The Post Anachronism: The Temporal Dimension of Facebook Privacy

Lujo Bauer†, Lorrie Faith Cranor*, Saranga Komanduri†, Michelle L. Mazurek*, Michael K. Reiter†, Manya Sleeper*, Blase Ur*

*Carnegie Mellon University {lbauer, lorrrie, sarangak, mmazurek, msleeper, bur}@cmu.edu
† University of North Carolina reiter@cs.unc.edu

ABSTRACT

This paper reports on two studies that investigate empirically how privacy preferences about the audience and emphasis of Facebook posts change over time. In a 63-participant longitudinal study, participants gave their audience and emphasis preferences for up to ten of their Facebook posts in the week they were posted, again one week later, and again one month later. In a 234-participant retrospective study, participants expressed their preferences about posts made in the past week, as well as one year prior. We found that participants did not want content to fade away wholesale with age; the audience participants wanted to be able to access posts remained relatively constant over time. However, participants did want a handful of posts to become more private over time, as well as others to become more visible. Participants’ predictions about how their preferences would change correlated poorly with their actual changes in preferences over time, casting doubt on ideas for setting an expiration date for content. Although older posts were seen as less relevant and had often been forgotten, participants found value in these posts for reminiscence. Surprisingly, we observed few concerns about privacy or self-presentation for older posts. We discuss our findings’ implications for retrospective privacy mechanisms.

Keywords
Privacy; Users; Access control; Temporality; Time; Facebook; SNS

Categories and Subject Descriptors
H.1.2 [User/Machine Systems]: Human factors

1. INTRODUCTION

Social networking sites like Facebook have become an important part of many users’ lives. Considerable research effort has been directed toward understanding users’ complex privacy preferences on these sites [1, 13], how users imagine their audience [5, 15], and many other aspects of privacy in social media.

However, beyond a recent small-scale, qualitative study [29] and a recent retrospective survey [3], little is known about how privacy preferences change over time. As each day passes, Facebook posts continue to accumulate. Users’ privacy preferences themselves may change over time, rendering preferences about older content obsolete. Content from months ago may no longer be relevant, or may even have become embarrassing. Users may not even remember old posts and may wish they would just disappear. At the same time, older content might help a user reminisce about the past.

In this paper, we examine temporal changes in users’ privacy preferences for their Facebook posts. Our primary investigation was a 63-participant longitudinal study in which participants completed surveys about their privacy preferences for the same set of posts at three different times: the week they made the post; one week later; and one month later. We implemented these surveys as a Facebook app, allowing us to select posts randomly and embed them in the survey. In a final survey, participants explained the importance of and rationale behind the changes we observed.

As a primary aspect of privacy on social networking sites involves bounding a post’s audience, we asked participants to specify who they wanted to be able to see a post at each time. Privacy can also involve making content more cumbersome to find, so we also asked questions about anyone the participant particularly hoped would see a post, as well as whether they wanted to emphasize a post visually or notify others of the post’s existence.

Informed by the limited past work on privacy temporality [3, 29], we tested the hypothesis that the audience participants wanted to be able to access a post would grow smaller over time. We also tested the hypothesis that participants would want similarly decreasing audiences for posts’ emphasis. We expected decreases would be driven by concerns about privacy, self-presentation, and relevance. We also investigated whether participants could, soon after posting, accurately predict how their preferences would change over time.

Instead, we found that the audience participants wanted to be able to access most of their Facebook posts remained relatively consistent over time. Participants did not want posts to fade away wholesale with age, and we observed about as many increases in audience breadth as decreases. Changes in preference that participants felt important were not captured by their actual privacy settings or actions, suggesting the need for more nuanced retrospective privacy mechanisms. While other researchers have proposed letting users give posts an expiration date at the time of posting [3], we found participants’ predictions about future changes to have surprisingly little association with their actual changes in preferences.

To compare our results more directly with the aforementioned retrospective survey [3], we also conducted a 234-participant retrospective study. We asked participants the questions from our longi-
tudinal study about posts they had made in the prior week, as well as in the same calendar week one year prior. We did not find the post’s age to have a significant impact on the audience participants chose. Although most participants had forgotten about the year-old posts, they often did not want these posts to disappear. Surprisingly, participants expressed few concerns about privacy or self-presentation for year-old posts. Older posts began to have greater value for the Facebook user himself or herself, mirroring qualitative results from Zhao et al. [29] and motivating our discussion about the need for reflective and retrospective privacy interfaces.

2. RELATED WORK

While the literature investigating privacy on social media is voluminous, only two recent studies have focused on privacy’s temporal dimension. In the first of these studies, Ayalon and Toch [3] conducted a 193-participant retrospective survey in which participants followed instructions to view Facebook Timeline posts from different eras. Participants then answered Likert-scale questions about their desire to keep the post on their timeline, delete the post, or change it. The authors found that the age of the post had a statistically significant, yet small, negative correlation with participants’ willingness to share the post on their timeline.

In contrast to Ayalon and Toch’s focus on Facebook’s existing mechanisms for sharing or deleting content, we investigated privacy preferences more abstractly. In particular, we focused on participants’ desired audience for a post, as well as ways to achieve privacy implicitly by making a particular post harder for others to find. Furthermore, our longitudinal study, the focus of this paper, investigated how privacy preferences for a single set of posts changed over time. This design allowed us to compare participants’ predictions with actual changes, as well as to ask participants to explain any changes we observed. To provide a more direct comparison to Ayalon and Toch’s work, we also conducted a retrospective study using our more nuanced questions about the audience for a post. Furthermore, we used the Facebook API to conduct our studies, enabling us to randomly select the posts we studied (avoiding participants’ biases in selecting a post and Facebook’s existing mechanisms for choosing which older posts to display), as well as to collect posts’ characteristics to correlate with behaviors. We also avoided drawing attention to the fact that the posts were old.

