Could This Be True? I Think So! Expressed Uncertainty in Online Rumoring

Kate Starbird+, Emma Spiro*^, Isabelle Edwards*, Kaitlyn Zhou+, Jim Maddock+, Sindu Narasimhan+

HCDE+, Information School*, Department of Sociology^

University of Washington, Seattle, WA

{kstarbi, espiro, iare, katezhou, maddock, sindu} @uw.edu

ABSTRACT

Rumors are regular features of crisis events due to the extreme uncertainty and lack of information that often characterizes these settings. Despite recent research that explores rumoring during crisis events on social media platforms, limited work has focused explicitly on how individuals and groups express uncertainty. Here we develop and apply a flexible typology for types of expressed uncertainty. By applying our framework across six rumors from two crisis events we demonstrate the role of uncertainty in the collective sensemaking process that occurs during crisis events.

Author Keywords

Rumoring; Twitter; social media; crisis informatics.

ACM Classification Keywords

H.5.3 Information Interfaces & Presentation: Groups & Organization Interfaces: Collaborative computing, Computer-supported cooperative work; K.4.2 Social Issues

INTRODUCTION

Rumors are a regular feature of crisis events [3,1]. Crisis contexts are characterized by extreme uncertainty and lack of information, conditions that lend themselves to the emergence of rumors [26]. During these non-routine situations, populations engage in collective sensemaking as individuals attempt to understand their environment, and uncertainty expressed throughout deliberation represents a key mechanism in this process. Today many of these processes occur on social media platforms, offering disaster scholars the opportunity to expand understanding of rumoring behavior [30,27,35]. However, despite the growing number of studies on rumoring during crisis events, limited efforts explicitly address expressed uncertainty. We aim to fill this gap by unpacking ideas of expressed uncertainty and demonstrating how they represent important dimensions of rumoring behavior on social media during times of crisis.

Our work is part of a larger project investigating rumoring behaviors on social media during crisis events. Using a mixture of qualitative and quantitative methods we identify, code, and analyze rumor-related tweets to understand how rumors grow, spread, change and are refuted on Twitter. Though our research project began by focusing on misinformation [29], in this study we address rumors more broadly, investigating emergent stories that have some uncertainty or dissonance in relation to a central narrative.

In the work that follows we identify and describe types of expressed uncertainty in social media posts on Twitter. We do so at different points in a rumor's lifecycle, across different rumors, and within multiple crisis events. Our analysis reveals specific phrases and linguistic patterns that appear in rumor-related posts that contain uncertainty. Through mixed inquiry methods we explore the meaning of these patterns and illustrate how they contribute to collective sensemaking.

BACKGROUND

Rumoring during Crisis Events

Researchers continue to debate the definition of rumor [22], though many conceptualize rumoring as a social process whereby information spreads in a population. Shibutani [26] argues that rumors are a tool for collective problem-solving when groups make sense of the uncertainties of their environment, and Rosnow [24] contends that rumors are public communications that reflect individual beliefs. Rumor scholars also link rumor definitions with credibility and evidence; for example, Allport and Postman [1] define rumors as propositions that pass from one person to the next without standards of supporting evidence. Likewise, rumors can represent conclusions based on unverified information that attempt to make sense of uncertain situations [37]. DiFonzo and Bordia [7] define rumor as "unverified and instrumentally relevant information statements in circulation that arise in context of ambiguity, danger or potential threat, and that function to help people make sense and manage risk" [7, p. 13]. Scholars agree that rumors emerge during situations characterized by high levels of uncertainty and anxiety where information (especially from formal sources) may be unavailable or insufficiently timely [3,24]. Notably, though the term rumor often implies false

ACM 978-1-4503-3362-7/16/05…$15.00
DOI: http://dx.doi.org/10.1145/2858036.2858551

information, none of the above definitions incorporate ideas of truth (or untruth). Indeed, because of the inherent uncertainty that surrounds rumors at the time of communication, rumors can in fact turn out to be true [27].

