Visualizations of Terms and Conditions and Privacy Policies

Noshin Nawar Sadat
University of Waterloo
Waterloo, Ontario

Bhavendra Budnarain
University of Waterloo
Waterloo, Ontario

ABSTRACT

We compare three different models of display of terms and conditions (T&Cs) and Privacy Policies (PPs) for web services with respect to their effect on a user's understanding and ability to recall information. This comparison focuses more on how much key information users are able to retain in the shortest amount of time. The first model displays the clauses of T&Cs and PPs in their original format; the second is based on a summarized version of the original text of clauses; the third is based on a similar condensed format, but also provides animated visualizations for the key clauses. We find that accompanying the summarized versions of T&Cs and PPs with display of animations results in participants having an overall better (though not significant) understanding of the T&Cs and PPs. The use of animations also does not have any significant effect on the time required to read the T&Cs and PPs.

INTRODUCTION

People almost always agree to terms and conditions (T&Cs) and privacy policies (PPs) of web services without reading them, not knowing how these service providers are gathering their data and also probably selling them under their noses. The clauses of T&Cs and PPs are too long to read and are often loaded with legal jargon, making them difficult to comprehend. Most people do not have enough time to read such long, verbose and complicated texts. Hence, by blindly trusting the services, agree to them without reading them.

According to McDonald and Cranor, an average American consumer comes across 1,462 privacy policies on average per year [16]. If people are indeed signing up for all these agreements without reading them, then they have legally bound themselves to thousands of clauses with these different services on the internet, without knowing what they entail. This is a huge risk to not only their personal but may be also to their work and social life.

We wanted to make it easier for people to read and understand T&Cs and PPs and also wanted to make sure that reading these did not take too much of time from their busy lives. We also wanted it to be less monotonous process. For this purpose, we wanted to come up with animated representations of different clauses to indicate what they were talking about. Then we planned to observe if visualizing the different terms of the web services in this way makes it easier for people to read and comprehend them.

The two main hypotheses of our research were:

1. Using animated visualizations to represent the clauses of T&Cs and PPs took less time for people to understand what they meant.

2. Animated visualizations made it easier for people to retain information conveyed by the clauses of T&Cs and PPs.

RELATED WORK

People do not read T&Cs and PPs

A comprehensive study on consumer behavior towards T&Cs was conducted by Elshout et al. in Europe, which showed that 90% to 95% people accepted T&Cs [14]. However, only 9.4% people opened the T&Cs if opening it was optional but about 77.6% people scanned through them if scrolling was the only option. This, however, does not mean that people are not concerned about their privacy and the use of their personal data. Another study conducted in Europe showed that 67% people were concerned about not having control over their own information and 57% agreed that providing personal information on the internet is a big issue [1]. Yet, only 18% of the respondents of that study read privacy statements fully, 49% partially and 31% never read them.

Obar and Oeldorf-Hirsch had conducted a study on 543 participants where they were asked to sign up to a fictitious social networking site [18]. It was found that 74% of the participants agreed to the T&Cs and privacy policies without reading them, where one of the clauses stated that they would have to give away their first born child as payment for accessing the website. Bakos et al. showed that only two out of every thousand retail software shoppers read license agreements [9]. Those who read, were mostly older in age and had more time to spare on reading the agreements. According to Bakos et al., it was easier for the older readers to understand the contents of the agreement as they were more educated. Milne and Culnan found through their experiment that most of those who read privacy notices, read them due to concerns about privacy. Some of the respondents mentioned that they at least skimmed through the notices to know about whether their personal information would be shared with any third parties or not and if so, which of them will be shared.

Obar and Oeldorf-Hirsch found through their experiment that among those who read the privacy policy (7977 words) and

CHI'16, May 07–12, 2016, San Jose, CA, USA

ISBN 123-4567-24-567/08/06...$15.00

DOI: http://dx.doi.org/10.475/123_4

T&C (4316 words), the average reading time was 73 seconds and 51 seconds respectively [18]. By studying the internet behavior of 48,154 monthly visitors of 90 online software companies, Bakos et al. also found that the visitors spent too little time to read the agreements [9]. However, according to McDonald and Cranor, 250 words per minute is the limit for any individual for thorough reading [16]. People, therefore, do not like reading much on the internet. In fact, Weinreich et al. argued that new pages with a lot of information and links are mostly viewed for brief periods [19].

Why people do not read T&Cs and PPs

The main reason behind people not reading T&Cs and PPs is the fact that they are too lengthy and difficult to understand. Obar and Oeldorf-Hirsch found that information overload (too long and takes up too much time) was a significant factor behind people's not reading T&Cs [18]. A survey conducted in Europe showed that 67% respondents thought PPs were too long, 38% thought they were too difficult and unclear, 15% thought their privacy would not be honored anyway, 14% thought that they would be protected by law, 14% thought just having a PP was enough, while 9% thought it to be of not much importance [1]. Elshout et al. stated that the length (ranging from 4,000 words to even up to 36,000 words) as well as the complex and technical language used in the T&Cs make it too time consuming for consumers and so they blindly agree to such T&Cs without reading them [14]. Böhme and Köpsell argued that internet users have been so well trained by the same age old design of highlighted 'I Agree' buttons on websites that they tend to click it without reading through the conditions even if they are very conscious about privacy issues [10].

Several studies were conducted to get an idea about the readability of the online T&Cs and PPs. McDonald et al. compared comprehensibility, psychological acceptability and time taken to answer questions of three format conditions of PPs of 6 companies [17].The formats were layered policies (short forms with standardized components along with full policy), Privacy Finder privacy report (standardized brief bulleted format of text descriptions of privacy policies) and the traditional non-standardized human readable policies. Participants took less time in both the standardized format conditions compared to the non-standardized format. Their accuracy scores were also higher than that of non-standardized format but the difference was very small. Participants disliked all formats equally. Anton et al. found that 28.3% of US adults are unlikely to understand PPs and would at least require high school education to understand fully [8]. While studying the PPs of health care websites, Ermakova et al. found that at least college level, 13th reading grade level or 16 years of formal education would be required to understand them[15].

