"I read my Twitter the next morning and was astonished" A Conversational Perspective on Twitter Regrets

Manya Sleeper*, Justin Cranshaw*, Patrick Gage Kelley†, Blase Ur*,
Alessandro Acquisti*, Lorrie Faith Cranor*, Norman Sadeh*

*Carnegie Mellon University	†University of New Mexico
{msleeper, jcransh, bur, acquisti, lorrie, sadeh}@cmu.edu	pgk@unm.edu

ABSTRACT

We present the results of an online survey of 1,221 Twitter users, comparing messages individuals regretted either saying during in-person conversations or posting on Twitter. Participants generally reported similar types of regrets in person and on Twitter. In particular, they often regretted messages that were critical of others. However, regretted messages that were cathartic/expressive or revealed too much information were reported at a higher rate for Twitter. Regretted messages on Twitter also reached broader audiences. In addition, we found that participants who posted on Twitter became aware of, and tried to repair, regret more slowly than those reporting in-person regrets. From this comparison of Twitter and in-person regrets, we provide preliminary ideas for tools to help Twitter users avoid and cope with regret.

Author Keywords

Twitter; regrets; messaging; conversation; survey

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

It is easy to say something you regret, angrily insulting a loved one or inadvertently letting a secret slip. However, Twitter, a social networking service, enables these types of regrettable messages to spread rapidly and broadly, and to remain available for extended periods of time. Twitter's ability to broadcast messages widely and retain them indefinitely potentially alters the dynamics of regretted communications. In extreme cases, Twitter has enabled highly-publicized instances of regret, like Rep. Anthony Weiner's infamous tweet that led to his resignation [5]. However, everyday Twitter use can lead to more mundane regrets. As in conversation, Twitter users insult others, accidentally reveal private information, and express emotion in heated moments.

Thus it is worthwhile to investigate regret both on Twitter and for in-person conversations. Past studies of in-person regret have identified factors that lead to regret, methods for becoming aware of regret, and strategies for repairing harm [8, 15, 16]. However, Twitter presents different features and limitations than offline conversation. Beyond offering wider audiences and increased message persistence, Twitter lacks face-to-face channels, such as body language, for transmitting apologies or indicating offense.

We explore regretted messages Twitter users posted on Twitter or said during in-person conversations. We aim to improve understanding of regrets on Twitter by comparing them with in-person regrets. By examining these regrets, as well as how people became aware of regrets in person and on Twitter, we also identify preliminary design directions for preventing and ameliorating regrets on Twitter.

Specifically, we examine four research questions:

- Q1: What states of being lead to regret on Twitter and in person?

- Q2: What types of regret occur on Twitter and in person?

- Q3: How do people become aware of regretted messages on Twitter versus in person?

- Q4: What repair strategies do people use to cope with regretted messages on Twitter and in person?

To address these questions, we ran a 1,221-participant online Mechanical Turk survey with two conditions. In one condition, we asked Twitter users to report on one message they regretted saying during an in-person conversation. In the other, we asked parallel questions about a message they regretted posting on Twitter. We collected information on the incident, the participant's emotional state preceding the incident, how the participant became aware of the regret, and any mitigation strategies employed. We used these answers to understand and compare drivers and consequences of regretted messages during in-person conversation and on Twitter.

BACKGROUND AND RELATED WORK

To demonstrate the conventions of posting on Twitter, we briefly highlight key features of the service. We then review related work on regret, first examining past analyses of regret during in-person conversations before discussing more recent work on social networking sites.

Twitter features

Twitter is an online social networking site where users post tweets, which are text-based messages of 140 characters or less. These messages are broadcast to a user's followers in relationships that are often asymmetric.

Twitter has several conventions that aid in sharing. Users can direct a message to a handful of specific users by crafting an @-reply. Users indicated by the @-reply will be alerted to the message through email or the Twitter client, but the message itself is public. A direct message (DM) allows a user to send a private message to a single person. A user can also add #hashtags to a tweet to categorize it, better enable searches as part of a trend, or provide contextual information. Tweets are publicly accessible unless an account is protected. Only a user's approved followers can view a protected user's tweets.

Related work

We first discuss research in the communications literature that has sought to understand many aspects of regret during in-person communication. We then outline work examining users' potentially regrettable behaviors and coping mechanisms on social networking sites.

In-person conversational regrets

Past studies of regrettable messages generally considered in-person messaging. Knapp et al. conducted 155 interviews, asking each participant about something they wished they hadn't said. The researchers noted eleven categories of regret, which we used in this study. They found that blunders, direct attacks, and group references were most frequently associated with regret, and that participants typically realized immediately when a message was regrettable [8].

Meyer surveyed 173 undergraduates about their cognitive states before and after saying regrettable things in person. Stress, frustration, and anger, as well as "having a lot on [their] mind," were most frequently associated with regrets. Participants commonly realized on their own that regrets had occurred, and regretted messages were rarely directed at more than one person [15]. Meyer separately examined efforts to repair the effects of in-person regretted messages. In a 204-participant survey, she found that nearly two-thirds of repair strategies involved apologies, while excuses and justifications were also common [16]. We drew heavily from these studies for our survey design.

McLaughlin et al. also examined regretted in-person messages as part of broader work on failure events. They evaluated concession, excuse, justification, refusal, and silence as failure-management strategies. Excuse was most commonly used, although increased guilt by the speaker tended to lead to concession [14]. We expect our work to align more closely with the narrower study of regretted in-person messages.