Given these characteristics of rumors and rumoring behavior, it is easy to see why crisis contexts have been a setting for rumor research. Preconditions for rumoring (high levels of anxiety and uncertainty) are core elements of post-disasters settings. Scholars have extended many theories of rumoring during crisis into online environments, where communication patterns and information diffusion during times of crisis are more visible. Studies of rumoring during crisis events on social media have studied processes of information production, distribution, and organization [30,4], citizen reporting and distributed volunteer efforts [33,31], participation of official organizations [12,2], and patterns of serial transmission of messages [35].

While rumoring behavior can alleviate anxiety and aid sensemaking, it can also be dangerous during event responses, leading people to potentially life-threatening decisions. Not surprisingly, emergency responders often view rumors and misinformation as a threat to be managed [14, 34]. In the context of social media use during disasters, emergency responders fear that social media platforms may amplify the effects of misinformation spread [15, 13].

Examining Uncertainty in Rumoring during Crises

Many researchers argue that *uncertainty* in the information environment [7,10] contributes to rumoring. Like rumor, conceptualizations of uncertainty can be applied in different ways. In research on managing uncertainty in illness, Brashers [8] extends work by Babrow and colleagues [5,6] to explain that "uncertainty exists when details of situations are ambiguous, complex, unpredictable, or probabilistic; when information is unavailable or inconsistent; and when people feel insecure in their own state of knowledge or the state of knowledge in general" (p. 478). When uncertainty represents potential danger, people actively engage in information seeking, which can lead to a reduction in uncertainty. Information seeking can also increase uncertainty, especially when sources are inconsistent or contradictory [8]. In the context of crisis, uncertainty can be viewed as part of a crowd communication and "sense-making" process. Weick & Sutcliffe [38] explain this as "an ongoing process of making sense of the circumstances in which people collectively find [themselves] and of the events that affect them." In other words, informal communication and rumoring are precisely the social mechanisms that allow expressions of uncertainty.

As research examining rumoring behavior within social media grows, few studies have included expressed uncertainty in their explorations. Bordia and DiFonza [7] studied the propagation of 14 Internet rumors using content analysis, categorizing each post with one or more of 14 codes. Several of their categories encompass some level of uncertainty, but their coding scheme does not focus specifically on uncertainty. Oh et al. [20,21] analyzed tweets from several crises through the lens of rumor theory, investigating the relationship of anxiety and informational certainty, and measuring their interactive dynamics with both quantitative and qualitative methods. Though Oh et al. examine content ambiguity, they do not single out expressed uncertainty as a specific dimension for analysis.

Starbird et al. [29] and Maddock et al. [16] examined the propagation of misinformation online in the aftermath of the Boston Marathon Bombings, identifying several salient features of rumor propagation. They utilized a one-dimensional coding scheme (speculation, misinformation, hedge, question, correction) where uncertainty was a factor in several codes, but not distinguished as a separate quality of information. Codes that encompassed uncertainty—speculation, hedge, and question—were limited by coder agreement, perhaps due to the ambiguous nature of uncertainty. Zhao et al. [39] also focused on uncertainty, or more specifically language that expresses skepticism (e.g. *"Is this true?"*, *"Really?"*), as indicative of rumor-related content. While their work contributes methods for early rumor detection, it does not go beyond identification to unpack the behavioral and social mechanisms behind this phenomenon. Limited work applies concepts of expressed uncertainty in rumor-related content across multiple rumors and, importantly, across multiple events. Moreover, none of these studies isolate uncertainty as a unique characteristic of shared information, at both textual and rumor levels.

METHODS

In this paper, we focus on "expressed uncertainty" in social media posts—i.e. explicit, linguistic articulations of uncertainty about the veracity of the information contained. We seek both to understand how uncertainty is expressed at the post level, and how uncertainty manifests across the lifecycle of different rumors.

Events and Data Collections

This study examines six rumors from two crisis events.