No regulations

Since majority of people do not read the T&Cs and PPs online, it is of great concern whether some website owners would take advantage of this situation and not be transparent about their use of personal information of the users. Yet, no kind of regulations have been imposed to keep it in check due to an "informed minority" hypothesis, which states that having a minimum number of conscious users is enough to discipline the website owners. However, Bakos et al. proved that PPs on average have too small readership to say to not require regulation [9].

Existing solutions

Many methods and tools have been proposed and are being used today in order to aid users in reading and comprehending the T&Cs and PPs. Many websites (for example, Github) are now providing summaries of each clauses alongside the original clauses [3]. Some (for example, Spotify) have also created Table of Contents for easy navigability among sections [6]. Google uses animated videos to explain about their privacy policies [4].

Conducting an experiment among more than 80,000 internet users, Böhme and Köpsell found that the very design of the agreement pages affect people's tendency to read the T&Cs [10]. They found that some people, if politely given the choice to accept to certain conditions, do read through the conditions.

In 2002, World Wide Web Consortium (W3C) established Platform for Privacy Preferences (P3P), which defines a standard for encoding PPs in an XML format [13]. Its components include entity (contact information of owner), access (which data are kept), disputes (how to resolve disputes with website owner), data (types of data collected), purpose (how the collected data is used and if opting is possible), recipient (under what conditions data can be shared and if opting is possible), retention (information about periodic purging of data), consequence (provides human readable explanation of site's data practices). Although a few tools and software now use P3P, it has not yet been widely accepted.

'Terms of Service; Didn't Read' (ToS;DR) is a non-profit service that aims to classify web services based on analyzing the T&Cs [2]. In the process, they also provide summaries of the most important points of those web services. They have a browser plug-in compatible with Firefox and Chrome. However, the list of important points they show is created based on the ratings and reviews provided by users. Moreover, web services keep updating their T&Cs often. As a result, several of the summaries found on ToS;DR are outdated. Many web services are still not classified.

Zimmeck and Bellovin proposed an architecture called Privee [20]. If policy analysis results are already available in ToS;DR, then Privee shows that result to user. Otherwise, it does a rule and ML based classification into six categories on the client machine and then shows the result. The six categories are - collection (of personal information), profiling (of users with own and third party information), ad tracking, ad disclosure, limited retention (period of keeping information), encryption (for information storage and transmission). After classification, the PPs are then graded into 3 classes - A, B, C.

Another data visualization tool called Polisis has been developed recently which uses machine learning to analyze the privacy policies of any website on the internet and provides a flow chart showing what data of the user the website collects, with which third parties it shares the data and also what the user can do about it [5].

Visualization of T&Cs and PPs

We did not come across enough studies that showed whether usage of summaries or navigability options like layered notices had any effect on the readability of T&Cs and PPs. We also found P3P, ToS;DR, Privee and Polisis to have one common flaw. They are all handled from the user's side instead of the websites' side. Hence, databases would be required to update regularly as every year millions of websites get created and the existing websites also make amendments to their T&Cs and PPs. It should be the responsibility of the websites themselves to ensure that their users can read and comprehend the T&Cs and PPs. They should try to improve their visualization techniques so that users can get all the necessary information in a short time and without any difficulty. Yet, we found very little research work done to improve the visualization of T&Cs and PPs. We wanted to express T&Cs and PPs in such a way that users would be able to comprehend it in a short amount of time and also be able to remember them later on.

PROPOSED TECHNIQUE

In order to improve recognition, recall and memorability of information, Borkin et al. suggested making use of colors, human recognizable objects (HROs), unique features and meaningful titles[11][12]. With their suggestion in mind, we decided to use graphic visualizations of the clauses of T&Cs and PPs.

Our idea is very similar to the method used by Google to show their privacy policies. However, what distinguishes our methods is that each of the animated videos that Google uses are more than 30 seconds long and they do not have videos for T&Cs. Moreover, it is not always possible for people to view 30 second long videos before signing up for every service. People can only increase or decrease the speed of the video play but have to view it consistently in order to get the information that matters them the most. On the other hand, what we proposed is the use of animated GIF (Graphics Interchange Format) images alongside summarized text of each of the original clauses of T&Cs and PPs, which could enable people to view and understand the clauses that matter to them at their own respective paces.

With our limited abilities in drawing we also attempted to give a slight comical touch to the animated GIF images to make them more interesting. we believe, if done right, they could draw attention of more people towards actually reading the clauses of T&Cs and PPs.

Although the initiatives taken by Privee, Polisis and ToS;DR are appreciable, it is not their responsibility to make users read the T&Cs of web services. The web service providers themselves should come up with ways to improve their presentation of T&Cs which will not only make users want to read the T&Cs, but also will require little time. They can, for example, follow the same approach as ToS;DR and ask their users to nominate the concerning points of the T&Cs. Moreover, instead of simply showing the summaries of those points and the original T&Cs, they can show graphic visualizations of those points as proposed by us. This could lead to increase in the numbers of people who actually go through the whole T&Cs or at least the main points that are of utmost concern.

IMPLEMENTATION

Method and Apparatus

We designed and implemented a web application following a client-server architecture model. The application was configured to display the terms and conditions according to the three test conditions which is described in the following section.

The web application was hosted at the University of Waterloo on a server with an open port making it accessible to participants. Participants were able to access the web application by pointing their web browser to the link which we provided. Participants brought their laptops and accessed the link through the Google Chrome web browser.