Like in-person conversation, Twitter is primarily focused on individual short messages, potentially with intended audiences. Thus, to explore the regrets that emerge in Twitter communications, rather than in person, we based our methodologies on previous in-person regrettable messaging work and results.

Regret in social media

Regretted tweets have not been studied extensively; however, Wang et al. performed a mixed-methods study of Facebook users' most regretted posts. These posts often contained sensitive or potentially offensive content, were created during highly emotional "hot states," or were seen by unintended audiences. The researchers also identified audience management and appearance management as major sources of potential regret [19]. While this work informed our analyses, differences between Facebook and Twitter usage patterns and audiences necessitated a different approach for investigating regrets on Twitter. Wang et al. looked at most-regretted incidents, a method often used to examine life regrets. We instead looked at unspecified regrets, a method used for in-person messaging regrets.

Although regretted messages have not been directly addressed for Twitter, a variety of factors have been investigated that could contribute to, or help ameliorate, regret on Twitter. Marwick and boyd found that Twitter users deal with "context collapse." Users fashion tweets that can simultaneously fit a variety of social contexts by tweeting to an "imagined audience," employing self-censorship, or aiming to balance authenticity with conscious identity management [13]. This tension between perceived and actual audiences, as well as the difficulty of balancing authenticity with self-censorship, may lead to regret on Twitter.

Lampinen et al. found that users adopted proactive coping strategies for managing the co-presence of diverse groups on Facebook [10]. Wisniewski et al. also examined interpersonal boundary management on social networks and found that users adopted ad-hoc boundary management mechanisms, such as ignoring information, blocking people, using aggressive behavior, or self-censoring [20].

Twitter presents a social environment with potential for regret. We seek to use methods from in-person messaging regrets research to understand the regrets that emerge, as well as the ad-hoc awareness and repair strategies used to address such regrets, by examining and comparing Twitter users' regrets on Twitter and from in-person conversations.

METHODOLOGY

Our goal was to analyze regrets that Twitter users had experienced on Twitter and during in-person conversations. We conducted a large-scale online survey from August to September 2012 using Amazon's Mechanical Turk (MTurk). We asked each of 1,221 MTurk Twitter users to describe one thing they had said and then later regretted (the regretted message) either during in-person conversation or on Twitter, depending on the condition to which the participant was assigned. We collected a description of the message, the context, how they became aware of the regret, and how they sought to repair the regret. It took participants 14.5 minutes on average to complete the survey, for which they were paid $0.75 (within the typical pay range for MTurk [6]).

Participant selection and conditions

We screened for US MTurk workers over 18 years old who self-reported English proficiency and relatively frequent

Twitter use (having had a Twitter account for at least a month and posting at least monthly). Of the 3,175 MTurk workers who started the survey, 946 did not meet these requirements. The majority (609) were disqualified for posting less than once a month on average.

Survey

Conditions
After the initial screening questions, participants were split into two conditions in a round-robin fashion. The first condition was conversational regret, which mirrored previously described work. The second condition asked parallel questions, slightly reworded to focus on Twitter regret. In both conditions, participants were asked to recall a time when they said or tweeted something and then regretted it, with the wording and format of the prompt based on Meyer's work on in-person messaging regrets [15, 16].

Our prompt for **conversational-regret** participants was:

"Please recall an occasion when you **said** something during an **in-person** conversation and then regretted saying it. This may be something that you regretted saying immediately or that you regretted saying later."

Our **Twitter-regret** prompt was similar:

"Please recall an occasion when you **tweeted** something and then regretted tweeting it. This may be something that you regretted tweeting immediately or that you regretted tweeting later."

Survey structure
Participants in both conditions who could not recall a regret were directed to an alternate survey that asked them about why they did not have regrets. We do not report the results of this survey, as the goal was only to ensure an equal workload for either positive or negative responses. Of the 1,879 participants who qualified for the study, 601 (456 for Twitter and 145 for conversational regret) could not recall regrets.

Participants who were able to recall regrets completed a survey about the regretted messages they reported in response to the initial prompt. The survey drew heavily on questions and structure from in-person messaging regrets work [8, 15, 16] and included several groups of related questions. We asked participants about the following:

Regretted message description: a series of essay questions that asked the participant to describe the message in detail, including the context, the reason why they said/tweeted it, the intended audience, the audience's reaction, why they regretted the message and any consequences

Circumstances: follow-up questions about their state when they delivered the message

Awareness: free response about how they became aware that they should not have said the message, followed by a multiple choice selection of how quickly after the message they realized they should regret the message

Repair strategies: a description of whether, how, and how successfully they tried to repair any harm caused by the message; participants were also asked to rate the seriousness of the regret before and after repair

Twitter specifics: questions on Twitter usage (e.g., client and device tweeted from, is/was the account protected)

Demographics: basic demographic questions

We based the general survey structure on the format used in previous work on in-person regrettable messaging [15, 16]. Specifically, we used Meyer's format of asking participants to provide one regret and then probing for details. Although this format has several weaknesses, as outlined in Limitations, it has been used repeatedly to examine in-person messaging regrets.

Quality control on Mechanical Turk
While MTurk has been shown to produce quality samples and results [6], surveys on MTurk should be designed to encourage quality responses. We took several quality control measures. First, we only used MTurk workers who had over a 95% approval rating on the site. Second, we front-loaded longer essay questions. By putting these questions earlier in the study, we encouraged lazy or unmotivated participants to drop out early or to enter nonsensical data where it was visible. It also made it easy for honest survey participants to return to the task, without feeling like they still needed to invest large amounts of time. We removed a small number of participants (25) from the dataset who provided nonsensical or non-English answers to the free response fields.

