Perceptions of Retrospective Edits, Changes, and Deletion on Social Media

Günce Su Yılmaz,[1] **Fiona Gasaway,**[1] **Blase Ur,**[1] **Mainack Mondal**[1,2]

[1] University of Chicago
[2] IIT Kharagpur
suyilmaz@uchicago.edu, fgasaway@uchicago.edu, blase@uchicago.edu, mainack@cse.iitkgp.ac.in

Abstract

Many social media sites permit users to delete, edit, anonymize, or otherwise modify past posts. These mechanisms enable users to protect their privacy, but also to essentially change the past. We investigate perceptions of the necessity and acceptability of these mechanisms. Drawing on boundary-regulation theories of privacy, we first identify how users who reshared or responded to a post could be impacted by its retrospective modification. These mechanisms can cause boundary turbulence by recontextualizing past content and limiting accountability. In contrast, not permitting modification can lessen privacy and perpetuate harms of regrettable content. To understand how users perceive these mechanisms, we conducted 15 semi-structured interviews. Participants deemed retrospective modification crucial for fixing past mistakes. Nonetheless, they worried about the potential for deception through selective changes or removal. Participants were aware retrospective modification impacts others, yet felt these impacts could be minimized through context-aware usage of markers and proactive notifications.

Introduction

Online social media sites like Facebook and Twitter have enabled users to share information for over a decade. A user's account can span many years and life phases, and these platforms have effectively become personal archives. Inevitably, users will want to modify some past content for reasons ranging from privacy (Zhou, Wang, and Chen 2016; Mondal et al. 2019) to fixing typos (Almuhimedi et al. 2013) to removing regrettable posts (Sleeper et al. 2013; Wang et al. 2011) to reflecting changes in life phase, such as changing jobs or transitioning identities (Bauer et al. 2013; Haimson et al. 2016). As a result, retrospective privacy management — changing the past through edits, deletion, or modification — is critical.

Nearly all social media platforms permit some form of retrospective content modification, such as manual deletion of a post, editing a post, or time-based automatic deletion. Platforms vary widely, though, in which *mechanisms* they support (Mondal et al. 2016). The ability to change the past raises questions about whether users perceive these mechanisms as socially acceptable and necessary, as well as how

perceptions differ based on the context or type of modification. These mechanisms have pros and cons. Retrospective edits that fix typos can improve clarity. Judicious deletions can address regrettable posts. However, removing abusive or hateful content may let the user evade accountability. Editing a post can change its meaning or fraudulently recontextualize a conversation. Facebook acknowledged this tension, tilting toward accountability, by removing the ability to edit photos retrospectively (Sanchez 2017). Our study is the first to systematically investigate perceptions of retrospective content modification on social media, as well as its resultant tensions.

Social media is a particularly complex domain for retrospective content modification since users other than the initial creator reshare posts, reply to them, like them, or are tagged in them. These other users may also be affected if a post is retrospectively edited, deleted, or modified. An individual who reshared or liked a benign post would be alarmed if it were edited to be hateful or controversial. If a post is later edited, deleted, or even anonymized, replies and other subsequent interactions could be recontextualized. While a post's creator might value the ability to delete replies, doing so can lead to an outcry from users whose replies were deleted.

Prior work has investigated why users delete content (Almuhimedi et al. 2013; Ramokapane, Rashid, and Such 2017), whether desired changes in post visibility over time can be predicted automatically (Bauer et al. 2013; Ayalon and Toch 2013, 2017; Mondal et al. 2019), and the causes of regret on social media (Zhou, Wang, and Chen 2016; Wang et al. 2011; Sleeper et al. 2013). This prior work, however, focuses on the creator of the post modifying content, not its impact on others. We instead examine how a platform's inclusion of particular mechanisms for retrospectively changing posts can impact both post creators and others who have interacted with the post. Another stream of work has examined how users manage multi-party privacy and content co-ownership, primarily in the context of photo tagging (Such and Criado 2018; Lampinen et al. 2011; Marshall and Shipman 2017; Besmer and Richter Lipford 2010; Wisniewski et al. 2016), but that work does not cover retrospective modification or the many subtle types of content co-ownership we examine.

Because understanding retrospective modification is complicated by the multitude of mechanisms available on different platforms and proposed in the literature (Mondal et al. 2016; Ayalon and Toch 2013), we first built a taxonomy of

retrospective mechanisms. We found that popular sites differ substantially from each other in the mechanisms they provide and what visual cues (*markers*) accompany changed content.

We drew on Communication Privacy Management (CPM) theory (Petronio 2002) and work on multi-party privacy (Such and Criado 2018; Lampinen et al. 2011; Wisniewski et al. 2016) to extend understandings of: (i) perceptions of the co-ownership of previously posted social content; and (ii) how retrospective changes contribute to boundary turbulence.

While our taxonomy covers stakeholders and mechanisms, it does not consider the acceptability of retrospective changes. We thus conducted a qualitative study of perceptions of these mechanisms centered on the following research questions:

- **RQ 1**: When do post creators deem different deletion, editing, and modification mechanisms useful, acceptable, and necessary for posts they themselves have made?
- **RQ 2**: Whom do participants believe to be affected by retrospective changes? Why?
- **RQ 3**: When content has been deleted or edited retrospectively, in which situations do participants think the platform should provide a marker or revision history?
- **RQ 4**: Do responses to RQ 1–3 change if considered from perspectives other than the post creator's?

We investigated these questions through semi-structured interviews of 15 long-term users of social media. Participants gathered posts both they and others had made in different eras. We based parts of the interview on these posts, representing both those the participant originally created and those with which they had only interacted (e.g., liked or commented on). Other parts of the interview elicited participants' attitudes more generally. In both cases, participants responded to questions about the acceptability of a post's creator retrospectively changing it using mechanisms from our taxonomy. We also asked whom the participant felt would be impacted by particular changes, as well as what types of markers (e.g., icons and edit histories) they felt should accompany those changes. Rather than studying specific platforms' implementations, we focused on perceptions about retrospective editing, deletion, and modification abstractly. For concreteness, we discussed examples deployed on current platforms.

We found that participants deemed the ability to retrospectively delete, edit, or modify content necessary for fixing past mistakes and remediating regrettable posts. Without such mechanisms, participants worried that post creators would lack both privacy and agency. However, they also identified these mechanisms' potential for causing boundary turbulence if the edit involved a large change or the loss of conversational context. Overall, participants noted how retrospective modification can impact other users, yet felt that these impacts could be minimized if the mechanisms provide revision histories, explicit markers, or proactive notifications. These indicators have limitations, however, in sometimes drawing excess attention to minor modifications. Thus, participants felt the markers should be displayed only in certain contexts based on the time between the post and the modification, the post's popularity, and the semantics of the change itself.

Related Work and Basis in Theory

Changing the past via retrospective mechanisms is common on social media. Users modify past posts for myriad reasons, including regretting a past statement (Wang et al. 2011; Sleeper et al. 2013), the passage of time (Liu, Kliman-Silver, and Mislove 2014), undergoing a life change (Child, Haridakis, and Petronio 2012), desiring to forget (Schönberger 2011), or desiring to update content (DeVito, Birnholtz, and Hancock 2017). Several empirical studies focused on mechanisms for modifying social media content through deletion or changing the audience. Researchers found more than 28% of six-year-old tweets had been removed (Mondal et al. 2016). Further, 17% of selectively deleted tweets were because of typos or a desire to rephrase the post (Almuhimedi et al. 2013). Users who delete tweets are more likely to be extroverted and neurotic (Bhattacharya and Ganguly 2016). Researchers have highlighted context-specific nuances; users like some content to remain visible for reminiscence, but other content to disappear (Zhao et al. 2013; Bauer et al. 2013). Some studies have focused on predicting changes (Volkova and Bell 2017) or changing user behavior through nudging (Wang et al. 2014; Wisniewski, Knijnenburg, and Richter Lipford 2017).