The back-end was implemented in Python3. The web server was able to process the HTTP requests sent from the client side. The web server supported the standard HTTP request method types such as, POST, GET, etc.

Participants were required to answer two questionnaires for which we used Google Forms.

The front end was implemented using HTML, CSS, and JavaScript. Additional libraries/frameworks that have been used included Bootstrap(HTML/CSS/JS library), jQuery (JS library) and AJAX for sending/receiving data asynchronously from the web server.

All of the data was logged using SQLite through Peewee which is a simple ORM for Python. SQLite is a lightweight and serverless database engine.

EXPERIMENT

Our aim was to test how our proposed visualization technique fared against other existing techniques of displaying T&Cs and PPs in terms of time taken and information retained by participants. For our experiment, we decided to compare our technique of displaying clauses of T&Cs and PPs with two existing techniques - displaying clauses in their original state and displaying a summarized version of the clauses. We ran a comparison of these three conditions in a controlled experiment. We hypothesized that users will be able to easily recall more information with visualizations of key parts of the terms and conditions in the shortest time. To design the experiment, we gathered ideas from the experiment conducted by McDonald et al. [17].

Procedure, Task and Design

The terms and conditions and privacy policies of Youtube were used for this experiment but we renamed Youtube to Terebi (Japanese word for television) so that participants do not know which website is being talked about. This will prevent their previous knowledge about the terms and policies of Youtube from affecting their answers.

We used only eight clauses of T&Cs and PPs of Youtube for our experiment. This was because we had to conduct the experiment in class and we only had 20 minutes to run our full experiment. We were concerned that using full T&Cs and PPs would exceed our time limit. We selected clauses of varying length ranging from 14 to 196 words in order to prevent bias. The clauses were then summarized and their

corresponding animated GIF images were created. Using an online animation tool Toonator, we were able to sketch out visualizations depicting each clause [7].

Therefore, our three test conditions were as follows:

1. **Original:** Clauses were displayed in their original form and were taken directly from the web service. There were no modifications made to the terms and conditions (figure 1).
2. **Summarized:** Summarized versions of the clauses were displayed without visualizations (figure 2).
3. **Visualization:** Summarized versions of the clauses were displayed along with animated visualizations (figure 3).

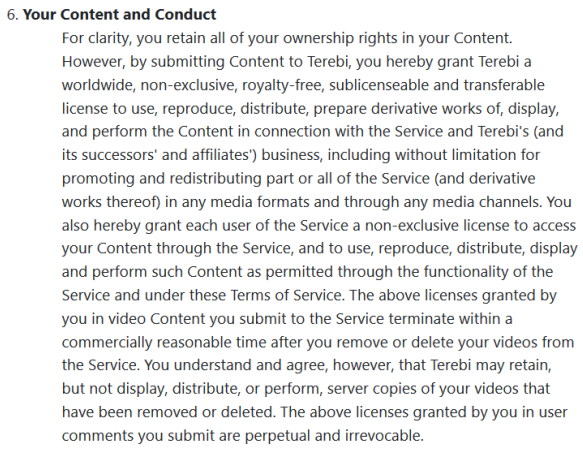

Figure 1. Example of a clause shown in Original condition

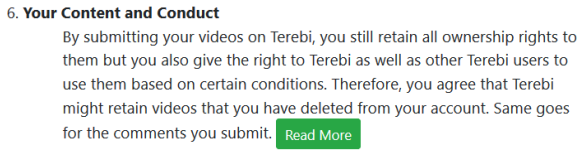

Figure 2. Example of a clause shown in Summarized condition

Figure 3. Example of a clause shown in Visualization condition

We conducted a pilot study with 2 participants before conducting the actual experiment, which allowed us to make amends to the web application, the questionnaires as well as the number of clauses to be used for the actual experiment. Our experiment followed a between group design to prevent carryover effects. Each participant was assigned a single test condition and was automatically assigned an alpha-numeric ID.

The experiment consisted of three main parts:

1. **Preliminary survey:** This was done to understand the participants' habits of interacting with T&Cs and PPs. The same questionnaire (See Appendix A) was used among all participants across conditions.
2. **Viewing Clauses:** Upon submission of the previous survey answers, the participants were directed to the web page that displayed the clauses corresponding to the condition that was assigned to them automatically. Since it was a between subject experiment, the same participant did not get to see more than one model of display. After they felt satisfied with their understanding of the clauses, they proceeded to the final questionnaire.
3. **Final Questionnaire:** The effectiveness of the approaches were evaluated through questionnaires provided to the participants. The questionnaire was designed after thorough review of existing literature. We evaluated their understanding of the key points of each clause through the questionnaire. Their were 8 such questions to test their understanding, one per clause. The questionnaire also included questions about the readability and acceptability of the clauses presented to them, and some demographic questions. The same questionnaire (See Appendix B) was used among all participants across conditions.

Since we wanted to know the effect of the three different conditions on time taken to read and ability to retain information, our dependent variables were time taken to read the clauses and the score of the participant out of 8 (8 points for 8 questions). Condition was our main independent variable. However, we also selected the following variables, the data for which we gathered through the questionnaires, as candidates for our independent variables:

- **Ever read full terms and conditions (ER):** Their previous experience in reading full T&Cs and PPs could affect their understanding of the clauses provided in a short time.
- **Ever skimmed through full terms and conditions (ES):** Their previous experience in reading full T&Cs and PPs could affect their understanding of the clauses provided in a short time.
- **Difficulty of terms and conditions in general (Diff):** If they found T&Cs and PPs to be difficult in general, chances were that they would find the provided clauses also difficult.
- **Age:** To find out if the claims of Bakos et al. [9] about age affecting the ability to understand T&Cs held in our experiment.
- **Native speakers of English (Native):** It would be easier for native speakers to understand the clauses in a short time.
- **Participant rating of own skills in English (Skill):** Those who rated themselves higher, would give better results.
- **Participant's confidence of understanding provided T&C (Confidence):** If they were more confident about their understanding of the provided clauses, then they would score better.