We also removed responses from 32 conversational-regret participants who responded about a regret on Twitter. We believe they did so because they were primed to think about Twitter when recruited as Twitter users. An additional 350 participants were removed for not completing the survey.

Data analysis
We surveyed MTurk users who posted on Twitter about a regretted message either said in-person or posted on Twitter. Although the surveys for each condition were designed to be parallel, the fundamentally different contexts preclude statistical comparisons between conditions. To explore characteristics of how regret on Twitter compares with in-person regret, we present the results of the Twitter- and conversational-regret conditions side-by-side. The proportions of participants reporting different answers are only meant to illuminate general themes and trends, not to be compared statistically.

Within a single condition, we perform statistical analyses. We use logistic regression to evaluate the relationship between types of regret and whether the audience was a group or individual, the relationship between awareness mechanisms and whether or not regret was experienced immediately, and the impact of repair strategy on the success level. Demographics were compared between conditions using a Wilcoxon test for numerical data and χ^2 tests for categorical data. All tests use a significance level of $\alpha = .05$.

Participant demographics

After quality-control removals, 1,221 people reported regrets: 747 for conversational regret (72% of those who started) and 474 for Twitter (41%). The mean age was 30.3 (28.2 for Twitter and 31.7 for conversational regrets). Overall, 53% of participants were female and 46% were male (10 preferred not to answer). The gender breakdown was almost identical for the Twitter- and conversational-regret conditions. Of the participants, 26% were students and 10% were unemployed. The remainder were primarily employed in science (9%), service (8%), and art (8%) occupations. There were no significant differences between the Twitter- and conversational-regret participants in age, gender or occupation, nor were there significant demographic differences between participants who did and did not report regrets.

ANALYSIS AND RESULTS

Q1: States of being leading to regret

States leading to regret

People often say things they later regret because of demands on mental capacity that impair thought processes. We found that both Twitter- and conversational-regret participants were often in negative, highly emotional states prior to regret. Meyer outlines several factors that contribute to "cognitive load," "physiological state," and "emotional state," which can potentially lead to regret [16]. We asked participants about these states. Based on Wang et al. [19], we also asked whether they were drunk at the time of the message. We asked participants to rate on a five-point scale how much or how little each factor applied immediately before they tweeted or spoke. A one indicated "Not at all" and a five indicated "Very much so." They rated each of the following: "I was fearful or frightened," "I had a lot on my mind," "I was feeling excited," "I felt ill," "I was worried," "I was nervous or anxious," "I was drunk," "I was angry," "I was stressed," "I was tired/fatigued," "I was happy," "I was hung over," and "I felt frustrated."

Consistent with Wang et al.'s work on Facebook regrets [19], we found that both in person and on Twitter, highly emotional negative states were most common prior to regret. Participants commonly reported a four or a five for stress (46% of Twitter and 50% of conversational participants), anger (51% and 43%), or frustration (58% and 53%) prior to the regrets. Participants also often had something on their minds (54% and 51%). Somewhat less common were positive emotions, including feeling excited (26% and 17%) or happy (22% and 21%).

Q2: Types of regret

We also looked at types of regrets participants reported for Twitter and for in-person conversations. In both conditions, participants most commonly reported regretting messages that were critical of others. However, on Twitter, participants more commonly regretted content that was expressive/cathartic and that was intended for groups of people.

Types of regret

We coded each regret described by participants into one of Knapp et al.'s categories for types of regretted in-person conversational messages [8], specifically:

Participant-Reported Types of Regret

	Twitter		Conversation	
Reveal too much	117	25%	105	14%
Direct criticism	96	20%	213	29%
Expressive	64	14%	15	2%
Direct attack	62	13%	108	14%
Blunder	51	11%	120	16%
Implied criticism	34	7%	84	11%
Group reference	13	3%	21	3%
Agreement changed	3	1%	10	1%
Behavior edict	2	0%	28	4%
Lie	1	0%	25	3%
Other	31	7%	18	2%

Table 1. Types of regret for Twitter and Conversation

- **Blunder:** "not normally perceived by a third-party observer as problematic"; mistakes, factual issues; includes typos or errors during conversation

- **Direct attack:** "critical statements directed at a person, the person's family, or the person's friends [...] general rather than specific"

- **Group reference:** stereotypical references about a group (e.g., ethnic, racial)

- **Direct criticism:** critical statements about "something specific" about a person

- **Reveal/explain too much:** telling "more than the situation calls for"; e.g., undesired personal information or a secret

- **Agreement changed:** agreeing to something, then later changing one's mind

- **Expressive/catharsis:** general "expressions of feeling and emotion"

- **Lie:** "knowingly lying to another person"

- **Implied criticism:** "critical remarks that are implicit" and can be "teasing remarks"

- **Behavioral edict:** telling someone to behave in a certain way

Two coders independently coded all the regrets based on Knapp et al.'s categories. Two coders reached a consensus for any regrets for which there were discrepancies.

Across both conversational and Twitter regrets, participants most commonly regretted critical statements (Table 1). Common critical statements included direct attacks and direct criticisms; 29% of conversational and 20% of Twitter regrets were direct criticisms, while 14% of conversational and 13% of Twitter regrets were direct attacks.