None of these previous studies investigated how retrospective changes impact users other than the post creator. These other users might find some mechanisms more acceptable than others in particular contexts. We turned to theory to begin building an understanding of how retrospective changes on social media might impact others. Boundary-regulation theories frame privacy as the selective control of access to the self or group by moving metaphorical boundaries. Altman proposed one of the first such theories for offline communication (Altman 1975). We focus on later works extending Altman's theory to modern digital communication. Specifically, we consider Petronio's Communication Privacy Management (*CPM*) theory (Petronio 2002) and Palen and Dourish's examination of boundary turbulence (Palen and Dourish 2003).

CPM theory (Petronio 2002) presents a framework for exploring the negotiation of privacy boundaries by multiple parties. Ownership, control, and turbulence are key components. We interpret CPM as follows in the context of retrospective modification. A post's creator owns that information and has the ability to grant access to others. When the owner grants access to a post, other individuals effectively become co-owners. Owners and co-owners have privacy rules for managing access. Implicit rules are a user's preferences, while explicit rules are specified and enforced by retrospective mechanisms. The coordination of rules among owners and co-owners is the process of boundary control. Mismatches in owners' and co-owners' preferences cause boundary turbulence. Thus, one needs to identify the full set of co-owners and their preferences to evaluate retrospective modification. While Petronio's work discusses abstractly how others who gain access to private information become co-owners, it did not specify how to identify co-owners in practice. Petronio also assumed that users view managing privacy as a rule-based system. Our study complements CPM theory by presenting a taxonomy of co-owners and uncovering that participants often did not decide acceptability based on fixed rules, but rather by navigating tensions in an ad hoc manner.

Palen and Dourish argue that one of the key privacy boundaries online is the identity boundary (Palen and Dourish 2003). In our context, we interpret this theory to mean that a post creator's privacy expectations might not match the expectations of others who can access the post. This mismatch creates tension at the identity boundary. Palen and Dourish, however, did not discuss co-ownership of information in their umbrella notion of "others," nor the tensions co-ownership causes. Researchers have begun expanding boundary-regulation theories of privacy to account for co-ownership (Lampinen et al. 2011), identifying an array of distinct boundaries in such scenarios (Wisniewski et al. 2016).

Co-ownership of content complicates privacy management. Woodruff identified particular problems with disclosure on social media when people other than the post creator are mentioned in a post or when the audience is large (Woodruff 2014). Such large, interconnected audiences are common (Bernstein et al. 2013). Most prior studies on multi-party privacy have examined how users navigate sharing photos of others, which is one form of content co-ownership. These situations are sometimes subtle, such as when co-owners' involvement is only implied by a photo's apparent context (Such et al. 2017). To mitigate these tensions, users adopt strategies including untagging (Wisniewski et al. 2015) and reporting (Such and Criado 2018; Fogues et al. 2017). They also negotiate offline (Wisniewski, Lipford, and Wilson 2012), in addition to trying to anticipate consequences of sharing co-owned content (Lampinen et al. 2011). Researchers have proposed interfaces to empower co-owners to co-manage content (Besmer and Richter Lipford 2010). Other work has viewed co-ownership of information through other lenses, finding context-dependent attitudes (Marshall and Shipman 2015; Hanson et al. 2020).

While existing work has thus begun to consider issues of shared ownership, it says little about retrospective modification of content, nor about extending conceptions of co-ownership to include participants in the discourse around a post (e.g., likes, comments, and reshares) on social media. Focusing on these two areas, we begin filling this gap.

Taxonomy of Stakeholders and Mechanisms

Changing content retrospectively on social media introduces nuanced issues. Correcting typos can improve clarity and deleting content can protect privacy, yet removing controversial content may let a user evade accountability. In this section, we present our framework of the stakeholders and mechanisms in retrospective modification. Current platforms differ substantially from each other in the mechanisms they provide and the markers, if any, that identify changed content.

Stakeholders

To create a taxonomy of the stakeholders in retrospective modification, we documented which other users could be explicitly identified in a post (e.g., as someone who liked or commented on a post) on popular social media platforms, as well as who could view or otherwise interact with a post (e.g., deleting it in moderation). We created our final stakeholder taxonomy by taking the union of these roles, further guided by theory and prior work.

Stakeholders are users who either take part in retrospectively changing content or who may be affected by others' changes. We identified six distinct roles. *Creators* are the users who initially made a post. Most prior work focuses on them. The *direct exposure set* contains users other than the creator who are explicitly associated with the post (e.g., tagged in a photo). The small literature on multi-party social media privacy considers them (Such et al. 2017; Wisniewski et al. 2015; Such and Criado 2018; Fogues et al. 2017). The *indirect exposure set* contains users taking part in the discourse around a post (e.g., liking or commenting) despite not being an explicit data subject. We are among the first to examine them. *Moderators* are users appointed by the platform who can change a post (e.g., Facebook content reviewers). In practice, moderators typically enforce the retrospective content-moderation policies set by platform owners or developers (e.g., removing hate speech from the platform). *Suggesters* flag content for review by moderators. The *impact set* contains everyone who can view a post, corresponding to co-owners in CPM theory (Petronio 2002).

Note that our stakeholder taxonomy is primarily applicable to users who are not public figures. Furthermore, in this paper we only examine content that is not of broad public interest. Examining perceptions of retrospective changes pertaining to widely distributed content or public figures (e.g., politicians) gives rise to an interesting policy debate (Orlowski 2014). Such cases are outside the scope of this work and complementary to our current exploration.

Mechanisms

We also cataloged the retrospective mechanisms and markers platforms currently provide. We examined the ten English-language social media sites with the most active users as of April 2018: Facebook, WhatsApp, YouTube, Facebook Messenger, Instagram, Tumblr, Reddit, Twitter, Skype, and LinkedIn. We also examined three other services (Viber, Snapchat, and 4chan) known to exemplify other mechanisms. We created test accounts on these thirteen platforms. We made test posts on each account, then tried to edit, delete, and modify these test posts. Our actions were guided by our own past experiences and web searches about how to perform modifications. We repeated this exercise using the authors' long-term, personal accounts. We also referred to each site's official documentation, including help and policy pages.

We collectively term deletion, editing, and modification *retrospective mechanisms*. Table 1 presents our taxonomy and indicates which platforms support which mechanisms. Figure 1 provides screenshots of a number of these mechanisms. Retrospective mechanisms fill a broad design space with seemingly little consensus on the best approach.

Deletion mechanisms permit content to be removed. Implementations vary in which responses (likes, shares, and comments) are deleted alongside a post, whether the creator can delete other users' responses, whether the creator can delete the post but not responses, and whether reshared versions of a deleted post remain. Some platforms permit automatic deletion (e.g., deletion after a set time) or permit moderators to delete content. A crucial difference across platforms is whether deletion is silent or indicated by a marker.

Mechanism	Variant	Platforms
Delete post and responses	Marker	Reddit *(Fig. 1(a))*
	No marker	Facebook, YouTube, Instagram, Tumblr, LinkedIn, 4chan
Delete the post; keep responses	Marker	Skype *(Fig. 1(b))*, Viber
	No marker	WhatsApp, FB Messenger, Twitter
Delete responses; keep the post	No marker	Facebook *(Fig. 1(d))*, YouTube, FB Messenger, Instagram, Reddit, LinkedIn, Tumblr
Delete post; keep reshared copies	No marker	WhatsApp, FB Messenger, Tumblr, Viber, Snapchat
Automatic deletion of post	Time	Instagram and Facebook stories, Snapchat *(Fig. 1(g))*
	Inactivity	4chan *(Fig. 1(h))*
Archive	No marker	Facebook, FB Messenger, Instagram *(Fig. 1(f))*
Deletion by moderator	Notification	Facebook, YouTube, Instagram, Tumblr, Twitter
	No marker	Reddit, 4chan
Edit post content	Edit history	Facebook *(Fig. 1(e))*
	Marker	Instagram, Viber, LinkedIn, Skype *(Fig. 1(b))*
	No marker	Tumblr, Snapchat, YouTube, Reddit
Anonymize	Marker	Reddit *(Fig. 1(c))*
No changes allowed	–	WhatsApp, FB Messenger, Twitter, 4chan

Table 1: Retrospective mechanisms different platforms implement.