- **Readability of provided T&C (Readability):** If they found the provided clauses t be more readable, then they would score better and also take less time.

We had a total of 15 participants (2 female), consisting of graduate students enrolled in the CS889 course and the Professor teaching the course, allowing for 5 participant per test condition. Their ages ranged from 20 to 39.

Results

Out of the 15 participants, 9 (60%) claimed they came across 1 to 3 terms and conditions in the month prior to the experiment, while the rest mentioned that they came across more than 3 (figure 4). 11 (73.3%) of the participants mentioned that they always agreed to T&Cs and PPs they came across. Among those who did not always agree, 3 mentioned that they had not read the terms and were thus concerned about their rights and privacy being violated, 2 mentioned that they went through the terms and did not agree with some and 2 mentioned about not having trust in the service provider.

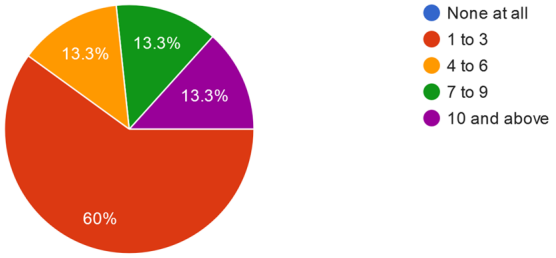

Figure 4. Number of times participants came across T&Cs and PPs in the month prior to the experiment

3 (20%) of the participants mentioned that they read each and every clause of terms and conditions of a web service at least once in their life and they spent average of 5.67 minutes (SD 4.04) to read them. However, they only read full T&Cs sometimes.

13 (86.7%) of the participants mentioned that they at least skimmed through the T&Cs and PPS to find clauses that matter to them and they spent an average of 4.08 minutes (SD 3.01) on it (figure 5). Out of them, only 2 mentioned that they always skimmed through T&Cs and PPs.

We wanted to know if the participants had ever noticed websites providing summaries of their T&Cs and PPs. 8 (53.3%) of them mentioned that they had seen summaries in websites before. Out of them, 4 always skimmed through the summaries.

In a Likert scale ranging from 1 to 5, where 1 meant very easy and 5 meant very difficult, the participants were asked to rate the difficulty of T&Cs and PPs in general. The result was a mean of 4.13 in difficulty (SD 0.83) .

A list of causes for not reading or skimming through T&Cs and PPs were provided to the participants. They were also asked to add their own causes, if any. Out of the 15 participants, 12 (80%) said the T&Cs and PPs were too long to read, 9

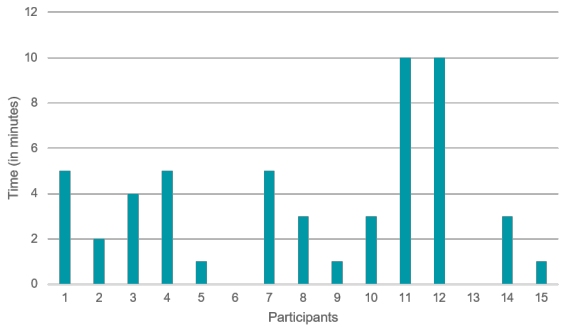

Figure 5. Time participants take to skim through T&Cs and PPs

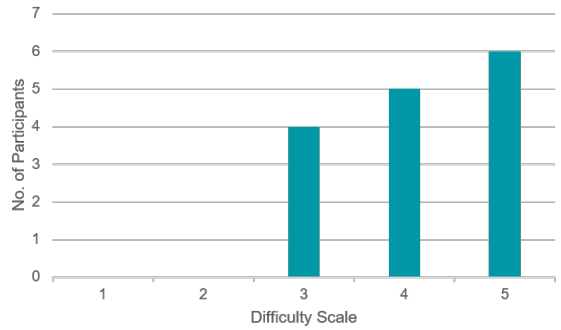

Figure 6. Difficulty of T&Cs and PPs as rated by participants

(60%) said they trusted the service they signed up for, 9 (60%) said they did not have time to read them, 5 (33.3%) said they could not understand the legal jargon, 3 (20%) said that their families and friends used the service and 2 (13.33%) said that they would have had to use the service anyway (figure 7). One interesting input from a participant was: "I'm too lazy to care that several websites probably own my soul and firstborn now".

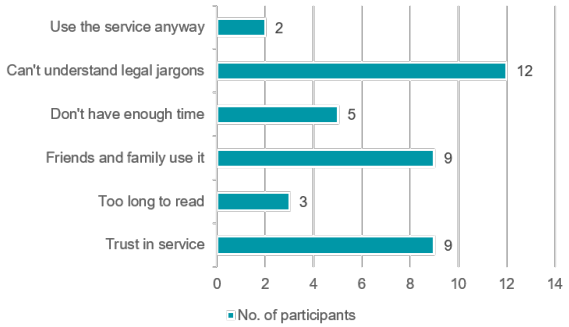

Figure 7. Causes for not always reading or skimming through T&Cs and PPs

When asked about the ideal length of time for them to read T&Cs and PPs, they responses of the participants averaged 3.36 minutes(SD 1.56) (figure 8.

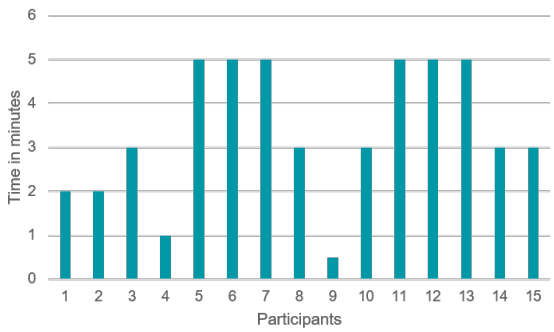

Figure 8. Ideal length of time for reading T&Cs and PPs for participants

4 (26.7%) of the participants were native speakers of English and most of the participants rated their English skills pretty high in a scale of 1 to 5 (mean - 4.53, SD - 0.64) (figure 9.