Blunders also arose frequently for both conversational and Twitter regrets, although more often for conversational (11% for Twitter, versus 16% for conversational). Although both Twitter- and conversational-regret participants reported some similar blunders, such as saying/posting messages they later

found out were false or that had been said/shown to someone who found them offensive, some blunders were unique to Twitter. On Twitter, time-delayed blunders sometimes caused participants to regret messages because of an event or change in context. For example, one participant regretted tweeting about a drive-by shooting in his friend's hometown when that friend was later killed in a drive-by shooting. Twitter, as an online interface, also allowed blunders caused by typos and broken links, which several participants found embarrassing. For example, one participant reported being "made fun of" for tweeting that he "used a lot of hags on [his] car."

Participants also regretted expressive or cathartic content more frequently on Twitter than in person (14% versus 2%). These expressive statements were typically tweeted when participants were angry or upset. They often served to vent or express frustration on topics such as work, relationships, or politics. Often, the goal was to allow others to sympathize or "know what [the participant] was going through." Participants tended to regret the message later after re-thinking how it would sound, or after someone who viewed it became upset. For example, one participant described tweeting "Last day of my internship, so excited to be done," because she "was unhappy with how the internship treated [her] and what had happened [...and] wanted [her] friends to see it because they knew [she] was having a rough time." However, she regretted the tweet when her internship coordinators saw it and sent her an email telling her she needed to delete the tweet. In contrast, expressive regrets during in-person conversations tended to be part of arguments or opinions.

Type and audience
Participants also specified whether they intended the messages to be seen or heard by individuals, or by multiple people. Twitter-regret participants were more likely to target multiple people (73% of Twitter regrets, versus 24% of conversational), likely because of Twitter's broadcast capabilities.

Certain types of regretted messages were more frequently intended for multiple people, especially on Twitter. When the intended audience comprised multiple people, rather than an individual, Twitter-regret participants were significantly more likely to report a blunder ($p = 0.008$), content that revealed too much ($p = 0.005$), or expressive/cathartic content ($p = 0.003$). Of Twitter blunders, 82% were intended for multiple people, versus 33% of reported in-person blunders. Twitter-regret participants often said that they wanted to tweet to friends, coworkers, or others interested in a specific topic, but regretted the tweet because they made an error that caused confusion or made them look bad. For example, one participant reported tweeting, "Congratulations to B for being elected ALA Councilor," intending the message for other librarians in South Carolina. She later realized that the individual was actually a candidate for the position, rather than having been elected, and regretted the tweet because "it was embarrassing."

Twitter-regret participants who regretted expressive or cathartic posts also tended to target multiple people rather than an individual (84% of expressive/cathartic regrets). Participants often hoped to share political or negative feelings with the general public or their friends because they "wanted to vent" or express their feelings "to anyone that would listen."

Regretted statements on Twitter that revealed too much also tended to be targeted at multiple people (80%). Many participants tweeted personal information, such as details about their lives or relationships, and then regretted sharing them on Twitter. Several participants also reported having both personal and professional accounts and regretting tweeting personal information on their professional Twitter accounts. For example, one participant said that he regretted tweeting "on my professional twitter account about a night of heavy drinking" because it seemed "unprofessional."

In contrast, conversational-regret participants were significantly more likely to report regrets that were direct attacks ($p = 0.024$) when the intended audiences were individuals (67%) rather than multiple people. Participants were typically angry or arguing with the recipient of the message. For example, one participant "screamed at my father that 'I hate him' in an argument" because his father kept him from attending a party. On Twitter, such attacks were commonly focused at groups (68%), and participants reported wanting their anger to be seen. For example, one participant had a conflict with a friend, and wrote "she's so annoying and whiny," intending "it to be seen by friends."

Unintended audience
We also coded for regretted messages having unintended audiences. In conversation, unintended audiences included people overhearing messages (e.g., by walking into a room) or being told about them. On Twitter, most of the tweets reported were public tweets. However, participants still had particular audiences in mind when they tweeted. Unintended audiences occurred because people other than the intended audiences saw or heard about the tweets.

For Twitter regrets, 13% had unintended audiences, compared to 5% of in-person regrets. Unintended audiences occurred most commonly on Twitter for regrets that revealed too much (23% of regrets that revealed too much), often because participants tweeted something private, insulting, or about work, which they later realized they didn't want everyone to know. For example, one participant described how she tweeted "something sexual and my [T]witter at the time was public, so I freaked out when I saw that my brother's screen name popped up on Recommended Twitter."

Level of regret
To measure level of regret, we asked participants "In your opinion, how serious of a problem was it that you said the messages, at the time you said it" (or tweeted it), based on a question from [15]. Participants responded from one ("Not at all") to five ("Very much so"). We consider participants who reported a four or a five to have had a high level of seriousness and below a four to have had a low level.

For Twitter, 18% of messages had high levels of seriousness. For conversational regrets, 38% had high levels of seriousness. However, the interpretation of the difference is somewhat ambiguous; the seriousness of regrets across contexts

Descriptions of Means of Awareness

Self realization	The individual realizes either by thinking about it or by just feeling bad that they should regret the message
Audience says something	The intended audience says something to imply that the person should regret the message
Audience takes an action	The intended audience does something to imply that the person should regret the message (e.g., stops speaking to the individual)
Audience body language	The individual realizes they should regret the message based on the intended audience's body language (e.g., smile, frown)
Third party says something	A person other than the intended audience says something to imply that the person should regret the message
Third party action	A person other than the intended audience does something to imply that the person should regret the message
Third party body language	A person other than the intended audience uses body language to imply that the person should regret the message

Table 2. Codes for means of awareness

Participant-Reported Means of Awareness

	Twitter		Conversation	
Self realization	58%	275	39%	294
Audience said	29%	138	17%	126
Audience action	7%	32	26%	191
Audience body lang	0%	1	19%	143
3rd party said	7%	33	5%	39
3rd party action	1%	5	1%	8
3rd party body lang	0%	1	0%	3
Other	1%	6	0%	3
Total		474		747

Table 3. Means of awareness for Twitter and Conversation

may not be directly comparable. For instance, a serious conversational regret may differ from one on Twitter.