The second closely related study is Zhao et al.’s 13-participant diary study [29] investigating uses for older Facebook posts. They proposed that the Facebook Timeline consists of three functional regions. The first region is for the performance of new content, building on Goffman’s classic theatrical framework for the presentation of self [8]. The second region is an exhibition in which content from both the past and the present is displayed as in a museum, building on Hogan’s modernization of Goffman’s work [11]. The authors proposed a third area, the personal region, which includes content that the user wants to keep for himself or herself, rather than display to others. While Zhao et al. leveraged qualitative methods and communication theory to provide a rich portrait of these regions, they did not explore users’ privacy preferences on a per-post basis. Our work grounds their theory in a larger, empirical investigation of the incidence of privacy-preference changes over time, the factors that are correlated with changing preferences, and how participants’ predictions of changes align with actual changes.

Researchers have examined many other aspects of privacy on social networking sites. Past studies have focused on how users regulate boundaries among different social groups [6]. Tufekci highlighted the problems of boundary regulation, finding minimal relationship between privacy concern and information disclosure [24]. In a seven-year longitudinal study, Stutzman et al. found Facebook users shared less with the public as time went on, yet more with their Facebook friends [23]. Researchers have also examined content that users regret posting on social networking sites [21, 26], as well as methods for automatically identifying Twitter posts about sensitive topics [17].

Self-censorship and the deletion of past content have been identified as privacy mechanisms for both current and past disclosures. Young and Quan-Haase found that users of social networking sites self-censor, use alternative communication channels, and change privacy settings [28]. Sleeper et al. explored self-censorship on social networking sites in greater detail [20], finding that users of social networking sites often choose not to post certain content due to the difficulty of targeting a particular audience.

Prior studies have found that users do curate past content to some degree. For instance, a Pew survey of 802 teens found that 59% of respondents reported deleting or editing past content [16]. Al-muhimedi et al. captured 1.6 million Twitter posts that were later deleted [2]. They identified spam, typos, and sensitive topics among those deleted tweets. Johnson et al. examined Facebook users’ strategies for dealing with privacy concerns, finding that 37% of participants were concerned about sharing items with friends. Their participants used a variety of strategies to deal with privacy concerns, including untagging themselves and deleting content [13]. Wisniewski et al. identified informal privacy “coping mechanisms,” including self-censorship and the use of multiple accounts [27].

Privacy for online social networking sites is deeply connected to the audience that can access a post, as well as who actually views a post. Many researchers have investigated the imagined audience—the audience users believe they have [15]. Bernstein et al. quantified the imagined audience on Facebook. In particular, they measured who among a user’s friends actually saw a post, comparing this real data with study participants’ expectations. Although their participants wished they had a larger audience, they had consistently underestimated their actual audience by a factor of four [5].

The idea of data that disappears over time has also been examined in other domains. For instance, Geambasu et al. created Vanish, which uses cryptography to make data unreadable after a specified time [7]. The Snapchat mobile app allows users to share photos that “disappear” from the recipient after a few seconds [22]. However, neither method completely prevents circumvention.

3. LONGITUDINAL METHODOLOGY

In our longitudinal study, we investigated how participants’ privacy preferences changed in the month after they made a post on Facebook. We examined a single set of Facebook posts for each participant throughout the study, asking parallel survey questions about these posts at three separate times: the week the post was made (initial survey), one week later (week-later survey), and one month later (month-later survey). We implemented our study using a Facebook app, enabling us to select randomly from among participants’ posts, retrieve rich metadata about posts, and embed posts in our surveys absent the context of surrounding posts.

“Posts” comprise all links, photos, and status updates participants posted to Facebook, but exclude comments they made on content posted by others. If a participant had made fewer than 10 posts in the prior week, we selected them all; otherwise we selected 10 randomly, ensuring when possible that at least two links, two photos, and two status updates were included. These 2–10 posts became the set we asked about in each survey.

3.1 Recruitment

We recruited participants from the classified-ad websites Craigslist and Backpage. For Craigslist, we recruited from Pittsburgh, PA
and Raleigh-Durham, NC, where our universities are located, as well as eight other locations randomly selected from the 25 most active Craigslist posts (Sacramento, Portland, Washington DC, Boston, Orlando, Phoenix, Austin, and Seattle). On Backpage, we recruited throughout the United States. Recruitment took place in April 2013.

We screened for participants 18 or older who live in the United States and primarily speak English with their friends and on Facebook. We required (verified by our Facebook app) that participants had posted on Facebook at least twice a week on average over the prior month, as well as at least twice in the last week. Our studies were approved by our Institutional Review Board (IRB).

3.2 Study procedure

After accepting the consent form, participants installed our Facebook application, which collected the links, photos, and status updates the user had posted. Participants first completed a demographic survey that included general questions about their typical Facebook use, including posting and deletion behaviors.

The remainder of the study consisted of additional surveys. Following the demographic survey, participants completed an initial survey about 2–10 Facebook posts from the prior week. As time passed, they completed week-later and month-later follow-up surveys about the same posts. They then completed a final survey investigating the reasons for any changes in preferences we observed. We have included these surveys as Appendix A. Participants were paid $30 for completing the study: $5 each for the initial and two follow-up surveys, as well as $15 for the final survey.

3.2.1 Initial survey

Following the demographic survey, participants answered questions about 2–10 posts from the prior week. Throughout this paper, we refer to these questions as the initial survey. Each post was covered separately, in random order. To investigate the extent to which participants wanted to disclose each post, we asked about the total audience they wanted “to be able to see” each post. Participants chose one of eight categories: “everyone / public,” “friends of friends,” “all of your Facebook friends,” “only some of your Facebook friends,” “only one of your Facebook friends,” “only you,” “no one; it should disappear from Facebook,” and “other.”