Event 1: The Boston Marathon Bombings

The first event was the Boston Marathon Bombings, which took place on April 15, 2013. Two bombs detonated near the finish line resulted in three fatalities and 260 injuries. A manhunt followed, with FBI releasing photos of the suspected bombers on April 18. From this event we identified three rumors that spread through Twitter (described in the findings). We collected data using the Twitter Streaming API to track several event-related terms including `boston, bomb, explosion, marathon`, and `blast`. Data collection began April 15 at 5:25pm EDT and ended April 21 at 5:09pm EDT. At several points data collection was rate-limited at 50 tweets per second, and we experienced two brief outages where no data was collected. The final dataset included about 10.6 million tweets.

Event 2: Sydney Siege

The second event occurred between December 15th and 16th, 2014 when a gunman took 18 customers and employees hostages in a Lindt chocolate café in Sydney, Australia. This resulted in a 16-hour standoff, with spectators and police surrounding the building. When a shot was heard from inside the building, police raided the café and shot the gunman. Three people were killed, including the gunman. We collected data on this event for the explicit purpose of examining rumoring behavior, again using the Twitter Streaming API to track several event-related terms, including: `sydneysiege`, `martinplace`, `sydney`, `lindt`, and `chocolate shop`. Data collection began on 15 December at 11:06am AEDT and ended two weeks later, resulting in a dataset of just over 5.4 million tweets.

Coding Tweets for Affirm/Deny and Uncertainty

Following [2,4], we manually code tweets associated with each rumor along two dimensions. The first dimension, which we designed to identify crowd corrections, consists of five mutually exclusive categories: *Affirm*, *Deny*, *Neutral*, *Unrelated*, and *Uncodable*. The second dimension captures expressed uncertainty—tweet text that suggests in some way that the veracity of the rumor is not completely established. Three trained coders manually coded every distinct tweet (removing retweets and very close matches) in each rumor corpus. Inter-rater reliability was computed to ensure reliable codes (κ > 0.65). We use a "majority rules" process for adjudication where agreement by two or more coders determines the final code. To explore nuances of expressed uncertainty we extended this scheme.

Rumor	Total	Affirm	Deny	Neutral	Uncrtnty
Proposal	3146	79%	16%	5%	2%
False Flg	3568	92%	4%	4%	43%
Accused	27,934	82%	16%	2%	12%
Hadley	2679	97%	1%	2%	23%
Lakemba	1338	38%	61%	~0%	11%
Belts	2583	71%	3%	26%	23%

Table 1. Tweet Count by Primary Rumoring Behavior Codes

Developing a Coding Scheme for Uncertainty

Combining inductive and deductive methods we iteratively developed a coding scheme for uncertainty. The scheme is informed by previous literature, including studies on linguistic shields in medical-related dialog [23] and milling behavior during crisis [36]. Researchers began with a sample of 100 tweets labeled with *uncertainty* from two previously coded rumors. We grouped tweets according to perceived similarities, including linguistic patterns, grammatical constructs, and punctuation choices, yielding three groups that we initially labeled: *questions*, *hedges*, and *deflections*. These groupings provided a starting point, but it was clear that subdivisions would reveal additional nuance. In addition, several tweets did not fit into any of the original categories and some tweets had features pertaining to multiple groups. Hedge and deflection groups reflected concepts of linguistic shields, which offered insight into a more cohesive structure for the scheme [23]. Prince et al. [23] examined how physicians communicate with one another when talking about the patient's conditions, and discussed two types of expressed uncertainty in this context: *approximators* and *shields*. Approximators deal with 'fuzziness' within propositional content, e.g. "his feet were sort of blue," and were not present in rumoring. However, a large number of the rumor-related tweets contained shields, which deal with 'fuzziness' in the relationship between content and speaker, e.g. "I think his feet were blue." Price et al. [23] divide shields into two types: *Attribution*, where the speaker's uncertainty relates to the information source; and *Plausibility*, where the speaker's uncertainty relates to reasoning about the information's plausibility. Mapping the groupings to these two concepts—deflections to attribution shields and hedges to plausibility shields—gave grounding in prior work.