(a) Deleted Reddit posts are denoted "[deleted]."

(b) Skype shows deletion with a trash can and edits with pencils.

(c) Reddit posts can be anonymized, changing the poster to "[deleted]."

(d) Facebook post creators can delete a post's comments.

(e) Facebook provides a public edit history.

(f) Instagram users can archive posts.

(g) Snapchat allows deletion after a set time.

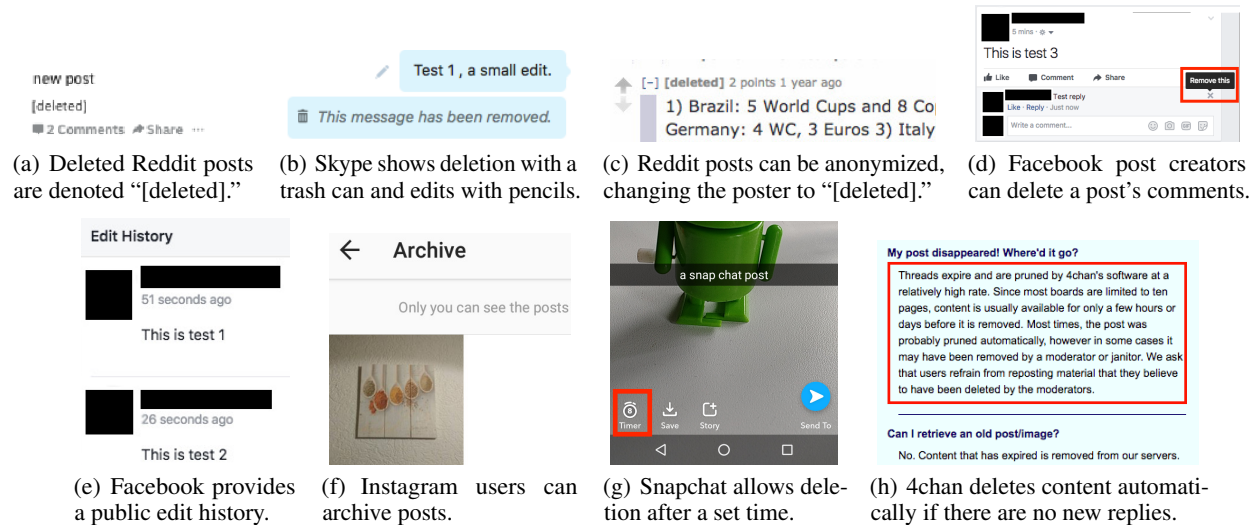(h) 4chan deletes content automatically if there are no new replies.

Figure 1: Example retrospective mechanisms and their associated markers as currently deployed on popular platforms.

Facebook, Instagram, and Tumblr permit creators to delete both their own content and other users' responses. Reddit allows the creator to delete their post, but not responses. Communication platforms WhatsApp, Skype, Facebook Messenger, and Viber allow the creator to remove their own content, but not content contributed by the indirect exposure set. Recently, Whatsapp and Viber began deleting other users' copies if the creator deletes their copy, yet old versions of these apps delete only the creator's copy (Warren 2017).

Snapchat specializes in auto-deletion after short time periods, while "stories" on Instagram and Facebook are auto-deleted after time periods like one day. Other types of Instagram and Facebook posts are not auto-deleted. Content is moderated on Facebook, YouTube, Instagram, Tumblr, Twitter, Reddit, and 4chan. However, perhaps because Reddit and 4chan are heavily moderated and frequently remove posts, they do not notify creators by default.

We find many platforms prioritize creators' agency in deletion, allowing them to remove posts and responses. However, heavily moderated platforms Reddit and 4chan do not allow creators to remove content from the indirect exposure set.

In addition, Instagram and other platforms permit content to be archived, meaning the content is available to the creator, but functionally deleted for everyone else. Platforms like Facebook with granular privacy settings permit content to be selectively hidden from only certain members of the indirect

exposure set. Furthermore, platforms can also make content harder to find (e.g., removed from a feed) even if the indirect exposure set remains the same. Making content difficult to find can be tantamount to removal.

Edit mechanisms permit content to be changed in place. Platforms again vary widely in the mechanisms provided. Twitter, WhatsApp, Facebook Messenger, and 4chan do not allow creators to edit posts. Tumblr, Snapchat, YouTube, and Reddit allow creators to edit content without leaving any marker. In contrast, Facebook provides the full edit history of a post, as shown in Figure 1(e). Further, Facebook does not allow photos to be modified at all (Sanchez 2017), though metadata like photo captions can be edited. Skype, Instagram, Viber, and LinkedIn allow creators to edit content, denoting edits with a visual marker as in Figure 1(b). An edit marker could indicate when an edit was made, but not the edit itself, though we did not observe this variant.

Modification mechanisms let stakeholders change a post's metadata, but not its primary content. Platforms like Reddit permit posts to be *anonymized*, keeping the content while removing the creator's identity. Figure 1(c) depicts anonymization on Reddit, which replaces the creator's username with "[deleted]." In contrast, on platforms like 4chan's /b/ forum, no modification whatsoever is allowed.

Completeness of Our Taxonomy

We used our taxonomy to unpack different stakeholders' perceptions of retrospective mechanisms and markers. As we explained earlier in this section, our taxonomy was driven by our own interactions editing, deleting, and modifying posts, as well as our observations of which other users could be explicitly identified in a post. In particular, we focused on the user interfaces provided by current social media sites. Consequently, a potential concern is that our taxonomy might contain only the most salient mechanisms, rather than all possible mechanisms. To that end, as part of our study, we presented this taxonomy to interview participants and asked them to identify mechanisms that appeared to be missing. Specifically, we explicitly asked if the participant recalled encountering any other retrospective mechanisms on any platform they use. The answer to this question was always negative. This result, along with our best-effort investigation of popular social media platforms' interfaces, gave us confidence about the completeness of our taxonomy.

Method

We investigated our research questions using rich qualitative data collected in IRB-approved interviews of 15 participants. Our protocol explored how social media users perceive different retrospective mechanisms and the reasoning underlying these perceptions. Each participant completed an online pre-survey and an in-lab, semi-structured interview. We asked about both specific social media posts and general perceptions. All interviews were audio recorded and transcribed.

Participant Recruitment

We recruited participants from a large city in the USA on Craigslist and by posting flyers. We required participants be at least 18 years old. So that we could ground the interview in social media posts from a variety of life phases, we required participants to have an account on Facebook, Twitter, Instagram, Reddit, or a similar site. We required that this account be at least three years old. We compensated participants with a $35 Amazon gift card, plus up to $10 for transportation costs. Participants took 75 minutes on average to complete the interview, which took place on our institution's campus.

We continued to recruit until reaching saturation in participants' explanations of how they used and perceived retrospective mechanisms and associated markers. In this work, we took a hybrid approach of achieving both *theoretical saturation* and *inductive thematic saturation* (Saunders et al. 2018). Specifically, after each interview during our data-collection phase, we checked if our aggregated data covered perceptions about all stakeholder roles in our taxonomy. Furthermore, for each stakeholder role, we also checked if any new meta-themes had emerged from the most recent participant's explanations. We stopped our data collection once our data covered all stakeholder roles (theoretical saturation) and additional meta-themes were consistently not emerging from the new data (inductive thematic saturation). While we reached saturation relatively quickly, our interviews covered substantial ground in their average of 75 minutes. Furthermore, the typical participant had wide-ranging experiences with, and opinions about, many different retrospective mechanisms.