As individuals can also achieve privacy by obscuring content, we investigated ways participants did or did not want to advertise the existence of a post. In particular, we asked about the targeted audience they “particularly hoped would look at” the post, mechanisms for adding emphasis to a post by highlighting the post for all or some of their friends, and hypothetical options for sending notifications to all or some of their friends about the post. To understand whether participants might be able to set an expiration date [3] that would match their preferences at a later point, we also asked participants to predict how they would want the noticeability of the post to change one week, one month, and one year later.

3.2.2 Week-later and month-later surveys

Each participant received an automated email both one week and one month after the initial survey to return for a follow-up survey. He or she received reminder emails every 24 hours until either the follow-up was completed or 72 hours had passed, at which point the participant was disqualified. Each follow-up survey asked about the total audience, targeted audience, and emphasis for the same posts as in the initial survey. These questions were parallel to those asked in the initial survey except for soliciting preferences “going forward,” rather than “at the time you made this post.” During each follow-up, the app also checked whether any posts collected during the initial survey had been deleted from Facebook.

3.2.3 Final survey

To understand changes in preferences we observed, we emailed participants to complete a final survey after they completed the month-later survey. For each post in which we observed any changes in total audience or targeted audience, we asked both multiple-choice and free-response questions about the importance of and reasons for the change. In addition, we asked multiple-choice and free-response questions about any posts that had been deleted.

3.3 Data analysis

We performed a number of statistical analyses investigating if participants’ preferences changed over time, as well as which characteristics of the post and the participant were correlated with any changes. As our data included multiple posts from each participant, observations were not independent. We therefore used the standard technique of including random effects in our models, effectively grouping each post with the participant who made the post.

Except where noted, we fitted the data to a cumulative-link (logit) mixed model, notated as CLMM throughout the paper. This type of model is appropriate for ordinal dependent variables [10], which was the case for most of our data. When our dependent variable was not ordinal, we created generalized linear mixed models, notated as GLMM. All models included four covariates about the post: the type of post (link, photo, status), the number of likes it had received, the number of comments it had received, and whether or not others had been tagged in the post. We also included three covariates about the participant: her age, her gender, and the number of Facebook friends she had. When we performed similar tests, we corrected p values using the conservative Bonferroni Correction.

We analyzed the accuracy of participants’ predictions about how their preferences for the noticeability of content would change over time using asymptotic linear-by-linear association tests. These tests compare the correlation between ordinal variables, which in our case were the directions of the predictions and the actual changes.

3.4 Limitations

Our methods have a number of limitations. First, even though privacy temporality deserves study in a number of domains, we restrict our investigation to Facebook posts in order to run a controlled study. We used a convenience sample that is not necessarily representative of all Facebook users, or any other large population. For instance, the proportion of female users in our study is higher than on Facebook overall, and our participants use Facebook more frequently than the general population [9]. We do not attempt to generalize the incidence of behaviors we observed to any larger population. Instead, our analyses are formative, seeking to understand the types of changes that occur over time in users’ privacy preferences, alongside characteristics associated with these changes and our participants’ rationale for changes.

We used the Facebook API to retrieve participants’ Facebook posts, enabling us to analyze characteristics about the posts, as well as to display them to the user during the surveys. This method excluded potential participants who chose not to use our Facebook app, and these users may be more privacy sensitive overall. Although we captured many posts that users made, we were unable to capture posts that participants made and then quickly deleted.

Using the Facebook API removed a number of biases from which alternative approaches would have suffered. Our app randomly selected from among a participant’s posts. Had we asked participants to locate their own older content as part of the study, they could have picked any post that caught their attention, or at least have been influenced by seeing other old posts.
4. LONGITUDINAL RESULTS

In this section, we report the results of the longitudinal study. We describe the participant pool and participants’ preferences for both each post’s audience and methods of advertising a post (highlighting and notification). We examine how these preferences changed over time, particularly in relation to participants’ predictions about these changes. We also discuss participants’ qualitative responses about how they use older content, as well as our observations of content being deleted over the course of the study.

4.1 Participants

Of 91 participants who completed the initial survey, 75 also completed the week-later survey. Of these 75 participants, 65 completed the month-later survey and 63 completed the final survey. In the remainder of this paper, we only analyze data from the 63 participants who completed all parts of the study.

The sample was not gender balanced; 27% of participants were male, 72% were female, and 1% declined to answer. Participants ranged in age from from 18 to 52, with a median age of 29. Among participants, 24% were students, 19% were unemployed, and the rest were employed in a variety of occupations, including education, arts, retail, and technology. Participants lived in 19 different U.S. states. Although we advertised on 10 of the 25 largest Craigslist sites, as well as all Backpage regions, 44% of our participants hailed from Pittsburgh, PA, where the university that advertised the study is located. This high concentration may indicate that people who recognized the university name in the study advertisement were more comfortable participating.

Our participants reported frequent Facebook use, with 89% saying they log onto Facebook multiple times a day, 9% reporting they log on daily, and the remaining 1% saying they log on a few times a week. Data collected by our Facebook app supported participants’ engagement with Facebook. According to the Facebook API, our participants had between 24 and 1,695 Facebook friends, with a median of 339 and mean of 466.2 ($\sigma = 382.4$). Participants were also relatively long-term Facebook users: 88% had used Facebook for more than three years, and the rest had used it for between one and three years. To log onto Facebook, 93% of participants reported using a desktop or laptop computer, 87% using a mobile device, and 27% using a tablet.

Our 63 participants answered questions about 462 different posts: 124 photos, 156 status updates, and 182 links. Each post had received between 0 and 42 “likes,” with a median of 1 and mean of 3.1 ($\sigma = 5.6$). The number of comments per post ranged from 0 to 20, with a median of zero and mean of 1.3 ($\sigma = 2.6$).