Though many tweets fit within these two broad categories, others did not. To expand our scheme, we use the concept of "verbal milling behavior," which Turner and Killian [36] define as a process through which people gather in times of crisis to discuss, hypothesize and attempt to understand the cause of the event. Incorporating ideas from this literature led to the identification of additional categories and the refinement of second-level subcategories. With the coding scheme defined, the team verified coder understanding of the categories. For each rumor, three researchers separately coded small subsets of random tweets, comparing codes to assess inter-coder reliability and flexed the coding scheme and coding definitions to ensure a shared understanding of the scheme (κ > 0.6 for all categories).

A Coding Scheme for Characterizing Expressed Uncertainty in Rumoring Tweets

The final coding scheme for expressed uncertainty in rumor-related tweets (illustrated in Figure 1) consists of two high-level (and non-exclusive) rumor-behavior categories: *shielding* and *milling*. Shields are employed by authors to protect themselves from making false statements. Milling behaviors include collective work to make sense of an uncertain space—e.g. interpreting, speculating, theorizing, debating, or challenging.

Shielding: Attributions

Attribution shields deflect responsibility for the information onto an external source—i.e. someone other than the author. We distinguish between three types of attribution shields: *named*, *unnamed*, and *implied*. Named attributions specify the person or entity who provided the information (*Boston Police said x*). Unnamed attributions note that there is a source, but do not specify exactly who that is (*reports state that x*). Implied attributions suggest that the author heard the information elsewhere without explicitly noting a source (*apparently x* or *allegedly y*).

Shielding: Plausibility

We identified two salient subgroups of plausibility shields: *personal* and *impersonal*. Personal plausibility shields locate the source of the uncertainty in the tweet author—

e.g. *I think x* or *it looks to me like y*. Posts with a personal plausibility shield often (though not always) also include building and doubting milling behavior. Impersonal plausibility shields contain language or punctuation that suggest some amount of uncertainty in the information space. Whereas a personal shield suggests that *"I do not know,"* an impersonal shield implies *"it is not known."* Common words and phrases in this category are: *possibly, may be, could be, unconfirmed*.

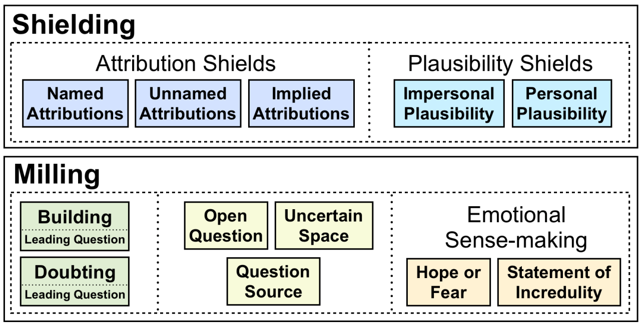

Figure 1. Coding Scheme for Expressed Uncertainty

Milling: Building and Doubting

Milling describes any type of behavior that built or shifted a rumor story while contributing to the uncertainty of the space. This category includes building (often speculative) behavior as well as challenging and doubting patterns. We theorize that these behaviors are similar because they are examples of users trying to make sense of events by proposing new theories to explain the information at hand. In our data, milling is very often seen in the form of leading questions, and we added specific sub-categories to distinguish these forms.

Milling: Open Question

Unlike leading questions, open questions do not attempt to confirm a hypothesis or theory. Instead they represent information-seeking behavior that generally corresponds to a neutral position in relation to the rumor. In contrast to leading questions that build or doubt a rumor, open questions reflect confusion or uncertainty about the information present. While leading questions imply that the tweeter has some opinion about the events but may be somewhat unsure about the veracity of available information, open questions suggest the author does not have an opinion one way or the other. Though this distinction is theoretically clear, in practice it may be difficult to differentiate between open and leading questions without making assumptions about users' intentions.