Study Structure

To ground part of the interview in specific posts, we required that participants complete an online survey prior to their interview. This survey asked them to identify their most frequently used social media site and to find eight past posts on that site based on criteria listed below. We asked participants to screenshot those posts to reference during the interview.

We aimed to elicit posts for which retrospective mechanisms might apply. Inspired by prior work on retrospective mechanisms (Bauer et al. 2013; Ayalon and Toch 2017), we asked participants to identify four posts from a year ago and four posts from the first year of their account. Inspired by CPM Theory (Petronio 2002) and privacy research (Litt and Hargittai 2016), we asked participants to select these posts based on role and activity. For the four posts in each time period, we instructed participants to choose a post in which they had been tagged or on which they had commented, a controversial post, a post they had retrospectively changed (or had considered changing), and a post that had generated many likes, shares, or comments. Participants completed an online metadata form for each post, and they brought screenshots of each post to the interview on their own device. While we acknowledge that this approach was biased toward posts more likely to be modified retrospectively, doing so helped elicit participants' perceptions of retrospective mechanisms.

The semi-structured interview itself consisted of two parts followed by an exit survey about demographics and social media usage. The first part focused on the posts selected in the pre-survey. For each post, participants answered a series of questions about what types of retrospective changes, if any, they had previously either applied or considered applying. We also asked whether they might now consider making

any changes, specifying that they could propose changes regardless of whether they are currently possible on their chosen platform. To further investigate content co-ownership, we asked whom the participant thought would be affected by any potential changes, as well as whether they expected these people would find the changes acceptable. Our questions about who would be impacted by retrospective mechanisms and our analysis of the tensions that surfaced were influenced by CPM's notions of boundary regulation and turbulence.

The second part focused on the necessity and acceptability of retrospective mechanisms in a more abstract setting. We asked participants to answer independent of any specific post or platform. We asked follow-up questions whenever the participant said their response would depend on the platform. We described each mechanism from our taxonomy, asking, "For posts that you made, would you find applying this mechanism to be acceptable always, sometimes, or never?" If they chose "sometimes," we asked about concrete situations or posts where it would be acceptable or unacceptable. If the participant answered "always" or "never," we tried to elicit the reasoning for this perception. We asked additional questions about how each mechanism might impact others.

To revisit all mechanisms in light of the entire conversation, we then handed participants a full list of the mechanisms previously introduced. We went back through this list, asking which mechanisms they would prefer to use in general and which ones they believed to be necessary for any platform. We specifically asked how platforms could resolve tensions between changing content in beneficial and harmful ways.

Data Analysis

To develop a holistic understanding of attitudes and perceptions, we performed qualitative analysis. Because our interview discussed both specific posts and abstract thoughts independent of platform, participants expressed attitudes about a given mechanism at many points in the interview. As a result, we chose to perform affinity diagramming (Harboe and Huang 2015) to identify and cluster themes in participants' explanations. We transcribed all 15 interviews. Two authors then independently read through the transcripts, flagging all quotes from the transcripts that were potentially explanatory of attitudes toward retrospective mechanisms and their impact on stakeholders. A total of 1,132 quotes from the 15 interviews were identified by at least one author as potentially explanatory. We then collaboratively created affinity diagrams of these 1,132 quotes, which resulted in 8 clusters representing high-level themes, with a total of 58 sub-themes.

Limitations

Most participants (8 of 15) chose Facebook as the platform they use most frequently for the first part of our interview, and some other platforms (e.g., Reddit) were only represented once in this first part. However, in the second part of the interview, participants answered independent of platform. Further, we could not directly investigate perceptions about already-deleted posts since they had been removed.

Our focus on retrospectively changing content likely biased participants toward considering changes at a higher rate than normal. That said, we were interested not in the frequency of changes, but in their potential impact.

Furthermore, we interviewed a convenience sample recruited primarily from among non-student community members living near a university. Our participants were younger and more educated than the broader U.S. population. Social media usage varies by demographics, location, and education (Correa, Hinsley, and De Zuniga 2010; Lenhart et al. 2010). Furthermore, prior work has found that the incidence of negative privacy experiences is correlated with education; individuals with lower educational attainment reported negative privacy and security incidents at equal or lower rates than those with higher educational attainment in a census-representative sample (Redmiles, Kross, and Mazurek 2017). Contrary to prevailing assumptions in the research community, that same work found socioeconomic status not to be correlated with the number of negative privacy and security incidents reported (Redmiles, Kross, and Mazurek 2017). Nonetheless, it is important to be aware of how the digital divide and potentially vulnerable populations are reflected in user study samples, even for small-scale studies. We did not formally ask participants about their racial identity or income. That said, our institution is located in a predominantly African-American area of our city, and the neighborhood in which it is located and all adjacent neighborhoods have median household incomes ranging from approximately $25,000 to $57,000; these are all lower than the US national median household income (roughly $63,000). We recruited participants primarily from the neighborhoods surrounding our institution and have no reason to believe that our participants' demographics deviated substantially from those of the surrounding neighborhoods.

Despite these limitations, we feel our data remains suitable to answer our research questions for two reasons. First, our focus was on developing a rich, qualitative understanding. To that end, we collected and analyzed a dataset that reached thematic and theoretical saturation. While it may seem surprising that we reached saturation in a 15-participant sample given that platforms differ widely, our focus was on eliciting abstract perceptions and reasoning, not gathering feedback on platform-specific implementations. Prior work has reported reaching saturation in meta-themes with a similarly sized sample (Guest, Bunce, and Johnson 2006). Second, our participants were long-term and frequent users of social media, giving them a rich array of experiences related to retrospective modification, the focus of our study.

Results

We begin by describing our participants before exploring which retrospective mechanisms participants found useful for themselves, which mechanisms they considered socially acceptable, what factors impact acceptability, and what tensions arise. We identify participants with labels comprising the participant number, the platform they were discussing at the time, and their role in that scenario. For example, P3-FB-C indicates Participant 3 talking about a post on Facebook from the perspective of a creator. Platforms and their abbreviations are Facebook (*FB*), Twitter (*TW*), Instagram (*IN*), Reddit (*RE*), and Tumblr (*TU*). We use *ANY* if no particular

platform was being discussed. Roles include membership in the direct exposure set (*DE*), indirect exposure set (*IE*), impact set (*I*), and sets of suggesters (*S*) and moderators (*M*). We also describe a meta-grouping of post non-creators (*NC*).

Participants

Fifteen people (nine female, six male) participated in our study. Regarding age, seven participants were 25–34, five were 18–24, two were 35–44, and one was 45–54. In contrast to typical academic convenience samples, only two participants were students. Seven participants held full-time jobs, five held part-time jobs, and one was unemployed. Only one participant had a technical background. While all participants had received at least some college credit, seven did not hold a degree (and only two were currently pursuing one), two held associate's degrees, three held bachelor's degrees, and three held graduate degrees. In sum, our participants were young and fairly well-educated non-student users of social media.

Participants completed the first part of the study with their primary social media account. Eight participants primarily used Facebook, three primarily used Twitter, two primarily used Instagram, one primarily used Reddit, and one primarily used Tumblr. Participants had used these primary accounts for a mean of 7.1 years (range: 4–13 years), with a mean of 490 friends or followers (max: 1,200). They reported daily social media usage ranging from five minutes to over ten hours, with a median of one hour. In the second part of the study, participants answered questions from the perspective of all platforms they used frequently. They used Facebook (12 participants), Instagram (6), Twitter (5), Snapchat (4), Reddit (3), Pinterest (2), LinkedIn (1), and Tumblr (1). Fourteen of the fifteen participants recalled deleting a post, ten recalled changing a post's audience, nine recalled asking a friend to delete a post, and eight recalled editing a post.