4.2 Access to a post

We asked participants about the total audience they wanted to be able to access a post at three different time periods. In the initial survey, we asked, “At the time you made this post, who did you want to be able to see it on Facebook?” Both one week later and one month later, we asked, “Going forward, who would you like to be able to see this post on Facebook?” Participants chose from among eight categories: “everyone / public,” “friends of friends,” “all of your Facebook friends,” “only some of your Facebook friends,” “only one of your Facebook friends,” “only you,” “no one; it should be deleted,” and “other.”

Contrary to our expectation, participants’ overall preferences remained relatively consistent over time (Figure 1). During all three time periods, participants said they wanted between 30% and 31% of posts to be public, while they wanted all of their Facebook friends to have access to between 49% and 53% of posts. Participants initially did not want anyone to see 3% of posts, which increased only to 4% after one week and 6% after one month.

As described in Section 3.3, we created a cumulative-link mixed model of the total audience chosen for a post. Participants’ choices of audience one week ($p = .135$) and one month ($p = .996$) later were not significantly different from the initial survey. Overall, male participants ($p = .001$) and participants with more Facebook friends ($p = .014$) picked broader audience categories. Participants also picked broader audience categories for links ($p = .011$) than status updates. Contrary to our expectations, the number of comments or likes a post had received, whether people were tagged in a post, and the participant’s age were not correlated with the audience category participants chose.

4.2.1 Changes in access

We also examined whether the size of the total audience for each post increased, stayed the same, or decreased over time. The majority of our participants had at least one post for which we observed a chance in audience. For example, from the initial survey to the survey one month later, 50 of the 63 participants (79%) chose a different audience for at least one post. Figure 2 illustrates audience changes on a per-post and per-participant basis, showing whether each participant had no audience changes for any post, only increases in audience breadth, only decreases, or both. Participants chose the same audience in the initial survey and one week later for 69% of posts. For 59% of posts, the participant chose the same audience across all three surveys.

Contrary to our expectation that privacy concerns would lead to audience preferences decreasing over time, audience changes were roughly split between audiences increasing and decreasing in size. For instance, one month later, participants chose the same audience...
as initially for 68% of posts, a smaller audience for 16% of posts, and a larger audience for the remaining 16%.

We found that participants did not feel strongly about many of the changes we observed. During the final survey, for each post for which a participant did not choose the same audience for all three longitudinal surveys, we said “we noticed your preferences changed” and showed the audience he or she had chosen at each time. To distinguish meaningful from ephemeral changes, we asked “How much does this change in preference matter to you?”

For 41% of cases where the audience shrank and 22% of cases where it grew, participants indicated that this change mattered a little, somewhat, or very much (Table 1). Although Facebook users can currently effect changes in audience for a post using privacy settings, we found this practice to be uncommon. In our final survey, participants reported changing the privacy settings for only 5% of posts for which the desired change in audience mattered a little, somewhat, or very much. As a result, for most of the posts where participants’ preferences changed, reality did not match preference, suggesting the need for retrospective privacy mechanisms.

We hypothesized that activity around the post (number of comments, number of likes), the type of post (link, photo, or status update), or demographic characteristics of the participant might be correlated with changes in audience. However, again using a cumulative-link mixed model, we did not find significant correlations between any of these factors and whether the participant chose a broader audience category, the same audience category, or a narrower audience category as time passed (all p ≥ .183)

4.2.2 Why the audience changed

We presented participants who attributed at least “a little” importance to a particular post’s change in audience with ten factors we hypothesized might have driven the change, shown in Appendix A. They rated their agreement or disagreement with each factor on a five-point Likert scale. For posts where the audience decreased in size, participants’ changes were most commonly driven by the occurrence of some specific event (participants responded “agree” or “strongly agree” for 43% of posts), feedback from others (40%), or the belief that a post was no longer relevant (37%).

Surprisingly few participants attributed changes in audience to privacy concerns, changes in relationship with people depicted in or relevant to a post, or beliefs that a “post did not depict [participants] in the manner [they] wanted to appear” – participants responded “agree” or “strongly agree” for at most 10% of posts for each reason. Increases in audience size were most commonly associated with a desire for more people to see the post (67%).

We observed similar trends in free responses. Many participants who wanted smaller audiences felt the content had lost relevance. For example, one participant said a post “appealed to the nerd in me, but a month later was no longer fresh.” Others tied relevance to an event, like one participant who said a post lost relevance once “tax day came and went.” Most participants who indicated an increase in audience size simply wanted more people to see the post. For example, one participant said, “At first I just posted so I could remember but then I want others to see.” Another participant explained, “I want my boyfriend’s friends to see us happy together.”

4.3 Advertising a post

While an individual can achieve privacy by restricting access to information, he or she can also achieve a degree of privacy by making information hard to find. This technique can be particularly appropriate when an individual neither wishes to bombard others with content nor to completely restrict access, but would like to make a post available to those who seek it out. Whereas Section 4.2 investigated participants’ preferences about restricting access, this section describes preferences about how conspicuously content is advertised. We first examine participants’ preferences for targeted audience, or who they particularly hoped would see a post. We then examine how participants wanted to visually emphasize, or highlight, a post. Finally, we investigate push vs. pull advertising by asking participants about actively notifying others about a post.

4.3.1 Changes in targeted audience

To understand whom participants particularly wanted to target with a post, we asked participants in each survey, “Is there anyone you particularly hope will look at [this post]?” Participants chose from the same categories as when choosing the total audience. We termed participants’ responses their targeted audience.

Participants’ preferences for targeted audiences were broad and were frequently the same as the total audiences. Overall (Figure 3), participants often hoped to target all of their Facebook friends or “everyone/public.” This result suggests that participants hoped to reach a wide audience, echoing results from Bernstein et al. [5].

As with the total audience, the distribution of participants’ preferences about the targeted audience was not significantly different one week (p = .157) or one month (p = .769) after the ini-
Most participants who wanted any extra emphasis wanted both types. Over time, the proportion of posts for which emphasis was desired dropped only slightly; even after one month, participants wanted 43% of posts to be emphasized in at least one way, suggesting that participants want to draw attention even to aged posts.