Milling: Uncertain Space

This type of tweet contains an explicit observation about the uncertainty or ambiguity in the information space—*e.g. "nobody knows what's happening"*. The author communicates that either facts about the rumor are not fully known or all the sources of information are unreliable, often by comparing contradictory information from different sources. This, unlike leading questions, is a neutral behavior, not seeking to affirm or deny the rumor and not weighing one side of the story more heavily than the other.

Milling: Question Source

Question source tweets challenge the legitimacy of a rumor or an argument about a rumor by questioning its source. This indicates that a user has, like in building or doubting behavior, chosen a side of the argument and is raising concerns over a source that conflicts with her beliefs. Doubting the credibility of a source is often done by calling out the lack of evidence and by using logic to undermine the position of the source.

Milling: Statements of Incredulity

This category is part of the broader emotional sensemaking category in which the uncertainty is expressed through an emotionally charged comment. Typical for statement of incredulity are phrases such as *"WTF???", "I can't believe this",* or *"Impossible!"*, expressing surprise over the information being presented. These phrases are often used in such a way that the tweet questions the possibility of an event happening; "*How could this happen?"* is a statement of incredulity, but it also doubts the likelihood that such an event happened. Unlike doubt or challenge discussed above, statements of incredulity simultaneously express some small measure of doubt while also accepting the truth, *"I can't believe this happened, but I know it did"*. This convergence of acceptance and doubt with emotional expression is distinct from other forms of uncertainty.

Milling: Hope or Fear

This category is the second branch we identified in emotional sensemaking. Hope or fear refers to statements such as *"Praying this isn't true"*, *"I hope this is just a rumor" or "I'm fearing the worst"* which give the sense that although the author is leaning towards believing the rumor, there is less than complete certainty about its veracity. The user hopes the information is false, or conversely fears it is true, thereby introducing the possibility of uncertainty.

Coding Tweets for Types of Expressed Uncertainty

For each tweet coded as having expressed uncertainty, we re-coded applying the above coding scheme. First, coders assessed whether the uncertainty in the tweet was *Related* or *Unrelated* to the rumor story. If *Unrelated*, no other codes were given and the tweet was removed from our analysis of uncertainty. Code categories are non-exclusive; a single tweet could have as many uncertainty codes as applicable. Each tweet was coded by one of three researchers (after establishing high reliability as discussed previously). Throughout coding researchers also kept track of specific phrases, words, punctuation or grammatical patterns associated with each code. This made it easier to see similarities between different categories and when the same phrase expressed different types of uncertainty.

Method of Analysis

In analyzing the coded rumors, we first graphed related uncertainty as it relates to affirm, deny and neutral signals from the original coding scheme, revealing peaks of the different signals over time. We examined tweets that had combined codes, e.g. affirm + uncertainty. We also created tables of totals for each first level code, related uncertainty, and the individual uncertainty codes, as seen in Table 1. We divided counts into original tweets and retweets, calculating retweet percentage as a proportion of the total. This way we could see when spikes were caused primarily by retweets or original tweets. Cross-comparison of tables from different rumors allowed us to see similarities and differences.

RESULTS

In this section we describe and analyze tweets related to the six rumors selected for this study. The first three rumors propagated during the aftermath of the 2013 Boston Marathon Bombings. The final three rumors relate to the 2014 Sydney Siege event. To provide context for each rumor, we describe each rumor holistically, though our goal is to highlight common patterns (as well as distinctions) across rumors in regard to expressed uncertainty.