The Utility of Retrospective Mechanisms (RQ 1)

Recall that RQ 1 focused on understanding the contexts in which post creators view different retrospective mechanisms as useful, necessary, and acceptable. We answer this question in two parts. In this section, we first examine perceptions of the utility of these mechanisms. The subsequent section presents perceptions of their acceptability and necessity. We observed three common reasons participants considered retrospective mechanisms broadly useful for post owners.

Maintaining accuracy: Typos and grammatical errors could be embarrassing for participants, and they wanted to be able to correct them. As P7-FB-C explained, *"I have edited posts in the past. I think they were mostly for spelling errors. I would consider doing it again if I see an error."*

Preserving control: Participants wanted to retain agency and control over content they had posted. Many participants felt such agency was an inherent right. P12-TW-C explained, *"If it is your post, you should be able to do something to it."*

Minimizing abuse and providing protection: Mechanisms like anonymization, deletion, and moderation were considered useful for minimizing abuse. P6-IN-C felt it important to be able to delete content that *"is harmful or hurtful."* Twelve participants deemed moderated deletion useful. In fact, P1-FB-S desired even more moderation: *"You know they*

say you can report people for certain type of pictures, but you really can't. . . there was no way for me to do it." Participants also found mechanisms like retrospective anonymization (e.g., on Reddit) useful for protection. P6-RE-C said, *"I anonymized my post, or my name, before. And I think it is very helpful if you don't want to associate yourself with it."*

Mechanisms' Acceptability and Necessity (RQ 1)

To further answer RQ 1, we asked whether participants find it 'always', 'never', or 'sometimes' acceptable to use each mechanism. 'Sometimes' was the most common response, indicating context-dependent tensions. For most mechanisms, participants identified at least some situations in which applying the mechanism would create tension. Prior work identified a few of these tensions, such as creators finding it acceptable to delete their own regrettable content (Sleeper et al. 2013). However, to our knowledge, the observations we detail below regarding post visibility, the size of the edit, and the perceived usage of the account are novel.

Responses (All mechanisms): Participants felt the amount and the types of responses to a post, specifically comments and tags, dictated acceptability. For example, they indicated that deleting a post or its comments would be unacceptable if it removes essential context in a debate. P6-IN-I explained, *"It is only deleting the one side of the story, so there is no way of referencing what was originally said."* Others felt it unacceptable to delete posts with many responses.

Participants felt the impact of archiving was similar to deletion. The types of activity in response influenced the acceptability of editing. P2-TW-C said, *"The only unacceptable example would be changing something that people have already discovered to be. . . offensive."* Participants noted a tension between the creator and those in the direct and indirect exposure sets because those stakeholders were either tagged or had seemed to endorse the content via liking or commenting. This tension increases for high-visibility posts.

Creator's intent (All mechanisms): Participants saw the deletion of all offensive posts, as well as some other posts, as acceptable. They supported the deletion of regrettable, embarrassing, harassing, or abusive content as long as they felt the intent of deletion was to reverse a mistake or ensure safety. P6-IN-I explained, *"I don't find it acceptable [to delete], just because to look back on the memories, but I know that if it was a stressful situation, or if it caused a lot of pain, I think that would be acceptable. Also if you offended somebody."* In contrast, participants found deletion that censors or deceives others unacceptable. For example, moderation based only on personal biases was unacceptable. P9-ANY-C said deletion was unacceptable *"if it's something that doesn't violate anything, but is just unpopular, or maybe they're targeting you specifically, or maybe it's the moderator's personal opinion."*

Participants supported edits with good intentions, yet felt editing had the potential for deception. The size of an edit was a deciding factor. P2-ANY-C said small edits are *"acceptable when you made a typo or just had an immediate change of thought or omitted something that you meant to include. Usually for mistakes."* However, large edits or edits to controversial content were unacceptable. P15-ANY-I said large edits cause confusion *"if you post something and people*

reply *'you're amazing,' 'I love you,'* and then you change it to something really bad, but you maintain those likes."

In sum, deletions and edits to correct mistakes and regrets, or to update content with good intentions, are acceptable. Unfortunately, intent is exceedingly difficult to measure.

Public vs. private platform (Deletion only): Participants were divided into two camps about deleting content based on their perception of a platform as public or private. Most participants saw Twitter as more of a public space than Facebook. When viewing social media as a public space to showcase content, participants preferred mechanisms that allow others to keep a copy of a deleted post. P10-TW-C felt that *"for them to be able to keep [retweets] on their wall, that's fine"* even if the original tweet was deleted. P2-TW-IE identified a different part of the direct exposure set as impacted: *"[A creator's own copy of a post] is his content that he shared, so if he no longer wants to share, then that would be acceptable...I would probably just find similar content to retweet."*

In contrast, perceptions of social media as a personal space that others can also view echoes conceptions of a "personal region" in social media (Zhao et al. 2013). P3-FB-C said that if friends saw her political posts *"on Facebook and they liked and commented on it, and then I deleted it, I think they may question why. But I don't think they would think it was unacceptable. I think the majority of them would think that it is her Facebook and she can almost do whatever she wants."* This sense of ownership and agency was common.

Family or close friends (Deletion only): If a post recalls a good memory, four participants felt it unacceptable even for the creator to delete it. Describing why a specific post should not be deleted, P7-FB-C explained, *"This is a close friend of mine. She wanted to celebrate, I guess, 7 years of Facebook friendship [with me]."* P14-FB-C would not delete content involving family members or close friends: *"They would take it quite offensively if I delete their things or untag them from this memorable event. They'd be like, why me, what did I do, what's wrong?"* However, the involvement of family members or close friends was sometimes mediated by negative factors like a post's irrelevance or lack of responses.

Stakeholder awareness (Editing only): Participants found editing more acceptable if the other stakeholders were aware of the edit. They felt that if others had already discovered, shared, or commented on a post, creators should be especially judicious about edits. They suggested such edits should be accompanied by markers or notifications to people in the direct or indirect exposure sets. P12-ANY-C stated, *"I think any edit can be made as long as people are aware of the change...because then they can have the choice of deleting their comments or removing their name."* However, creators do not always want to advertise edits that cover up a mistake. P6-FB-IE mentioned that a marker *"just brings more attention to the fact that they had to edit it."* A tension exists between transparency and drawing attention to errors.

Safety vs. accountability (Anonymization only): Some participants noted anonymization's importance for safety. P14-ANY-I stated, *"You can anonymize if it will come back to bite you, if it can get you in serious trouble...to protect yourself."* Other participants, however, mentioned that anonymization opens the door for abuse, so such mechanisms

need moderation. P9-ANY-I said, *"I'd still find it to be unacceptable if they could anonymize and it was something really offensive...I think maybe rules about what you can post anonymously, [involving] moderators,"* would help. Some participants, though, felt anonymization was necessary for debate, venting, or healthy arguments. P11-TU-C implored, *"If you vent something, you shouldn't have to sign your name."*

Control (Platform-centric mechanisms only): A few current platforms permit automatic deletion or moderation. Others do not permit any retrospective modification. Participants' opinions of these approaches revolved around the degree of control retained. All participants considered it unacceptable to forbid any retrospective modification because it takes away creators' control. P11-ANY-C said, *"Everyone makes some sort of mistake. Even if it is not a big deal, someone will write something that they didn't necessarily intend to. They have a spelling or grammatical error that they want to go back and fix."* Nine participants deemed moderated deletion necessary even though it takes away control. They felt, however, a notification and reason must be given. P7-ANY-C said, *"There should be moderation in all social media because of activity which is inappropriate, and illegal, unfair, and violates other people's rights."*

Participants found platform-controlled automatic deletion unacceptable unless such a feature was the main purpose, such as for Snapchat or Facebook Stories. P13-ANY-C mentioned, *"If it is the post owner [setting a deletion time], I think that it is fine...I feel like the platform shouldn't make the decision to delete things after a certain amount of time."*

Lessening control might also be desirable for convenience. P2-ANY-C supported inactivity-based deletion because *"I could see practical uses for that, like maybe you are promoting something that was no longer valid."* While our participants did not want to lose control to subjective moderation or automatic deletion, they sometimes found mechanisms that take away some control acceptable during discussions (preventing deletion) or for convenience (automatic deletion). Control depends on context in nuanced ways.