Relative to the initial survey, our CLMM showed that participants wanted fewer posts to be visually highlighted both one week later (p = .012) and one month later (p < .001) later. Compared to status updates, participants wanted links to be visually highlighted at a significantly higher rate (p = .023). Participants also appeared to want photos to be visually highlighted at a higher rate than status updates, yet this difference was not quite statistically significant (p = .056). Participants also wanted fewer posts to appear at the top of the timeline both one week (p = .016) and one month (p < .001) later. Participants wanted links (p = .004) and photos (p = .004) to be emphasized at a higher rate than status updates.

4.3.3 Notification

Managing audience for Facebook posts extends beyond who can see the content and how obvious it is. Privacy preferences also include how and whether audience members are notified that the content is available. In the initial survey, we asked participants whether they wanted all, some, or none of the total audience to receive each of five different types of notification that the post was available. Participants wanted 78% of posts to appear in all or some of their Facebook friends’ newsfeeds. Furthermore, participants wanted 63% of posts to appear at the top of all or some friends’ newsfeeds, and 41% of posts visually highlighted in the newsfeed. Participants wished to send Facebook notifications for 37% of posts, and either an email or text message for 30% of posts. For each notification method, participants wanted to notify friends tagged in the post for an additional 8%–20% of posts.

Table 2: How much participants said the change they indicated in targeted audience matters, on a per-post basis.

<table>
<thead>
<tr>
<th>Audience intended</th>
<th>Audience grew</th>
</tr>
</thead>
<tbody>
<tr>
<td>No change</td>
<td>22%</td>
</tr>
<tr>
<td>Not at all</td>
<td>45%</td>
</tr>
<tr>
<td>A little</td>
<td>13%</td>
</tr>
<tr>
<td>Somewhat</td>
<td>17%</td>
</tr>
<tr>
<td>Very much</td>
<td>3%</td>
</tr>
<tr>
<td></td>
<td>32%</td>
</tr>
<tr>
<td></td>
<td>44%</td>
</tr>
<tr>
<td></td>
<td>13%</td>
</tr>
<tr>
<td></td>
<td>9%</td>
</tr>
<tr>
<td></td>
<td>2%</td>
</tr>
</tbody>
</table>
4.4 Predicting changes

We also investigated how well participants could predict changes that would occur over the month. Overall, participants were quite poor at making predictions, casting doubt on the feasibility of letting users set an “expiration date” when they make a post.

The first time we asked participants about a post, we asked how they expected their desired noticeability of that post to change over time. Figure 6 illustrates participants’ predictions. Compared to their predictions about noticeability one week later, participants predicted they would want posts to be less noticeable both one month ($p < .001$) and one year ($p < .001$) later in our CLMM. They predicted they would want status updates to decrease in noticeability more than their photos ($p < .001$), while they predicted that they would want posts in which they had tagged other users to decrease in noticeability more than posts without tags ($p = .047$).

We found that participants’ predictions of how they would want the noticeability of posts to change over a month were not well aligned with their actual decisions one month later. We performed asymptotic linear-by-linear association tests to investigate the correlation between participants’ predictions and the actual changes (increase, same, decrease) in the size of the total audience, the size of the targeted audience, and the amount of emphasis a participant wanted to place on a post. Participants’ predicted changes in noticeability were not significantly correlated with the actual changes in total audience ($p = .705$), targeted audience ($p = .632$), or emphasis ($p = .056$), even without correcting for multiple testing.

Figure 7 illustrates the mismatch between participants’ predictions and their decisions one month later. For example, among the 285 posts participants predicted they would want to be “a little less noticeable,” “much less noticeable,” or “not accessible at all” one month later, participants actually chose a smaller total audience for only 54 posts (19%) one month later, whereas they chose a larger total audience for 42 posts (15%). Predictions of increases and decreases in noticeability were similarly disjoint from actual changes in total audience, targeted audience, and emphasis. Most of the posts for which the size of the total and targeted audiences actually decreased were posts users had expected they would want to become less noticeable, but the analogous relationship did not hold for posts whose intended audience increased in size.

To evaluate whether extreme predictions would be more accurate, we examined the 24 posts participants predicted they would want to become “not accessible at all” one month later. Participants actually selected “no one” as the audience one month later for only three of these posts (13%). In contrast, for 17 posts (71%), the participant selected a targeted audience of everyone/public, friends of friends, or all Facebook friends, contradicting the prediction.

4.5 Deletion

We also investigated whether participants deleted any of the posts we asked about in the study. During both the week-later and month-later surveys, our Facebook app verified that the content we asked about was still available. Even if the post had been deleted, participants saw a cached copy of the post while completing the survey.

We observed nine of the 63 participants (14%) deleting content during our study. One week after the initial survey, only four of the 462 posts (1%) had been deleted. After one month, 31 additional posts (7%) had been deleted. Notably, 31 of these 35 deleted posts were status updates (89%), while the other four were photos.

In our final survey, we showed participants any posts they had deleted over the course of the study and asked how important it was to them that the post disappear from Facebook. Most participants did not feel strongly about the posts being deleted; participants said that deleting the post mattered “very much” for only two of the 35 deleted posts (6%), “somewhat” for only one post (3%), “a little” for only one post (3%), and not at all for the remaining 89%. When asked why they deleted these posts, participants wrote phrases like “no one seemed to care” and “[it was] obsolete.” A participant who deleted a photo “already had copies.” Another participant wrote, “I re read it and it doesn’t make sense and it just sounds rude. Idk.” A final participant confessed, “I was just rambling.”

4.6 Uses for historical posts

We used participants’ free-response answers to understand why they seemed to find old posts useful. During the demographic survey, we asked participants if they had “ever used Facebook to look at a post someone made more than a few months ago.” Fifty participants (79%) had done so. Several participants looked at older posts to learn more about people. They would look at older posts of someone they met as “a background check, to see what he/she is all about.” Other participants revisited posts when they became relevant again due to new comments. A number of participants used old posts “to reminisce” or “to look back on memories.”