Q3: Awareness of regret

Individuals must become aware of regrets to address them. Conversational-regret participants tended to become aware of regret more quickly and relied more on audience actions, such as body-language cues. Twitter participants more often reported realizing regrets themselves or had audience members tell them they should regret the message.

Means of awareness

We asked each participant to describe in a free response how they became "aware [they] shouldn't have said the message." Two coders created a set of codes for means of awareness based on types of awareness outlined in Meyer's work on regretted messaging [15] using a set of 100 regrets (Table 2). The same two coders then independently coded the regrets based on these codes. A third coder also independently coded the regrets to break ties. In cases where all three coders disagreed, two coders reached a consensus. A regret could be coded for multiple, different means of awareness.

Participants became aware of regret using different means on Twitter and in person (Table 3). This is partially explained by the different contexts for Twitter and conversational regret. Audience body language is usually immediately available in person but typically absent on Twitter. Thus, 19% of conversational-regret participants described using audience body language to become aware of regret. Participants often realized the regret immediately when they saw their audiences' facial expressions. For example, one participant reported calling "his cousin an asshole in-front of our entire family" and realized he should regret it "[w]hen everyone glared at me."

Conversational-regret participants were also more likely to report relying on audience actions to become aware of regret (26% for conversation, versus 7% for Twitter), also likely due to the intended audience's physical presence. Such actions included storming out of a room, laughter, or sitting silently, which are difficult to convey over Twitter. Offline followups to Twitter messages, such as job termination or laughter, led to awareness for Twitter regrets, as did Twitter-specific online actions, such as being unfollowed or ignored.

Comparatively, Twitter-regret participants more frequently became aware of regret on their own (58%, versus 39% for conversational regrets). Participants in both conditions would often realize that the regretted message was something that they should not have said or tweeted, either after thinking about it or because they felt bad. As one participant put it: "Something inside just told me it was wrong." However, on Twitter, messages also remain available over time. Several Twitter-regret participants reported re-reading the message later and realizing that they should regret it, an option that is rarely available in person. For example, one participant tweeted, "Absolutely pointless," about her relationship and realized she should regret it when she "read over [her] tweets the next morning and thought it was dumb."

Twitter-regret participants were also more likely to report that their intended audience said something to imply that they should regret the message (29% of Twitter, versus 17% of conversational). This may partly reflect the wider audiences targeted by Twitter users but also how, on Twitter, people helped participants realize they should regret a message. Often, a friend or co-worker saw the message and contacted the participant to tell them that they should regret it. For example, one participant tweeted "Having fun on my day off. #callinginsick" and realized he should regret it when "[o]ne of [his] friends told [him] it wasn't a good idea."

Time until awareness

Conversational-regret participants also became aware of regrets more quickly than participants on Twitter. Based on wording used by Meyer [15], we asked participants "how much time passed between" when they tweeted or spoke and when they became aware they shouldn't have tweeted or said the message. We found that the majority of conversational respondents became aware immediately (62%), with many of the remaining participants becoming aware within a few minutes (18%). Of the remaining 20%, the majority became aware the same day or the next day (13%). On Twitter, participants reported taking longer. Only 11% were immediately aware, while 29% realized within a few minutes, 33% at some point the same day, and 16% the next day. The majority of the remaining 11% became aware of the regret within a few days.

For some types of awareness, participants were more or less likely to become aware immediately. On Twitter, participants were significantly less likely ($p = 0.028$) to become aware of the regret immediately (4%), rather than later, when the audience said something to imply that they should regret the tweet. This is consistent with users tweeting and audience members later informing them that they should regret the content, implying a time delay. For conversational regrets, participants were significantly more likely ($p < 0.001$) to learn immediately (84%) from audience body language about a regret. They often reported realizing as soon as they spoke that they should regret the message due to the audience's physical reactions. As one participant reported, "The moment it slipped out, I knew I shouldn't have. The awkward looks and silence that followed confirmed that it was as bad as it sounded." In contrast, conversational-regret respondents were significantly less likely ($p < 0.001$) to become aware immediately (13%) when a third party told them something to imply that they should regret the message. The person about whom they were talking, or who was impacted by the message, often contacted them, delaying awareness. For example, one participant "told a coworker that I intended to leave my job in an open area" and regretted it "[w]hen I went to meet with my boss she told me she had heard rumors."

Q4: Repair strategies

After becoming aware of a regretted message, people often employ strategies to repair the impact, or potential impact, of the message. We asked participants about the repair strategies they used after tweeting or saying the messages, as well as the impact of these repair strategies. We found that conversational-regret participants most often chose to apologize, while Twitter-regret participants most often chose to delete regretted tweets. As occurred in regret awareness, Twitter-regret participants also took longer to repair regrets than conversational-regret participants.