Modifications' Impact on Stakeholders (RQ 2)

Next, we move onto answering RQ 2, which asks which stakeholders (aside from post creators) participants believe to be impacted by retrospective changes. Participants felt that non-creator stakeholders can be impacted by retrospective modification in certain situations.

Impact depends on content type: Participants reported that retrospective modification of content that was family-related, embarrassing, or captured good memories would impact close friends and family. They felt members of the direct and indirect exposure sets (those who interacted with the content) might lose memories from modification. P8-FB-C explained, *"Certain family members [would be impacted] because they may decide, oh, I want to go back and look at that post, I really like the post."* However, five participants felt that both the exposure and impact sets would be positively impacted if embarrassing content disappeared. P14-FB-C did not *"think anybody would be [adversely] impacted...If it is deleted, no one could say anything about our silly faces."*

Impact depends on the relationship to the post: Even within the direct exposure set, participants distinguished between users who liked the content, commented, shared it, and who were tagged. This facet arose particularly for family and close friends, especially regarding content involving them. Explaining who would be impacted by deleting a post and its comments, P9-FB-DE said, *"Probably all the people who are tagged in it, if they don't have a copy. Like I don't think I saved any of the pictures on the computer. . . so I would lose the picture."* P8-FB-C disagreed, saying, *"I don't think people who comment and liked it would be as impacted as someone who shared it because I feel, like, if you're gonna share it to your page, then it must really mean something."* Furthermore, participants thought if the post relates to family or close friends, then those close connections will be significantly impacted by retrospective changes. In contrast, only two participants felt modification would impact users in the impact set (those who only viewed the content).

No impact for old or redundant content: Some participants felt that no one would be impacted if the content had been posted long ago or if similar content existed elsewhere. P13-FB-IE explained, *"It was long enough ago that people wouldn't really care. . . There are other pictures of that night, similar pictures, so I don't think anyone would be impacted."*

Markers, Revision Histories, Notifications (RQ 3)

Markers, revision histories, and notifications to stakeholders can mitigate some tensions of retrospective modification. For example, Facebook's revision history provides accountability for edits. To answer RQ 3, we now investigate the context-dependence of participants' preferences for markers, revision histories, and notifications.

Changes in the meaning of a post: Participants were reluctant to receive more notifications than they already do on social media. They wanted markers, revision histories, or notifications mainly for large edits that might change a post's meaning. P13-FB-IE *"would only want a notification if the context of the post is changed."* The post's relevance also influenced preferences. P10-FB-IE said, *"From the majority of the tweets that I endorsed or retweeted, I wouldn't want that notification, but in this minority situation I would."*

Mistakes should not have markers: When fixing embarrassing posts, accidents, or mistakes, participants wanted the option to have no revision history. Some did not even want a marker. P14-FB-C found deletion without a marker acceptable in some situations, saying, *"Even if it is 15, 20 years ago, I don't [need markers] because it's a good thing to get rid of, so even people who are tagged don't have to know. I did you a favor: gone."* Some participants felt markers were superfluous since the act of deletion acknowledges a past mistake. P6-ANY-M explained, *"People make mistakes all the time. The problem with the internet is that it is there for everyone to see. . . People should be allowed to fix their mistakes."*

Markers, revision histories, and notifications are not a silver bullet for resolving tensions, yet are useful for preserving context when silent modifications could corrupt coherence.

Perceptions of Post Non-Creators (RQ 4)

Finally, we answer RQ 4, which concerns the degree to which perceptions change when considering retrospective modification from perspectives other than the post creator's. Contrary to our hypotheses, across all mechanisms participants indicated that their responses would not have changed had the post creator and other stakeholder roles been reversed. In other words, when answering from the perspective of a post non-creator, they reported that they would feel the same as if they were the post creator. For example, when discussing non-creator roles, P3-FB-NC expressed, *"I think I would feel the same way. I would say they could do whatever they want."* For P3, respect for the ultimate agency of the post creator, regardless of whether or not she herself was the post creator, guided her perception. Similarly, when talking about deletion from the perspective of a post non-creator, P2-TW-NC explained, *"I think it's kind of the same rules that apply, where it is fine unless they are just kind of trying to hide something."*

This lack of differentiation carried over to markers. After discussing her belief that edit histories are necessary for maintaining the integrity of posts she had created and subsequently modified, P10-TW-NC clarified that she felt the same when she was not the post creator: *"If I was part of a conversation, then I would want all of that information about the edit history. . . If I didn't have it. . . it could be confusing."*

Across all cases we discussed with participants, they expressed a consistent belief that their views applied equally to situations in which they were and were not the post creator. This finding suggests that perceptions and underlying tensions are not dependent on stakeholder roles, but rather on the context of usage for the retrospective deletion, editing, or modification mechanisms.

Implications for Mechanism Design

In this section, we use participants' comments on mitigating the tensions of retrospective modification to synthesize and highlight design principles for respectful retrospective mechanisms. P1-FB-C captured this respect, saying, *"If I edited, I could say that 'I edited what you see and I hope you don't mind.' I would be mindful."* However, courtesy alone is insufficient for resolving tensions systematically.

Adding markers to preserve discussion context: For platforms that are heavily conversation-based or for which different stakeholders' contributions build off of each other, it is critical to deploy markers to preserve the context of a discussion in spite of potential retrospective changes like deletion. Participants felt deleting comments without providing a marker or deleting a post without deleting its responses removes context. Discussing their aversion to deleting a post but leaving comments, P8-ANY-C explained, *"People are gonna wonder what are these comments for."* Thus, a direct and simple way to mitigate this tension is to add a marker in place of the deleted comment or post, which preserves context while respecting the retrospective change's potential for addressing mistakes or regrettable content.

User-controlled timers for automatic deletion: While participants considered the ability to retrospectively modify posts necessary to retain control, at the same time they

thought some retrospective mechanisms themselves could impair control. For example, participants opposed automatic deletion if the platform sets the deletion date. In contrast, they expressed that automation deletion is useful as long as the user controls the time of deletion. P6-ANY-C explained, *"A predetermined date. . . I would do that if it was set by the user."* Thus, automatic deletion mechanisms should always include a user-controlled timer to make them useful and acceptable.

Auto-detecting the intent of an edit: We found that the timing and magnitude of edits are important in selecting appropriate markers. A small edit made immediately to a post might indicate fixing a mistake, making edits acceptable even without markers. P5-ANY-M said, *"I would give them a chance. . . a few hours to decide to change that."* In this case, we envision that social media platforms should automatically check for small edits (e.g., based on Levenshtein edit distance) and abstain from sending notification for them. In fact, earlier work that already developed automated methods to detect typos and minor rephrasing in social media could be leveraged for this purpose (Almuhimedi et al. 2013). However, for larger edits, the platform should notify the creator about adding a justification to reduce tensions.

Detecting and informing impacted stakeholders: We found that participants felt that members of the *direct exposure set* and *indirect exposure set* are impacted stakeholders in many circumstances of retrospective modification, though the impact on each set differs. Whereas CPM theory acknowledges viewers as a post's co-owners (members of the impact set), our findings extend CPM's concept of co-ownership to encompass additional stakeholders who should be treated differently depending on the situation and their relationship to the modified content. Participants expressed that markers or notifications should be mandatory only for posts with many comments or many users tagged. P6-ANY-DE said, *"To delete the post and its comments, I think I would like a marker if I was tagged in it or someone else was was tagged in it."* P9-FB-C emphasized considering the number of stakeholders, saying, *"If someone posts. . . a non-controversial topic. . . and all the comments are really vague like 'I support this,' someone could just change the post content, but those comments would just stay there."* We also found that, for posts related to family and close friends, participants perceived retrospective modification's potential impact as especially powerful.