During our final survey, we again asked participants how they felt about content they had posted “a long time ago.” Participants described using old content to keep a record of themselves, referring to it as a “scrapbook.” One participant explained, “Sometimes it seems a bit dorky, like anything you write within a certain window of time in the past. But I hope these details about my life and thoughts remain available in the future, for posterity.” Other participants revisited deleted older content. One participant explained, “I want it gone. Especially with the feature on your profile that allows you to go to a specific year. I don’t want new friends, future employers, and romantic interests to see me in 2006.”
5. RETROSPECTIVE METHODOLOGY

We used a retrospective study to examine even older posts than in our longitudinal study. In a single survey, we asked participants about the posts they had made in the week preceding the survey in May 2013, as well as during the same week one year prior.

To increase sample size, we recruited participants for the retrospective study from Amazon’s Mechanical Turk (MTurk) crowdsourcing service. We again screened for English-speaking participants 18 or older. To align with average rates on MTurk, we paid participants $3. We only accepted U.S. workers with an approval rating over 95%. We manually examined participants’ free-response answers. All participants appeared to take the study seriously, possibly because our app’s installation and automatic checks were onerous. As we were recruiting anonymous workers, we did not download the content of the posts, nor names or other PII.

We again used a Facebook app to conduct the study. Participants were required to have made at least one post in the past week, as well as at least one post in the same week one year earlier, verified by our app. If a participant had made more than five posts in a week, we randomly selected five posts. We asked participants to express their preferences about audience and emphasis for each post “going forward.” We also asked them to rate their agreement or disagreement on a five-point Likert scale with statements covering motivations for a post to remain on, or be removed from, Facebook.

5.1 Limitations

Our retrospective study has many of the limitations of our longitudinal study, as well as a few additional ones. Participants may have already deleted posts from a year ago, precluding us from asking about them. Deleted posts might be more privacy sensitive overall, but they constitute only a fraction of posts [2], and users’ privacy preferences for all posts are interesting. Recruiting participants from MTurk introduces additional limitations. Although MTurk workers are generally younger and more educated than the general population [12], they can still provide high-quality data [4,14], and our app verified the age of their Facebook account.

Using the Facebook API again enabled us to randomly select and display posts from the past in isolation. However, for the retrospective study only, an additional limitation of the Facebook API itself biased our sample. In our experience, and as documented on developer forums,1 the Facebook API appears to limit the historical posts that can be returned to the few hundred most recent. Users who post more frequently therefore have less historical content available. As a result, we were unable to enroll the most frequent Facebook users (more than a few hundred posts in a year) in the retrospective study. Nevertheless, our participants were still active Facebook users.

6. RETROSPECTIVE RESULTS

Whereas our longitudinal study investigated privacy preferences for a given post at three different times, we were also interested in preferences for even older posts. In this section, we compare participants’ privacy preferences for recent and year-old content.

6.1 Participants

A total of 234 MTurk workers participated in our retrospective study. This sample was more gender-balanced than the longitudinal study: 55% of participants were female, 44% were male, and 1% declined to answer. Participants ranged from 18 to 60 years old, with a median age of 26. They hailed from 44 U.S. states. Among participants, 27% were students, 11% were unemployed, and the remainder held various jobs, from teachers to filmmakers.

Despite the different recruiting pool, the retrospective participants resembled the longitudinal participants in terms of Facebook experience. A total of 93% of participants reported using Facebook for more than three years, while 7% reported using Facebook for one to three years. While 71% of participants said they looked at Facebook multiple times a day, another 23% said they did so daily.

6.2 Comparison of current and year-old posts

As in our longitudinal study, we asked participants “Who do you want to be able to see [this post] on Facebook?” and “Is there anyone you particularly hope will look at [it]?” for each post they had made in the past week (termed 2013 posts), as well as in the corresponding week one year prior (termed 2012 posts). We collected 237 posts from 2012 and 392 posts from 2013.

As in our longitudinal study, we observed the distribution of participants’ preferences about privacy to be relatively similar over time, suggesting that whether a post is a few days old or a year old does not strongly impact these preferences. Figure 8 shows participants’ preferences for both the total audience and targeted audience for 2012 posts and 2013 posts.

We again created CLMMs of these data. Contrary to intuition about the privacy of old data, the year was not a significant factor (p = .606) in participants’ preferences about the total audience for posts. Participants were more likely to share links (p = .033) and photos (p = .007) than status updates with a broader audience. Furthermore, male participants wanted to share with a broader audience than female participants (p = .033).

We found similar results for the targeted audience. Had we excluded covariates from the CLMM, participants would have appeared to target 2013 posts to significantly larger audiences (p = .022) as nearly twice as many 2012 posts were targeted to “no one” (Figure 8). However, once we controlled for covariates, we no longer observed the year of the post to have a significant effect (p = .721). Instead, we observed a number of significant covariates. Compared to status updates, participants were more likely to target links (p < .001) and photos (p = .007) to a broader audience. Furthermore, older participants (p = .048) and male participants (p = .005) wanted to target posts to a larger audience than younger and female participants, respectively.

As in the longitudinal study, we asked participants to rate their agreement or disagreement on a five-point Likert scale with opinions and reactions we hypothesized might apply to Facebook posts from the past or the present (Figure 9). We constructed a Generalized Linear Mixed Model (GLMM) for each statement and performed Bonferroni Correction (labeled BC) to correct for multiple
Percent of posts with audience change for which the user agreed or strongly agreed.