Frequency of repair strategy

We asked each participant to select repair strategies they used from a list taken directly from the conversational-regrets literature [16]. Participants in both conditions were provided with the options: "I tried to say something to offset the harm done," "I tried to justify or defend what I said to minimize its offensiveness," "I apologized for saying it," "I just acted

Participant-Reported Repair Strategies

	Unsuccessful		Successful	
	Twi.	Conv.	Twi.	Conv.
Delete	111	–	134	–
Apology	53	173	72	218
Act like nothing hppnd.	44	70	38	42
Excuse	36	92	34	55
Justify	38	89	30	64
Say something to offset	17	77	22	67
Deny	10	50	10	31
Non-verbal behavior	–	40	–	30
Other	11	21	5	21
Apology and delete	30	–	38	–
Apology and justify	15	49	16	43
Apology and offset	5	52	12	45
Apology non verbal	–	25	–	19
Total (participants)	191	329	196	302

Table 4. Repair strategies for Twitter and Conversation

like nothing had happened," "I denied or tried to take back what I said," "I offered an excuse for why I said it," "I didn't do anything." Conversational-regret participants were also offered the option "I employed a nonverbal behavior to indicate that I regretted it" (from the regrets literature), while Twitter participants were offered "I deleted the tweet."

Overall, we found that a similar proportion of Twitter- and conversational-regret participants took actions (did not report doing nothing) to repair regrets (82% and 84%, respectively). However, the distribution of repair strategies varied (Table 4). Conversational-regret participants most frequently chose to apologize (34% of strategies). Twitter-regret participants most often chose to delete regretted tweets (37%), an option unavailable in person. Both conversational and Twitter participants were relatively likely to try to make an excuse (11% of Twitter and 13% of conversational strategies), justify their messages (10% and 13%), and act like nothing had happened (12% and 10%). However, conversational participants were more likely to try to say something to offset the harm (12%, versus 6% for Twitter).

Success of repair strategies

These different repair strategies also met with varied levels of success (Table 4). Participants rated, on a five-point Likert scale, how successful or unsuccessful their repair strategies were. Participants who ranked their strategies as "successful" or "very successful" were categorized as having successfully repaired the regret. Approximately half of each of Twitter- and conversational-regret participants who took repair actions were successful. Controlling for seriousness of regret at the time of the message, several repair strategies emerged as significantly more likely to be successful or unsuccessful.

On both Twitter and in conversation, using an apology significantly increased the probability of success ($p = 0.043$ and $p < 0.001$ respectively). In person, making an excuse significantly decreased the probability of success ($p = 0.002$),

while on Twitter, deleting the tweet significantly increased the probability of successful repair ($p = 0.038$).

Participants who apologized on Twitter varied in their use of online and offline apologies. Online, they apologized using a variety of means, including tweets, instant messages, and text messages. Offline, they apologized face-to-face or by calling impacted individuals. This choice of online or offline strategy seemed to depend on level of personalization and context. Several participants chose to apologize offline because they were confronted about a regretted tweet in an offline environment. For example, one participant apologized when his tennis coach confronted him about an insulting tweet and and told the coach that he "would delete the tweet immediately." Other participants reported apologizing in person to make the apology more personal, writing, "It was personal," so "I called them personally."

Twitter is often a relatively public forum, and, as the regretted tweets often reached wide audiences, apologizing online could also allow participants to reach larger audiences. Participants reported using online apologies to add additional information to their original tweets or add corrections. For example, one participant described accidentally posting misinformation about an animal rescue. After realizing her mistake, she tweeted a correction and an apology. Online apologies were also used to reach large groups of people. One participant described how she "tweeted back so everyone could see my apology and called the person" that she had upset.

Apologies after regretted tweets were also often paired with other online actions. Of the regretted tweets participants apologized for, 54% were also deleted. After posting "something passive-aggressive about someone," one participant described how she tried to repair the situation by telling her "friend that I'd acted immaturely and that I was sorry." She also "deleted the tweet because [I] was embarrassed by my actions."

For in-person regrets, apologies tended to be offline and verbal, often face-to-face to a single person involved with the regret. For instance, one participant jokingly "insulted a friend only to find out his mother had passed away earlier in the week and hadn't told anyone." Once he found out, the participant "immediately apologized stating that [he] didn't know and offered [his] condolences." Such apologies were often paired with justifications (23% of conversational apologies) or explanations that tried to offset the harm (25%). One participant described criticizing how her husband had done the household chores. She explained that she "apologized, and I think maybe explained that I hadn't meant to sound as rude and critical as it sounded. I also thanked my husband for the work he had done and said that I was glad he was so helpful."

Time to repair
Varied amounts of time passed before participants addressed the regretted messages. Participants responded in free-text to "When did you take these actions?" Two coders coded responses for all participants who used repair strategies other than acting like nothing had happened (1127 participants), based on the indication of the first repair. The coders reached a consensus on any disagreements. The categories were: Im-

mediately/a few minutes after the regret (15 minutes or less), the same day, the next day, more than a day but less than a week, more than a week but less than a month, and one month or more. For 32 participants (29 for Twitter and 3 for conversation), the time period was unclear.

Conversational-regret participants tended to respond more quickly, as might be expected because they also become aware of the regret more quickly. Of conversational-regret participants who actively tried to repair their regrets, 392 (67%) did so within a few minutes. The majority of the remainder did so the same day (78 participants, 13%) or the next day (49 participants, 8%). Alternatively, only 98 Twitter-regret participants (26%) who actively tried to repair their regrets did so within minutes; 131 (34%) tried to do so the same day, and 74 (19%) did so the next day. The majority of the remaining 10% took less than a week.