Thus, we envision a design where the social media operators automatically identify stakeholders and rank their closeness for posts that are retrospectively modified. In appropriate contexts, the site should automatically notify members of the most impacted groups. Our exploration identified the following groups as most impacted for a given post: tagged users, users who commented, and close social ties (family and close friends) referenced in the post. There is a huge body of social media research focused on community detection, tie-strength detection, and topic detection (Papadopoulos et al. 2012; Gilbert and Karahalios 2009; Lai and To 2015). We envision that social media platforms could easily leverage this prior work to automatically prioritize which users should receive notifications in case of a retrospective modification.

Designing nudge-based notifications : Many of the design directions suggested by our study involve sending notifi-

cations to impacted users. To that end, we envision that social media platforms should design behavioral nudges to help the user who is modifying content use their unique knowledge about contexts and situations to proactively notify stakeholders they believe will be impacted. Prior work on privacy nudging focused on post creation (Wang et al. 2014). However, analogous nudges during retrospective modification might also help users recognize a modification's potential impact. As they delete a post, users could be encouraged, yet not forced, to send archival copies to the most impacted stakeholders, who perhaps could be detected automatically as suggested above. As they edit a post, users could be shown photos of the other stakeholders and be encouraged to notify those who may be impacted. The exact design of such nudge-based notifications requires further research, however.

Discussion and Conclusions

Our study shed light on user perceptions of retrospective mechanisms in social media. We found that co-owners (those who liked, commented, or shared a post someone else made) are affected by retrospective changes. They sometimes deem certain retrospective modifications unacceptable. This creates tensions among stakeholders, causing boundary turbulence.

Agency of content creators and co-owners: Participants noted the importance of content creators' agency over their content. Their responses indicated that a no-modification-allowed policy is unacceptable since it might result in a loss of control for creators. They also acknowledged the agency of co-owners, as predicted by CPM theory. However, this creates a tension between respect for co-owners and the content creator's agency in having the freedom to make changes.

Fortunately, our results also point out ways to mitigate this tension. Participants did not see co-owners as a monolithic group and felt that not all stakeholders have equal rights or expectations to manage content. This finding deviates from CPM theory in our context since CPM does not distinguish among classes of co-owners. Participants generally felt that users in the direct exposure set (e.g., those tagged), family members or close friends, and users who comment on a post are more affected than others by retrospective mechanisms. Notably, the set of users who comment on a post and thereby enter into a post's surrounding discourse, but are not otherwise data subjects, has been mostly overlooked in prior work (Such et al. 2017; Wisniewski et al. 2015; Such and Criado 2018; Lampinen et al. 2011; Besmer and Richter Lipford 2010; Marshall and Shipman 2017; Wei et al. 2020).

Participants found it acceptable to give substantially impacted stakeholders more information about retrospective modification than those who like or simply view a post. This information could include markers, revision histories, and push notifications. In fact, our findings suggest platforms should consider more customized, nuanced, and contextual retrospective mechanisms that enable owners to set different markers or notifications for different stakeholders.

Balancing a "right to be forgotten" and accountability: Participants strongly supported a post creator's right to be forgotten, as evident in their opposition to a no-modification-allowed policy. This desire to let content creators make retrospective changes aligns with the desire to let a content creator

exercise freedom of expression. However, balancing accountability with the right to be forgotten gives rise to key tensions. Participants identified ways to mitigate these tensions.

First, our results suggest that the deletion of past mistakes (even for regrettable, offensive, or abusive content) is often viewed positively. Even though three of our participants found the denial of these past mistakes unacceptable, they still wanted the content to be removed from the platform to prevent further harm. Thus, participants' views aligned somewhat with legislation about the right to be forgotten. However, for selective deletion of posts and comments from a conversation, as well as editing, participants sought accountability and wanted to maintain the coherence of discussions. Without a revision history, conversations embedding selective modifications could become incoherent. Thus, platforms should likely implement the right to be forgotten with broad mechanisms.

Second, tensions between accountability and a right to be forgotten are influenced by the creator's intent in making the modification and how the modification recontextualizes co-owners' contributions. For example, participants hoped to minimize misinformation. They did not find editing or deletion acceptable when comments and endorsements could be misrepresented. They felt that modifications reflecting benign intent, such as fixing typos or adding clarification, should be allowed without any markers. In contrast, participants demanded markers, revision history, or notifications for potentially malicious modifications, even for non-stakeholders. The distinction between accountability and agency for self-presentation depends on a modification's intent.

Unfortunately, intent is highly subjective and is therefore difficult to fully automate. However, participants indicated possible avenues for partial automation. For example, editing a post immediately after uploading it, or just changing a few characters, might indicate benign intent. Platforms could explore requiring a content creator to indicate their intent.

Finally, participants identified user-initiated strategies that implicitly enforce accountability. These strategies include co-owners taking and sharing screenshots of conversations, saving the post content via copy-pasting, and the intervention of moderators to enforce civility. Co-owners could play a role in assessing intent and enforcing accountability, such as by asking the content creator in public about modifications.

Future work: We identified tensions between accountability and the right to be forgotten in current retrospective mechanisms. Future research should design and evaluate retrospective mechanisms that empower users to protect their privacy while minimizing boundary turbulence. Another avenue for future work is to compare these perceptions along different demographic axes to design and evaluate retrospective mechanisms tailored to different social groups. Our results also hinted that platform-specific considerations and functionality impact the acceptability of retrospective modification. For example, distinctions about the acceptability of deletion in spaces perceived as public (versus private) demonstrate that some tensions may manifest differently on different platforms. Our mechanism-centric investigation documents perceptions of retrospective modification broadly in social media. Future work should explore why and when people would use particular mechanisms on particular platforms.

References

Almuhimedi, H.; Wilson, S.; Liu, B.; Sadeh, N.; and Acquisti, A. 2013. Tweets Are Forever: A Large-Scale Quantitative Analysis of Deleted Tweets. In *Proc. CSCW*.

Altman, I. 1975. *The Environment and Social Behavior: Privacy, Personal Space, Territory, Crowding*. Brooks/Cole.

Ayalon, O.; and Toch, E. 2013. Retrospective Privacy: Managing Longitudinal Privacy in Online Social Networks. In *Proc. SOUPS*.

Ayalon, O.; and Toch, E. 2017. Not Even Past: Information Aging and Temporal Privacy in Online Social Networks. *HCI* 32(2).

Bauer, L.; Cranor, L. F.; Komanduri, S.; Mazurek, M. L.; Reiter, M. K.; Sleeper, M.; and Ur, B. 2013. The Post Anachronism: The Temporal Dimension of Facebook Privacy. In *Proc. WPES*.

Bernstein, M. S.; Bakshy, E.; Burke, M.; and Karrer, B. 2013. Quantifying the Invisible Audience in Social Networks. In *Proc. CHI*.

Besmer, A.; and Richter Lipford, H. 2010. Moving Beyond Untagging: Photo Privacy in a Tagged World. In *Proc. CHI*.

Bhattacharya, P.; and Ganguly, N. 2016. Characterizing Deleted Tweets and their Authors. In *Proc. ICWSM*.

Child, J. T.; Haridakis, P. M.; and Petronio, S. 2012. Blogging Privacy Rule Orientations, Privacy Management, and Content Deletion Practices: The Variability of Online Privacy Management Activity at Different Stages of Social Media Use. *CHB* 28(5).

Correa, T.; Hinsley, A. W.; and De Zuniga, H. G. 2010. Who Interacts on the Web?: The Intersection of Users' Personality and Social Media Use. *CHB* 26(2).