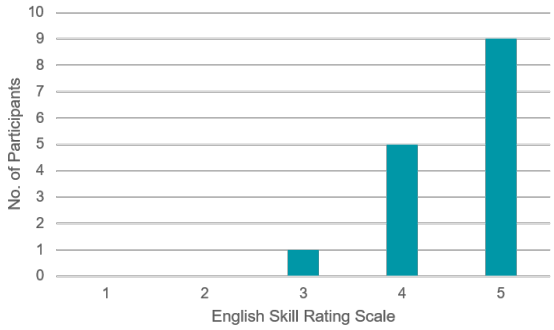

Figure 9. English skills of participants rated by themselves

We also asked participants about their confidence in their understanding of the clauses that were provided to them as well as the readability of those clauses. It was found that participants from the visualization condition were more confident about their understanding but found the display of clauses less readable.

Condition	Mean of Conf	SD of Conf	Mean of Read	SD of Read
1	3.8	0.45	3.6	0.55
2	3.0	1.22	3.2	0.45
3	4.2	0.46	2.6	0.89

Table 1. Confidence (Conf) and Readability (Read) scores (out of 5) by condition. Here, condition 1 : Original, condition 2 : Summarized and condition 3 : Visualization

In order to see the effects of the independent variables on time and score, we aggregated the data of time and score by the independent variables. Since condition was our main independent variable, all other independent variables were tested in combination with the condition variable.

The aggregation of time and score by condition are shown in Table 2. Although participants in the visualization condition

Condition	Mean of Score	SD of Score	Mean of Time	SD of Time
1	5.2	2.95	119.99	52.99
2	4.6	2.7	110.17	143.89
3	6.2	2.17	183.11	122.99

Table 2. Aggregation of score (out of 8) and time (in seconds) by condition. Here, condition 1 : Original, condition 2 : Summarized and condition 3 : Visualization

scored higher than the other two conditions, participants in the summarized condition unexpectedly scored less than those in the original condition. Moreover, we found that participants took longer time to read the clauses in the visualization condition compared to the other two conditions. Since the mean time and mean score still showed some difference across conditions, we performed statistical analysis to see if there is any significant effect of condition on either of them.

We first determined the normality of the time and score data. Through density plots, we found that time and score did not follow normal distribution, as shown in figures 10 and 11. We also performed Shapiro-Wilk normality tests on time and score data, where they yielded W = 0.88 (p.value = 0.05) and W = 0.83 (p.value < 0.05) respectively, further confirming their non-normal distribution.

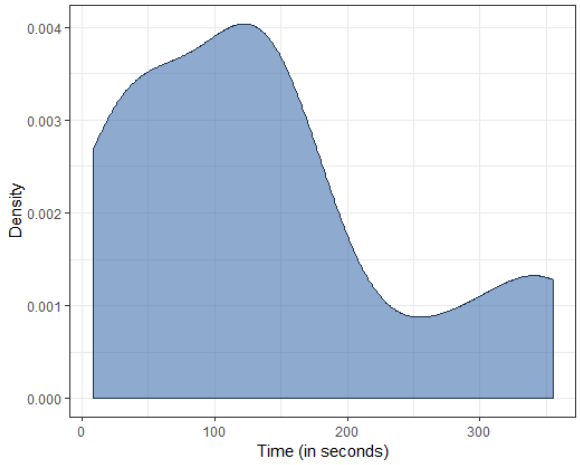

Figure 10. Density plot of time (in seconds) data

Since both the time and score data were not normally distributed and the experiments were done between-subject, we performed Kruskal Wallis test (a non-parametric test) on both time and score with condition as the independent variable. No significant effect of condition on either time (chi-squared = 1.82, df = 2, p.value = 0.4025) or score (chi-squared = 1.5919, df = 2, p.value = 0.4512) was found.

Out of the eight other candidate variables, we found that compared to others within the same condition, those who read full T&Cs before (See table 3), took less time and scored better, and those who were native English speakers (See table 4), took less time. So we also checked the interactions of these

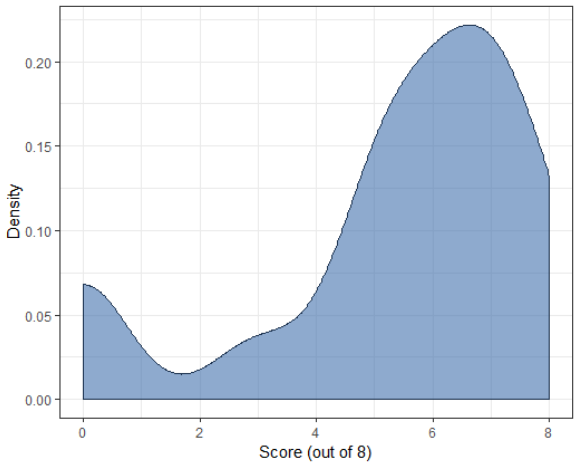

Figure 11. Density plot of score (out of 8) data

variables with condition variable to see if they had any effect on score or time.