LIMITATIONS
There are limitations in our study design. We performed this study using Mechanical Turk. Although this potentially biases our sample, MTurk's population biases have been documented [18]. Samples and results from MTurk workers have also proven comparable to other online sources [6, 7]. We also took several measures to ensure quality responses. However, such quality control measures may also have biased our participant pool, potentially electing for more diligent or intelligent workers. It is unclear how this impact might differ from quality-control measures used for other survey methodologies. However, previous conversational-regrets work drew from an undergraduate population [15, 16]; using MTurk allowed us to expand to a large, cost-effective sample relative to offline pools or alternative online sources.

Our survey design had additional, inherent limitations. We used the basic design from the conversational-regrets literature [15, 16] in which each participant recalled a single, regretted message. Thus, we don't have a true analysis of the frequency of different types of conversational or Twitter regrets. Based on the conversational-regrets design, we asked participants for the regret that first came to mind, rather than the most recent or strongest regrets. However, certain regrets may come to mind more easily or may be more or less embarrassing to detail in a survey. Thus, we may have an overrepresentation of memorable regrets and an underrepresentation of deeply shameful regrets.

The survey format was also a limitation. We asked participants for self-reported, recalled data. Participants may attribute more meaning to events occurring in the past when reporting on them in a survey. There was also potential for reverse causality issues. We tried to limit causality questions, but participants may have attributed factors like states of being to the regret, when they were actually caused by the regret. We could offer more conclusive results if we tracked participant behavior over time and noted actions, like repair strategies, as they occurred. For example, a diary-study approach could be used to supplement this work.

DISCUSSION

We found that Twitter- and conversational-regret participants differed in the types of messaging regrets they reported, how they became aware of the regrets, and how they tried to repair the harm caused by the regrets. Time delays on Twitter, as well as lack of face-to-face communication with audiences, also caused awareness and repair on Twitter to occur more slowly than for conversational regrets. Based on these findings, we offer several early potential design directions for helping users prevent and repair Twitter regret.

Detecting and preventing regret on Twitter

Although our participants took measures to repair harm caused by the regretted messages, they often would have liked not to have tweeted the messages. One way to potentially prevent regret on Twitter would be to develop tools to detect potentially regrettable messages and provide users with suggestions for when they might want to reconsider tweeting. Behavioral economics offers a potential direction to help prevent users from sending such tweets by using behavioral "nudges" to help people identify tweets they might not want to post [1, 4]. Such nudges are cues that suggest that users should alter a behavior without forcing them to do so.

We found that several negative emotions, including anger, stress, and frustration, tended to lead to regret on Twitter. A recent study of deleted tweets also found a slightly higher frequency of negative-sentiment keywords in tweets that were deleted [3], a common strategy for coping with regretted tweets. Prior to a tweet being sent, such negative states could potentially be detected using tools like sentiment analysis or word frequency. Word analyses could potentially also be combined with environmental cues, such as location, especially when users tweeted from mobile devices; 45% of regrets reported by Twitter-regret participants were made from mobile devices. Once a negative mood was detected, it might be possible to provide feedback to the user about the negative emotion, or, in a manner similar to Google Mail Goggles [17], lock them out until they could think more clearly.

We also found that certain types of regret related to broadcasting thoughts to wide audiences were more common on Twitter. Twitter-regret participants tended to report regretting revealing too much, revealing expressive/cathartic thoughts, and sharing with unintended audiences. Such types of regret might be preventable through better audience awareness or management on Twitter. Participants often regretted tweets that revealed too much or that were expressive/cathartic because they were seen by people they didn't want to see them, or because people saw the tweets and were hurt. For these regrets, it might be possible to indicate more clearly who might view a tweet, for example by showing images of a user's followers. Interestingly, several tweets were sent by participants who had protected accounts at the time of the regretted message (25% overall, and 21% for unintended audience). Participants tended not to accidentally tweet to the general public. Rather, their tweets were viewed by people they didn't initially anticipate would view the posts. This is in line with Acquisti and Gross' concept of "imagined communities" [2] and the concept of tweeting to an "imagined" audience [13]. One

way to visualize the actual audience might be to show images of people who could view the tweet, potentially prioritizing by interaction level. For instance, Lieberman and Miller's Facemail prototype uses this approach for email [11].

Promoting regret awareness

To address a regretted message, users must first realize that they should regret the tweet. We saw several methods for becoming aware of regretted messages that were unique to in-person conversation and could potentially be adapted for Twitter, as well as several techniques that were unique to Twitter and could be further emphasized.

In person, participants often quickly became aware of regretted messages, typically through physical cues. For instance, one conversational-regret participant experienced regret after his girlfriend "instantly became upset and started to cry." Other participants saw audiences storm out of the room or laugh. Twitter users, physically separated from their audiences, usually lack instant audience feedback.

One possibility for improving Twitters users' awareness of regret would be to improve their abilities to gauge potential audience reaction absent physical feedback. Work has been performed to visualize sentiment conveyed in electronic communications. For example, Liu et al. prototyped an "EmpathyBuddy" for email that presents a line-drawn face that reacts to the emotion in the text [12]. Similar visualizations showing the sentiment conveyed by tweets might help Twitter users more quickly become aware of potentially regrettable tweets before tweeting them. A visualization that persisted after a user tweeted might also allow awareness to occur more quickly after a tweet.

We also found that Twitter-regret participants often reported being informed by their communities (e.g., friends, family, and co-workers) that they should regret messages, often over electronic means like text messages, or on Twitter itself. Lampinen et al. discussed how users of social networks collaboratively control disclosure [9]. Their participants used collaborative strategies to protect each others' privacy. Similarly, other individuals helped our participants become aware of regretted content. In some cases, these individuals were impacted by the message. In other cases, they were not. Developing easy mechanisms for people to tell someone about potentially regrettable tweets could mitigate potential regret.