DeVito, M. A.; Birnholtz, J.; and Hancock, J. T. 2017. Platforms, People, and Perception: Using Affordances to Understand Self-Presentation on Social Media. In *Proc. CSCW*.

Fogues, R. L.; Murukannaiah, P. K.; Such, J. M.; and Singh, M. P. 2017. Sharing Policies in Multiuser Privacy Scenarios: Incorporating Context, Preferences, and Arguments in Decision Making. *TOCHI* 24(1).

Gilbert, E.; and Karahalios, K. 2009. Predicting Tie Strength with Social Media. In *Proc. CHI*.

Guest, G.; Bunce, A.; and Johnson, L. 2006. How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability. *Field Methods* 18(1).

Haimson, O. L.; Brubaker, J. R.; Dombrowski, L.; and Hayes, G. R. 2016. Digital Footprints and Changing Networks During Online Identity Transitions. In *Proc. CHI*.

Hanson, J.; Wei, M.; Veys, S.; Kugler, M.; Strahilevitz, L.; and Ur, B. 2020. Taking Data Out of Context to Hyper-Personalize Ads: Crowdworkers' Privacy Perceptions and Decisions to Disclose Private Information. In *Proc. CHI*.

Harboe, G.; and Huang, E. M. 2015. Real-World Affinity Diagramming Practices: Bridging the Paper-Digital Gap. In *Proc. CHI*.

Lai, L. S. L.; and To, W. M. 2015. Content Analysis of Social Media: A Grounded Theory Approach. *JECR* 16(38).

Lampinen, A.; Lehtinen, V.; Lehmuskallio, A.; and Tamminen, S. 2011. We're in It Together: Interpersonal Management of Disclosure in Social Network Services. In *Proc. CHI*.

Lenhart, A.; Purcell, K.; Smith, A.; and Zickuhr, K. 2010. Social Media & Mobile Internet Use Among Teens and Young Adults. *Pew Internet & American Life Project* .

Litt, E.; and Hargittai, E. 2016. "Just Cast the Net, and Hopefully the Right Fish Swim into It": Audience Management on Social Network Sites. In *Proc. CSCW*.

Liu, Y.; Kliman-Silver, C.; and Mislove, A. 2014. The Tweets They Are A-Changin': Evolution of Twitter Users and Behavior. In *Proc. ICWSM*.

Marshall, C. C.; and Shipman, F. M. 2015. Exploring the Ownership and Persistent Value of Facebook Content. In *Proc. CSCW*.

Marshall, C. C.; and Shipman, F. M. 2017. Who Owns the Social Web? *CACM* 60(5).

Mondal, M.; Messias, J.; Ghosh, S.; Gummadi, K. P.; and Kate, A. 2016. Forgetting in Social Media: Understanding and Controlling Longitudinal Exposure of Socially Shared Data. In *Proc. SOUPS*.

Mondal, M.; Yılmaz, G. S.; Hirsch, N.; Khan, M. T.; Tang, M.; Tran, C.; Kanich, C.; Ur, B.; and Zheleva, E. 2019. Moving Beyond Set-It-And-Forget-It Privacy Settings on Social Media. In *Proc. CCS*.

Orlowski, A. 2014. Google De-listing of BBC Article Broke UK and Euro Public Interest Laws. *The Register*.

Palen, L.; and Dourish, P. 2003. Unpacking "Privacy" for a Networked World. In *Proc. CHI*.

Papadopoulos, S.; Kompatsiaris, Y.; Vakali, A.; and Spyridonos, P. 2012. Community Detection in Social Media. *Data Mining and Knowledge Discovery* 24(3).

Petronio, S. 2002. *Boundaries of Privacy: Dialectics of Disclosure*. State University of New York Press.

Ramokapane, K. M.; Rashid, A.; and Such, J. 2017. "I Feel Stupid I Can't Delete...": A Study of Users' Cloud Deletion Practices and Coping Strategies. In *Proc. SOUPS*.

Redmiles, E. M.; Kross, S.; and Mazurek, M. L. 2017. Where is the Digital Divide? A Survey of Security, Privacy, and Socioeconomics. In *Proc. CHI*.

Sanchez, A. 2017. Facebook No Longer Allows You to Edit Images in Posts. *E-Marketing Associates Blog*.

Saunders, B.; Sim, J.; Kingstone, T.; Baker, S.; Waterfield, J.; Bartlam, B.; Burroughs, H.; and Jinks, C. 2018. Saturation in Qualitative Research: Exploring its Conceptualization and Operationalization. *Quality & Quantity* 52(4).

Schönberger, V. M. 2011. *Delete: The Virtue of Forgetting in the Digital Age*. Princeton University Press.

Sleeper, M.; Cranshaw, J.; Kelley, P. G.; Ur, B.; Acquisti, A.; Cranor, L. F.; and Sadeh, N. 2013. "I Read My Twitter the next Morning and Was Astonished": A Conversational Perspective on Twitter Regrets. In *Proc. CHI*.

Such, J. M.; and Criado, N. 2018. Multiparty Privacy in Social Media. *CACM* 61(8).

Such, J. M.; Porter, J.; Preibusch, S.; and Joinson, A. 2017. Photo Privacy Conflicts in Social Media: A Large-scale Empirical Study. In *Proc. CHI*.

Volkova, S.; and Bell, E. 2017. Identifying Effective Signals to Predict Deleted and Suspended Accounts on Twitter Across Languages. In *Proc. ICWSM*.

Wang, Y.; Leon, P. G.; Acquisti, A.; Cranor, L. F.; Forget, A.; and Sadeh, N. 2014. A Field Trial of Privacy Nudges for Facebook. In *Proc. CHI*.

Wang, Y.; Norcie, G.; Komanduri, S.; Acquisti, A.; Leon, P. G.; and Cranor, L. F. 2011. "I Regretted the Minute I Pressed Share": A Qualitative Study of Regrets on Facebook. In *Proc. SOUPS*.

Warren, T. 2017. WhatsApp Now Lets You Delete and Revoke Messages You Sent By Mistake. *The Verge*.

Wei, M.; Stamos, M.; Veys, S.; Reitinger, N.; Goodman, J.; Herman, M.; Filipczuk, D.; Weinshel, B.; Mazurek, M. L.; and Ur, B. 2020. What Twitter Knows: Characterizing Ad Targeting Practices, User Perceptions, and Ad Explanations Through Users' Own Twitter Data. In *Proc. USENIX Security Symposium*.

Wisniewski, P.; Islam, A.; Richter Lipford, H.; and Wilson, D. C. 2016. Framing and Measuring Multi-Dimensional Interpersonal Privacy Preferences of Social Networking Site Users. *CAIS* 38(1).

Wisniewski, P.; Lipford, H.; and Wilson, D. 2012. Fighting for My Space: Coping Mechanisms for SNS Boundary Regulation. In *Proc. CHI*.

Wisniewski, P.; Xu, H.; Lipford, H.; and Bello-Ogunu, E. 2015. Facebook Apps and Tagging: The Trade-Off Between Personal Privacy and Engaging With Friends. *JAIST* 66(9).

Wisniewski, P. J.; Knijnenburg, B. P.; and Richter Lipford, H. 2017. Making Privacy Personal: Profiling Social Network Users to Inform Privacy Education and Nudging. *IJHCS* 98.

Woodruff, A. 2014. Necessary, Unpleasant, and Disempowering: Reputation Management in the Internet Age. In *Proc. CHI*.

Zhao, X.; Salehi, N.; Naranjit, S.; Alwaalan, S.; Voida, S.; and Cosley, D. 2013. The Many Faces of Facebook: Experiencing Social Media As Performance, Exhibition, and Personal Archive. In *Proc. CHI*.

Zhou, L.; Wang, W.; and Chen, K. 2016. Tweet Properly: Analyzing Deleted Tweets to Understand and Identify Regrettable Ones. In *Proc. WWW*.