<table>
<thead>
<tr>
<th>Reason</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>... primarily for my friends to look at it</td>
<td>14%</td>
<td>23%</td>
</tr>
<tr>
<td>... primarily to look at it myself</td>
<td>36%</td>
<td>16%</td>
</tr>
<tr>
<td>... both</td>
<td>50%</td>
<td>62%</td>
</tr>
</tbody>
</table>

Table 3: Participants’ per-post completions of the sentence, “I want this post to remain on Facebook...” Participants more commonly felt that year-old posts were valuable for themselves.

testing. Relative to 2013 posts, participants were more likely to have forgotten 2012 posts, have had relationship changes with individuals relevant to the post, find the post irrelevant, and find the post helpful in reminiscence (all \( p < .001 \), BC).

Several responses we expected to be popular for older posts were rarely chosen by participants. Self-presentation is often considered an important aspect of social network sharing [29], yet only 9% of 2012 posts were identified as no longer depicting their authors accurately. Similarly, participants reported having privacy concerns for only 9% of 2012 posts. While participants considered 47% of 2012 irrelevant, they felt 68% of 2012 posts helped them reminisce.

We also asked participants about the value of each post, as shown in Table 3. Participants more often wanted posts from the past week to remain on Facebook for friends to look at them, whereas they more often wanted posts from a year prior to remain on Facebook for themselves to look at. To confirm these observations, we created separate GLMMs representing “for friends” and “for myself.”

Participants wanted 2013 posts to remain for friends significantly more than 2012 posts (\( p < .001 \), BC), while they wanted 2012 posts to remain for themselves at a significantly higher rate (\( p = .024 \), BC). We observed significant covariates only “for friends.” Links were intended for friends at a greater rate than status updates (\( p < .001 \), BC), and the number of likes a post received was positively correlated with intending it for friends (\( p = .005 \), BC). Older participants (\( p = .023 \), BC) and male participants (\( p = .014 \), BC) were also more likely to intend posts for friends.

7. DISCUSSION AND CONCLUSIONS

We presented a longitudinal and a retrospective study examining how users’ privacy preferences and desired audience for Facebook posts change over time. While we found that participants did not want content to fade away wholesale with age, participants did want a handful of posts to become more private over time, as well as a handful of posts to become more visible or to be readvertised.

Notably, popular sites currently make this task difficult. Changing old posts’ privacy settings on Facebook involves burrowing through the timeline and navigating cumbersome interfaces, while Google+ does not permit any post’s audience to be changed after the fact beyond deleting and reposting the content.

The apparent mismatch between our participants wanting to limit the disclosure of only a few posts as time passed and the wide adoption of tools like Snapchat [22] to make data disappear can likely be explained by users’ maturing understanding of privacy on social networking sites. While the early literature on social networking sites found users to have little understanding of sites’ privacy implications [1], users are reaching an increasingly nuanced understanding and practice of privacy protection online [13, 23]. Users still make posts they later regret [26], yet they also do so offline [21], suggesting that online social networks are not solely to blame.

Given that most Facebook posts in our study did not appear to become more privacy sensitive as they aged, one might suggest that not doing anything to address old content is an acceptable solution. However, a need exists to address old content whose privacy sensitivity has changed, even though these posts represent only a fraction of the old posts on Facebook. For instance, in the longitudinal study, participants said they wanted 42 posts (9.1%) to be visible to either no one or only themselves after one month. For 9 of these posts, the four different participants who made the posts said these changes mattered to them. However, in only one case had the participant deleted the post, and the participants had not changed the privacy settings for any of the posts, leaving a gap between preference and reality. Similarly, in the retrospective study, participants wanted twice as many year-old posts as posts from the past week to be accessible by no one or only themselves (5.9% vs. 2.8% of posts). Although participants preferred that these posts not be accessible, they still were accessible.

The percentage of posts for which privacy preferences changed drastically was small, yet given the large number of Facebook posts an average user makes, even a small percentage should not be ignored. While we did find some correlations between characteristics of the post or user and changes in privacy preferences, these associations were not strong enough to automate changes.

Previously proposed interfaces for retrospective privacy do not appear to support the preferences we observed, suggesting the need for novel interfaces for adjusting privacy settings for content as it
For instance, Ayalon and Toch [3] proposed a mechanism with which users could set an automatic expiration date when they posted. Unfortunately, we found that participants were very poor at predicting their future changes in preference. Both Zhao et al. [29] and Ayalon and Toch [3] proposed an “archive” feature for restricting posts from all friends. However, we found that participants wanted to restrict only a small fraction of posts to themselves, raising questions about the usefulness of extensive archiving features.

Instead, a way forward might be to design interfaces that promote reflection about older content. While a user browses a social networking site, the site could highlight a post that user had made in the past, either in a sidebar or visually set off from recent content. This interface might also remind users of who can see their post, taking guidance from work on privacy “nudging” [25]. This interface could give the user choices about decreasing or increasing access to the post, as well as the way the post is advertised. Based on our results, users would likely restrict the audience for the handful of posts whose sensitivity, relevance, or other characteristics changed with the passage of time. However, we found that participants often wish to allow others to access even those posts whose primary purpose has become personal, rather than social. Users will likely increase the access to, or the visual emphasis of, additional posts. In these cases, a post from the past might have great value in promoting reminiscence [18, 19]. Users might want to choose particular posts to readvertise as if they were new, interspersing recent posts with pleasant “blasts from the past.”

8. ACKNOWLEDGMENTS

This research was supported in part by the Facebook graduate fellowship program, the ARCS Foundation, the Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program, and National Science Foundation grants CNS-0831407, DGE-0903659, and 0946825.