Condition	ER	Mean of Score	SD of Score	Mean of Time	SD of Time
1	Yes	6.5	0.71	135.44	17.66
1	No	4.33	3.79	109.69	71.16
2	No	4.6	2.7	110.17	143.89
3	Yes	8.0	-	347.76	-
3	No	5.75	2.22	141.95	94.19

Table 3. Aggregation of score (out of 8) and time (in seconds) by condition and ER. Here, condition 1 : Original, condition 2 : Summarized and condition 3 : Visualization

Condition	Native	Mean of Score	SD of Score	Mean of Time	SD of Time
1	No	5.2	2.95	119.99	52.99
2	Yes	6.0	1.41	22.69	20.73
2	No	3.67	3.21	168.48	168.64
3	Yes	5.0	2.83	102.12	69.06
3	No	7.0	1.73	237.11	130.12

Table 4. Aggregation of score (out of 8) and time (in seconds) by condition and Native. Here, condition 1 : Original, condition 2 : Summarized and condition 3 : Visualization

Kruskal Wallis test on time for condition x ER interaction (chi-squared = 2.9375, df = 4, p.value = 0.5683) and condition x Native interaction (chi-squared = 4.4017, df = 4, p.value = 0.3544) showed no significant effect. Kruskal Wallis test on score for condition x ER interaction (chi-squared = 3.5557, df = 4, p-value = 0.4695) and condition x Native interaction (chi-squared = 3.3557, df = 4, p-value = 0.5002) also showed no significant effect.

DISCUSSION

From the analysis of the results of our experiment, we concluded that both our first and second hypotheses were invalid. However, when asked if they would read T&Cs and PPs when presented to them in the same format as their respective condition, 2 out of the 5 participants from the visualization condition said that they would always read them, while the other three said that they would sometimes read them. All 5 participants of the summarized condition and 4 participants from the original condition said that they would sometimes read them. 1 participant from the original condition said that they would never read them. Moreover, though it took longer mean time for participants to read the clauses in the visualization condition, they scored better in average compared to the other conditions.

We found the differences in mean scores and mean time to be not significant through the Kruskal Wallis test but the fact that we were constrained by the number of students in the CS889 course who were willing to participate in the experiment might have resulted in us not getting significant results. We had 15 participants in total which meant 5 participant per condition, which is a very low number to get any significant results through statistical analysis.

We tried to find out why participants from the visualization condition took longer time compared to those of other conditions to read the clauses when visualization condition had a lot less texts (the same amount of text as the summarized condition) with only added animations. Since participants in the summarized condition took the shortest amount of time and the only difference between the summarized and visualization condition was the animations, the delay might have been due to the animations themselves. The fact that they took longer time to go through the clauses could also mean that the animations had grabbed their attention and they spent more time on viewing them, which would be a good thing if it made people read clauses out of interest. We had not asked the users any specific questions about the animations themselves (if they were confusing or if they found them interesting, for example) and we also did not have any way to track down the participants from the visualization condition to ask them later. Perhaps, if we had told the participants that their time to read the clauses was being calculated, then we could have seen a different result. However, we wanted them to read the clauses at their own comfortable paces without thinking about time and so, we had not mentioned it.

We had expected participants in the original condition would get lower mean score compared to both the other conditions. However, we found that they did better than participants of the summarized condition. We observed that out of the three people people who had read full T&Cs and PPs at least once before, 2 participated in the original condition, while none of the participants in the summarized condition had ever read full T&Cs and PPs. The experiences of the participants who read full T&Cs and PPs before might have contributed to them getting better scores. Moreover, one of the participants in the summarized condition only took 8 seconds to read through the clauses, whereas the summarized clauses of 227 words were supposed to take 54 seconds to read, if we consider 250 words per minute as typical reading speed. This participant also

scored less compared to others in this test condition. Perhaps, they had not read all the clauses properly.

Moreover, we also observed that although non-native English speakers took more time compared to native speakers in both the summarized and visualization condition, the former scored higher in average compared to the latter. Perhaps, the animations made it easier for them to internalize the key points of the clauses. Although the number of participants in our experiment is too small to draw any concrete conclusions, this could be an indication that if the target users of a service includes non-native English speakers, then use of animated visualizations for T&Cs and PPs could make it easier for them to understand.

The large standard deviations of time also raised some concerns since it could mean that some of the participants were either too good at reading or too bad. It could also mean that some of the participants did not even read the clauses properly. Their scores could have been a good indicator of whether they had read the clauses or not, but we are also not sure if they had answered the questions based on what they had read or if they had made any guesses.

Furthermore, although the initiative taken by ToS;DR is appreciable, it is not their responsibility to make users read the T&Cs of web services. The web service providers themselves should come up with ways to improve their presentation of T&Cs which will not only make users want to read the T&Cs, but also will require little time. They can, for example, follow the same approach as ToS;DR and ask their users to nominate the concerning points of the T&Cs. Moreover, instead of simply showing the summaries of those points and the original T&Cs, they can show graphic visualizations of those points as proposed by us, that is, if further experiments proved our hypotheses as valid. This could lead to increase in the numbers of people who actually go through the whole T&Cs or at least the main points that are of utmost concern.

FUTURE WORK

For proper statistical analysis of any data, a minimum number of data points is necessary. In our case, we had only 15 data points, 5 per condition. Yet a minimum number of 8 data points per condition needs to be there in order to perform statistical analysis of a between subject experiment. Moreover, the participants in our experiment were all Computer Science majors, which could have yielded biased results. Therefore, one possible direction for future work is to perform this experiment with a larger number of participants with diverse educational and employment background.

We had initially planned to compare our technique with the technique used by ToS;DR and the original T&Cs and PPs of Youtube for our experiment. However, we were constrained by the time allocated to us for running the experiment. Since ToS;DR only shows the key concerning points of T&Cs and since we were also calculating the time required for reading the T&Cs and PPs, we had to make users of the other conditions read the full T&Cs to identify the key points. We were concerned that it might take longer than the allotted time to finish reading the original version of the T&Cs and PPs and also to answer all the questions of the questionnaires. Hence, we improvised and instead of using ToS;DR's technique for comparison, we used the summarization technique that several web services now use.