Throughout our results, we saw that Twitter had a time delay compared to conversation, both in terms of time to awareness and time to repair. This was somewhat due to the lack of immediate audience feedback; in cases where Twitter regret was informed by others, this response often came hours or days later. On Twitter, users cannot typically see immediate feedback, and audiences sometimes cannot immediately access messages, delaying regret awareness and potential repair. However, unique to Twitter, even when there was no negative reaction, participants regretted tweets because of the record provided by Twitter. Participants re-read their tweets and realized the message was regrettable. Creating tools that better help users review past tweets may also help them become aware of, and purge, possibly regrettable content.

CONCLUSION

We examined Twitter users' regrets for in-person conversations and on Twitter. We found that, on Twitter, participants tended to report regretted messages targeted at broad audiences, including messages intended to be expressive or cathartic, that revealed too much, or that reached unintended audiences. In general, we also saw that Twitter-regret participants became aware of regret more slowly than conversational-regret participants, more often relying on others to tell them about the regret or eventually realizing themselves that the message should be regretted in the absence of physical audience cues. Once aware of the regret, Twitter users tended to delete the regretted tweet and/or apologize. Based on the findings, we offer several early design suggestions, including behavioral nudges for helping Twitter users realize potentially regrettable posts either before or after tweeting and for better audience management.

REFERENCES

1. Acquisti, A. Nudging privacy: The behavioral economics of personal information. *IEEE Security & Privacy 7*, 6 (Nov./Dec. 2009), 82–85.

2. Acquisti, A., and Gross, R. Imagined communities: Awareness, information sharing, and privacy on the Facebook. In *Privacy Enhancing Technologies*, Springer (2006), 36–58.

3. Almuhimedi, H., Wilson, S., Liu, B., Sadeh, N., and Acquisti, A. I wish I hadn't tweeted that! Large-scale quantitative analysis of deleted tweets. In *Proc. CSCW 2013*, ACM (2013).

4. Balebako, R., Leon, P., Almuhimedi, H., Kelley, P. G., Mugan, J., Acquisti, A., Cranor, L. F., and Sadeh, N. Nudging users towards privacy on mobile devices. CHIPINC 2011 (2011).

5. Bosker, B. The Twitter typo that exposed Anthony Weiner, 2011. `http://www.huffingtonpost.com/2011/06/07/anthony-weiner-twitter-dm_n_872590.html`.

6. Buhrmester, M., Kwang, T., and Gosling, S. D. Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science 6*, 1 (2011), 3–5.

7. Jakobsson, M. Experimenting on Mechanical Turk: 5 how tos. `http://blogs.parc.com/blog/2009/07/experimenting-on-mechanical-turk-5-how-tos/`, July 2009.

8. Knapp, M. L., Stafford, L., and Daly, J. A. Regrettable messages: Things people wish they hadn't said. *Journal of Communication 36*, 4 (1986), 40–58.

9. Lampinen, A., Lehtinen, V., Lehmuskallio, A., and Tamminen, S. We're in it together: Interpersonal management of disclosure in social network services. In *Proc. CHI 2011*, ACM (2011), 3217–3226.

10. Lampinen, A., Tamminen, S., and Oulasvirta, A. All my people right here, right now: Management of group co-presence on a social networking site. In *Proc. GROUP 2009*, ACM (2009), 281–290.

11. Lieberman, E., and Miller, R. C. Facemail: Showing faces of recipients to prevent misdirected email. In *Proc. SOUPS 2007*, ACM (2007), 122–131.

12. Liu, H., Lieberman, H., and Selker, T. Automatic affective feedback in an email browser. Tech. rep., MIT Media Laboratory Software Agents Group, 2002.

13. Marwick, A. E., and boyd, d. I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society 13*, 1 (Feb. 2011), 114–133.

14. McLaughlin, M. L., Cody, M. J., and O'Hair, H. D. The management of failure events: Some contextual determinants of accounting behavior. *Human Communication Research 9*, 3 (1983), 208–224.

15. Meyer, J. R. Regretted messages: Cognitive antecedents and post hoc reflection. *Journal of Language and Social Psychology 30*, 4 (2011), 376–395.

16. Meyer, J. R., and Rothenberg, K. Repairing regretted messages: Effects of emotional state, relationship type, and seriousness of offense. *Communication Research Reports 21*, 4 (2004), 348–356.

17. Perlow, J. New in labs: Stop sending mail you later regret, 2008. Official Gmail Blog. `http://gmailblog.blogspot.com/2008/10/new-in-labs-stop-sending-mail-you-later.html`.

18. Ross, J., Irani, L., Silberman, M. S., Zaldivar, A., and Tomlinson, B. Who are the crowdworkers?: Shifting demographics in Mechanical Turk. In *Ext. Abstracts CHI 2010*, ACM (2010), 2863–2872.

19. Wang, Y., Norcie, G., Komanduri, S., Acquisti, A., Leon, P. G., and Cranor, L. F. "I regretted the minute I pressed share": A qualitative study of regrets on Facebook. In *Proc. SOUPS 2011*, ACM (2011).

20. Wisniewski, P., Lipford, H., and Wilson, D. Fighting for my space: Coping mechanisms for SNS boundary regulation. In *Proc. CHI 2012*, ACM (2012), 609–618.