Rumor #1: Proposal

This rumor claimed that a man was planning to propose to his girlfriend at the finish line, but that she had been injured or killed by the blasts. Using the search string ("propos" or "marry") and filtering to exclude (manually coded) unrelated tweets, we identified 3,146 tweets in our collection related to this rumor. It began with several affirming tweets with no expressed uncertainty, e.g.:

@userA[1] (Apr 15 6:22pm): This guy was gonna propose to his girlfriend today then she got killed by the bomb omg I'm crying

This rumor later became associated with a photo picturing a man attending to an injured woman after the bombings:

@userB (Apr 15 7:48pm): His girlfriend ran the Boston Marathon. He was waiting at the end for her, to propose. She died. #PrayForBoston http://t.co/UaB72fxHyF

Primary propagation occurred over a 36-hour period. Examining the temporal signature (Figure 2), we see the primary spike of rumor-affirming tweets—220 tweets per hour (TPH) at 10pm EDT April 15 (Point A)—is followed by a steady decline punctuated by several subsequent peaks of decreasing volume over time. We were rate-limited during this rumor's first few hours, and evidence from a complementary collection [18] suggests overall volume was about twice as high during its initial peak.

The denial signal for this rumor is weak, especially during its primary propagation window. 16% of total tweets in the rumor set were denials, and a large portion of these came very late in the rumor's lifecycle (Figure 2, Point B)—after CNN posted an article refuting event-related rumors.

Some Early Uncertainty, Especially in the Denial

There was little uncertainty (2.4%) in this rumor. Following its initial trend, the rumor largely propagated as what appeared to be factual statements that later turned out to be false. However, uncertainty was far more likely to appear in tweets denying the rumor (7.2% of denies) than tweets affirming the rumor (1.2% of affirms).

@userC (Apr 15 11:23pm): If the girl in the picture was running the marathon and her boyfriend was gonna propose at the end.. Then why was she behind the gate

The above denial tweet was posted just after the first peak and was coded as milling in the form of a doubting question—though, like 4% of all milling questions in our data, it does not contain a question mark. This tweet also uses a conditional statement */(If)(\w+)(then)(\w+)/*. Conditional statements such as these were a prominent pattern in milling behavior, both building and doubting, appearing in 15.3% of tweets across all rumors.

89% of the 36 rumor denial tweets with expressed uncertainty contained milling in the form of doubting. These were concentrated in the primary propagation window. Later in the cycle, denial tweets also often included attribution shields—for example, the below tweet has an implied attribution with the "apparently" and linked attribution to an article stating this rumor was false:

@userD (Apr 16 8:42pm): The story about a man who would propose to his gf at Boston finish line, apparently fake -> <link>

Using a shield within a denial tweet may be a strategy for avoiding direct confrontation. The rumor also included seven distinct tweets posing open questions, including:

@userE (Apr 16 7:02pm): The story about the guy almost proposing to his girl he lost during the Boston marathon...is it true?

Among the 31 affirming tweets with uncertainty, 35% contained implied attribution shields.

Rumor #2: False Flag by Navy Seals

This rumor claimed that the marathon bombings were a "false flag" attack perpetrated by agents of the U.S. government and designed to be blamed on someone else. Early versions of this rumor identified, from digital photos of the marathon crowd, men who were said to be Navy Seals. Later posts assigned blame to professional mercenaries. We scoped this rumor using ("navy seal" or "blackwater" or "black ops" or ("craft" and "security")).

The first tweet claiming a false flag connection to the bombings was sent just hours after the bombings. Over the next 24 hours, users posted a total of eight tweets related to this rumor. Of these, three had uncertainty, including:

@userF (Apr 16 9:42am): The Boston Marathon Bombings.. is this another false flag..or Black Ops..terror attack. by the US Fascist regime.. all to gain more control

[1] In the reporting presented here, we maintain real account names for response organizations, media organizations and professional journalists. We anonymize all other account names.