If a larger number of participants can be used and if there are no time constraints as there was during our experiment, then we could also compare our proposed technique with other techniques, such as, ToS;DR, Google's animated videos and Spotify's Table of Contents. In such cases, we could use the full T&Cs and PPs instead of using a few clauses from it. Such experiments could also help us to find answers to the questions that were raised during this experiment.

The longer times spent on our visualization condition could be explained by the possibility of the animations catching the user's attention. We did notice participants who were assigned the visualization condition appeared to be interested in understanding what message the animations were conveying as opposed to the text based conditions where not much time was spent. Although we hypothesized the time being less for the visualization condition, the usage of animations acting as an attention grabber may also be something worth looking into.

CONCLUSION

In order to make it easier for people to read the clauses of terms and conditions and privacy policies of web services, we proposed our own techniques of displaying the clauses. Our goal was to ensure that the clauses were easier to remember and also took less time to read. We proposed displaying the summarized versions of each clause along with animated GIF images that depicted the contents of those clauses. We compared our display technique with the conventional technique of displaying the clauses in their original form and with the technique of displaying only summaries of the clauses. It was a between subject experiment with 5 participants per test condition. The time taken to read the clauses were calculated and the the understanding of the clauses was scored through some questions asked after the reading.

It was found that there was no significant difference in either time or score for the different test conditions. However, participants who got our proposed technique, scored better compared to those who got the other two techniques.

Since the number of data points gathered from our experiment were very low, they did not show any significant results. However, we are optimistic that if a larger number of data points can be gathered from a more diverse participant group by replicating our experiment, then we might get significant results in favor of our hypotheses.

REFERENCES

1. 2015. Special Eurobarometer 431 "Data protection". (June 2015). DOI:http://dx.doi.org/10.2838/552336
2. 2016. Terms of Service; Didn't Read. (2016). Retrieved October 22, 2018 from https://tosdr.org/index.html.
3. 2018. GitHub. (2018). Retrieved November 8, 2018 from https://help.github.com/articles/github-privacy-statement/.

4. 2018. Google. (2018). Retrieved December 8, 2018 from https://policies.google.com/privacy?hl=en.

5. 2018. Polisis. (2018). Retrieved November 09, 2018 from https://pribot.org/polisis.

6. 2018. Spotify. (2018). Retrieved November 8, 2018 from https://www.spotify.com/ca-en/legal/privacy-policy/.

7. 2018. Toonator. (2018). Retrieved December 8, 2018 from https://toonator.com/.

8. A. I. Anton, J. B. Earp, Q. He, W. Stufflebeam, D. Bolchini, and C. Jensen. 2004. Financial Privacy Policies and the Need for Standardization. *IEEE Security and Privacy* 2 (2004). DOI: http://dx.doi.org/10.1109/MSECP.2004.1281243

9. Y. Bakos, F. Marotta-Wurgler, and Trossen D. R. 2014. Does Anyone Read the Fine Print? Consumer Attention to Standard Form Contracts. *New York University Law and Economics Working Papers* (2014).

10. R. Böhme and S. Köpsell. 2010. Trained to accept?: a field experiment on consent dialogs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 2403–2406. DOI: http://dx.doi.org/10.1145/1753326.1753689

11. M. A. Borkin, Z. Bylinskii, N. W. Kim, C. M. Bainbridge, C. S. Yeh, D. Borkin, H. Pfister, and Oliva A. 2016. Beyond Memorability: Visualization Recognition and Recall. *IEEE Transactions on Visualization and Computer Graphics* 22 (2016), 519–528. Issue 1. DOI: http://dx.doi.org/10.1109/TVCG.2015.2467732

12. M. A. Borkin, A. A. Vo, Z. Bylinskii, P. Isola, S. Sunkavalli, A. Oliva, and H. Pfister. 2013. What Makes a Visualization Memorable? *IEEE Transactions on Visualization and Computer Graphics* 19 (2013), 2306–2315. Issue 12. DOI: http://dx.doi.org/10.1109/TVCG.2013.234

13. L. F. Cranor. 2003. P3P : Making Privacy Policies More Useful. *IEEE Security and Privacy* 99 (2003), 50–55. Issue 6. DOI: http://dx.doi.org/10.1109/MSECP.2003.1253568

14. M. Elshout, M. Elsen, J. Leenheer, M. Loos, and J. Luzak. 2016. "Study on consumers' attitudes towards Terms and Conditions (TCs). (2016). Retrieved October 22, 2018 from https://ec.europa.eu/info/sites/info/files/terms_and_conditions_final_report_en.pdf.

15. T. Ermakova, B. Fabian, and E. Babina. 2015. Readability of Privacy Policies of Healthcare Websites. In *12th International Conference on Wirtschaftsinformatik (WI 2015)*. 1085–1099.

16. A. M. McDonald and L. F. Cranor. 2008. The Cost of Reading Privacy Policies. *I/S: A Journal of Law and Policy for the Information Society* (2008).

17. A. M. McDonald, R. W. Reeder, P. G. Kelley, and L. F. Cranor. 2009. A Comparative Study of Online Privacy Policis and Formats. In *Privacy Enhancing Technologies (PETS 2009)*. Springer, 37–55.

18. J. A. Obar and A. Oeldorf-Hirsch. 2018. The biggest lie on the Internet: ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society* (2018). DOI: http://dx.doi.org/10.1080/1369118X.2018.1486870

19. H. Weinreich, H. Obendorf, E. Herder, and M. Mayer. 2008. Not quite the average: An empirical study of Web use. *ACM Transactions on the Web (TWEB)* 2 (2008). Issue 1. DOI:http://dx.doi.org/10.1145/1326561.1326566

20. S. Zimmeck and S. M. Bellovin. 2014. Privee: An Architecture for Automatically Analyzing Web Privacy Policies. In *Proceedings of the 23rd USENIX Security Symposium*.

