new notices make it easier for consumers to understand, compare policies, 2005. http://www.investors.com/breakingnews.asp?journalid=24821135.