9. REFERENCES

APPENDIX

A. LONGITUDINAL STUDY MATERIALS

Initial survey:

At the time you made this post, who did you want to be able to see it on Facebook?
( ) Everyone / Public   ( ) Friends of friends   ( ) All of your Facebook friends
( ) Only some of your Facebook friends. (Who?): ________________________
( ) Only one of your Facebook friends. (Who?): ________________________
( ) Only you   ( ) No one; it should have disappeared from Facebook   ( ) Other: ________________________

At the time you made this post, was there anyone you particularly hoped would look at it?
( ) Everyone / Public   ( ) Friends of friends   ( ) All of your Facebook friends
( ) Only some of your Facebook friends. (Who?): ________________________
( ) Only one of your Facebook friends. (Who?): ________________________
( ) Only you   ( ) No one   ( ) Other: ________________________

At the time you made this post, how would you have felt if a person you didn't want to view it were able to see it?
( ) Not upset at all   ( ) A little upset   ( ) Upset   ( ) Very upset

At the time you made this post, how would you have felt if a person you particularly hoped would look at it were not able to see it?
( ) Not upset at all   ( ) A little upset   ( ) Upset   ( ) Very upset

You indicated that you wanted [PEOPLE] to be able to view this post. Of these people, how many would you have liked to see it in the following ways:

All of them   Some of them   Only people who are tagged in the post, if any   No one
...when they looked at your Facebook Timeline? ( )  ( )  ( )  ( )
...visually highlighted when they looked at your Facebook Timeline? ( )  ( )  ( )  ( )
...near the top when they looked at your Facebook Timeline? ( )  ( )  ( )  ( )

How noticeable, compared to now, do you expect you'll want this post to be to people who can see it on Facebook...

Much more noticeable   A little more noticeable   Equally as noticeable   A little less noticeable   Much less noticeable   Not accessible at all
...one week from now? ( )  ( )  ( )  ( )  ( )
...one month from now? ( )  ( )  ( )  ( )  ( )
...one year from now? ( )  ( )  ( )  ( )  ( )

You indicated that, at the time you made this post, you wanted [PEOPLE] to be able to view it. At that time:

All of those people   Some of those people tagged in the post   Only people   No one
Who would you have liked to receive an email or text message notification about this post? ( )  ( )  ( )  ( )
Who would you have liked to receive a notification about this post in Facebook's notification box? ( )  ( )  ( )  ( )
Who would you have liked to see this post in their newsfeed? ( )  ( )  ( )  ( )
Who would you have liked to see this post near the top of their newsfeed? ( )  ( )  ( )  ( )
Who would you have liked to see this post visually highlighted (for example, twice as large) in their newsfeed? ( )  ( )  ( )  ( )

Are there any other ways you would have liked to notify anyone when you posted the content? Please describe.

How would you feel if a person you didn't want to notify about this post received a notification?
( ) Neutral   ( ) A little upset   ( ) Upset   ( ) Very upset

How would you feel if a person you wanted to notify about this post didn't receive a notification?
( ) Neutral   ( ) A little upset   ( ) Upset   ( ) Very upset
One week and one month later:

[POST APPEARS]

Going forward, who would you like to be able to see this post on Facebook?
( ) Everyone / Public   ( ) Friends of friends   ( ) All of your Facebook friends
( ) Only some of your Facebook friends. (Who?): _____   ( ) Only one of your Facebook friends. (Who?): _____
( ) Only you   ( ) No one; it should disappear from Facebook   ( ) Other: ________

Going forward, is there anyone you particularly hope will look at this post?
( ) Everyone / Public   ( ) Friends of friends   ( ) All of your Facebook friends
( ) Only some of your Facebook friends. (Who?): _______   ( ) Only one of your Facebook friends. (Who?): _______
( ) Only you   ( ) No one   ( ) Other: ________

Going forward, how would you feel if a person you didn't want to view this post were able to see it?
( ) Not upset at all   ( ) A little upset   ( ) Upset   ( ) Very upset

Going forward, how would you feel if a person you particularly hoped would look at this post were not able to see it?
( ) Not upset at all   ( ) A little upset   ( ) Upset   ( ) Very upset

The following questions ask about how you would like your post to appear (some of the options you will be asked about currently exist, while others are hypothetical).

You indicated that you wanted [PEOPLE] to be able to view this post. Of these people, how many would you like to see it in the following ways:

<table>
<thead>
<tr>
<th>All of those people</th>
<th>Some of those people</th>
<th>Only people tagged in the post</th>
<th>No one</th>
</tr>
</thead>
<tbody>
<tr>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
</tbody>
</table>

...when they look at your Facebook Timeline?
...visually highlighted when they look at your Facebook Timeline?
...near the top when they look at your Facebook Timeline?

Final Survey: (We displayed analogous questions for each of: “who you wanted to be able to see the post,” “who you particularly hoped would look at the post,” and posts “you deleted.”)

[POST APPEARS]

We asked you who you wanted to be able to see the post shown above on Facebook at three different times. We noticed that your preferences changed.

Initially, you said you wanted the following people to see your post: [PEOPLE]

One week later, you said you wanted the following people to see your post: [PEOPLE]

Four weeks later, you said you wanted the following people to see your post: [PEOPLE]

How much does this change in preference matter to you?
( ) Very much   ( ) Somewhat   ( ) A little   ( ) Not at all   ( ) N/A (I didn't mean to indicate a change in preference)

Please describe why your preference for who you wanted to be able to see this post on Facebook changed.

Please indicate how much you agree or disagree that each of the following impacted your change in preferences.

<table>
<thead>
<tr>
<th>Strongly agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly disagree</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>This post did not depict me in the manner I wanted to appear on Facebook</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>This post was no longer relevant</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>I had forgotten that this post was still on Facebook</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>I was concerned about privacy for this post</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>No one was interested in the post's content</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>I received feedback from others about the post</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>I wanted more people to see the post</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>The sensitivity of the post changed</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>A specific event occurred</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>Something not included in this list impacted my preferences</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
<tr>
<td>My relationship with the people depicted in, tagged in, or particularly relevant to this post changed</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
<td>( )</td>
</tr>
</tbody>
</table>

(Optional) Did any other factors impact your change in preferences?

Did you use Facebook's privacy settings to change who could view this post?   ( ) Yes   ( ) No   ( ) I don't remember