APPENDIX A : Preliminary Survey Questionnaire

Survey on Terms & Conditions and Privacy Policies

Please participate in this short survey for our project on Terms & Conditions (T&Cs) and Privacy Policies(PPs).

* Required

1. **ID (Please do not edit)**

_____

2. **How many times in the last month have you come across T&Cs and PPs while using any kind of product or service on the internet? Make an estimate if you are not sure.** *
   Mark only one oval.

   - ◯ None at all
   - ◯ 1 to 3
   - ◯ 4 to 6
   - ◯ 7 to 9
   - ◯ 10 and above

3. **Do you always agree to these T&Cs and PPs you come across?** *
   Mark only one oval.

   - ◯ Yes *Skip to question 5.*
   - ◯ No *Skip to question 4.*

Skip to question 5.

4. **What are the reasons behind not agreeing to certain T&Cs and PPs? (Select all that apply)** *
   Check all that apply.

   - ☐ Went through some or all clauses and found a few I cannot agree with
   - ☐ Did not read but was worried they might violate my privacy and/or other rights
   - ☐ Got distracted and never signed up
   - ☐ Other: _____

Skip to question 5.

5. **Have you ever read each and every clause of T&Cs and PPs of any online service or product?** *
   Mark only one oval.

   - ◯ Yes *Skip to question 6.*
   - ◯ No *Skip to question 8.*

6. **How long (in minutes) do you think you take to read them?** *

7. **How often do you read the whole T&Cs and PPs?** *
Mark only one oval.

- Sometimes *Skip to question 8.*
- Always *Skip to question 11.*

8. **Have you ever skimmed through the clauses of T&Cs and PPs to find specific issues that matter to you?** *
Mark only one oval.

- Yes *Skip to question 9.*
- No *Skip to question 11.*

9. **How long (in minutes) do you think you take to skim through them?** *

10. **How often do you skim through the clauses of T&Cs and PPs?** *
Mark only one oval.

- Sometimes
- Always

11. **Have you ever noticed summaries of clauses of T&Cs and PPs on any website?** *
Mark only one oval.

- Yes *Skip to question 12.*
- No *Skip to question 15.*

12. **How many websites did you notice provided such summaries?** *
Mark only one oval.

- All of them
- Most of them
- Some of them

13. **Have you ever read/skim through all the summaries of such services?** *
Mark only one oval.

- Yes *Skip to question 14.*
- No *Skip to question 15.*

14. **How often do you read/skim through such summaries?** *
Mark only one oval.

- ◯ Sometimes
- ◯ Always

conclusion

15. **What is the ideal length of time (in minutes) you are willing to spend in reading T&Cs and PPs?** *

_____

16. **In general, what do you think of the level of difficulty to understand the clauses of T&Cs and PPs?** *
Mark only one oval.

	1	2	3	4	5	
Very Easy	◯	◯	◯	◯	◯	Very Difficult

17. **If you do not read/skim through T&Cs and PPs, what would you say are the reasons behind it? (Select all that apply)**
Check all that apply.

- ☐ I trust the service I sign up for
- ☐ My friends/family use the service
- ☐ Don't have so much time
- ☐ Can't understand the legal jargon
- ☐ Too long to read
- ☐ Other: _____

APPENDIX B : Final Questionnaire

Demographics

Please answer the following questions to help our analysis.

* Required

1. **What is your age?** *
   Mark only one oval.

   - 15 - 19
   - 20 - 24
   - 24 - 29
   - 30 - 34
   - 35 - 39
   - 40 - 44
   - 45 - 49
   - 50 - 54
   - 55 - 59

2. **Which gender do you identify most with?** *
   Mark only one oval.

   - Male
   - Female
   - Prefer not to say

3. **What is your highest qualification?** *
   Mark only one oval.

   - High school
   - Undergraduate
   - Master's
   - PhD
   - Other: _____

4. **What is your major?** *

   _____

5. **Is English your native language?** *
Mark only one oval.

- Yes
- No
- Other: _____

6. **How would you rate your reading and writing skills in English?** *
Mark only one oval.

	1	2	3	4	5	
Very bad						Very good

Acceptability Questionnaire

7. **How would you rate your confidence about your understanding of the clauses of Terebi?** *
Mark only one oval.

	1	2	3	4	5	
Very low						Very high

8. **How would you rate the readability of the clauses of Terebi?** *
Mark only one oval.

	1	2	3	4	5	
Very easy						Very difficult

9. **Would you read Terms & Conditions and Privacy Policies if they are presented to you in this format when signing up for any website?** *
Mark only one oval.

- Yes, always
- No, never
- Yes, sometimes
- Other: _____

Comprehension Questionnaire

Thank you for going through the terms and conditions of Terebi! Based on your understanding from what you read and saw, please answer the following questions. Select "I don't know" if you are not sure.

10. **ID (Please do not edit)**

_____

11. **Can Terebi suddenly discontinue any aspect of their service?** *
Mark only one oval.

- Yes
- No
- I don't know

12. **Will Terebi deny access to your account if you are held responsible for copyright infringements once?** *
Mark only one oval.

- Yes
- No
- I don't know

13. **Does Terebi share your personal information with others?** *
Mark only one oval.

- Yes
- No
- I don't know

14. **Does Terebi collect your location information?** *
Mark only one oval.

- Yes
- No
- I don't know

15. **Does Terebi retain content that you delete?** *
Mark only one oval.

- Yes
- No
- I don't know

16. **Will Terebi always notify you when changes to their terms and conditions are made?** *
Mark only one oval.

- Yes
- No
- I don't know

17. **Is Terebi liable for the content or practices of third party websites found through links from visiting Terebi?** *

Mark only one oval.

- Yes
- No
- I don't know

18. **Are you liable for losses of Terebi or other companies due to unauthorized use?** *

Mark only one oval.

- Yes
- No
- I don't know