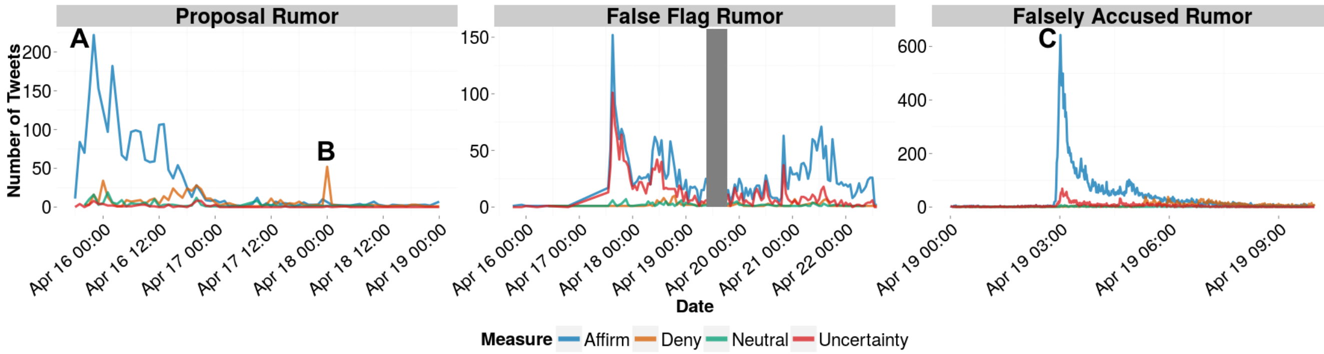

Figure 2: Tweet volume by rumor stance and expressed uncertainty for Boston marathon bombing rumors

Building a Rumor Through Leading Questions

The above tweet is a milling/building tweet in the form of a leading question without the question mark (*is this another false flag*), a pattern we saw previously in the Proposal rumor. This tweet contains another interesting pattern /(\w+)(..|…)(\w+)(..|…)/ that appeared several times in milling/building tweets across different rumors.

The first major peak in the rumor's temporal signature took place between 3:30pm and 8pm EDT on April 17. 531 tweets were sent during this period, with a maximum of 158 TPH at 4pm EDT (Figure 2). This surge in activity around this rumor began with an affirm tweet linking to an article on InfoWars, a political news outlet. The publication of that article appears to have set off a wave of commentary, including speculative tweets such as:

@userG (Apr 17 2:40pm): [infowars] Navy SEALs Spotted at Boston Marathon Wearing Suspicious Backpacks? <link> #nwo

Like many tweets in this rumor, this one contains uncertainty related to milling/building in the form of a leading question and an impersonal plausibility shield (the question mark). Tweets with impersonal shields in the form of leading questions can be viewed as a way of "hedging" or stating something potentially false or controversial without saying it firmly or factually. Of the 390 uncertainty tweets (73% of the total) in this wave, 323 were coded as milling/building, and almost all contain a leading question.

Early Uncertainty, Yet Very Weak Denial

This is a unique rumor in our study. The other five rumors have a finite window of propagation, typically with one large spike followed by a gradual but steady drop-off, punctuated by a few smaller spikes. For the Navy Seals rumor, after its initial spike, the rumor returns to relatively high volume (about 60 TPH) several times, often days apart. At the end of our collection window, the rumor is still propagating. Like the first major spike, many downstream peaks are coupled with tweets linking to online articles making these claims. Compared to the other rumors in this study, the percentage of retweets is relatively low (40%).

There are very few denials of this rumor (140 tweets or 4% of total volume). However, the uncertainty signal in this rumor is strong, especially early in its propagation window. A total of 1,533 tweets (43%) in this rumor have expressed uncertainty, including a strong majority in the initial spike. Interestingly, as the rumor continues, the uncertainty signal fades—in the final 24 hours of the collection period, only 13.7% of tweets had uncertainty.

Rumor #3: Falsely Accused

This rumor falsely asserted that Sunil Tripathi, a Brown University student who had gone missing March 16, 2013, was one of the suspected marathon bombers. Sunil's disappearance was fairly well publicized in the Boston area, and after the FBI released photographs of men they suspected were involved in the bombings, several social media sites—including Reddit and Twitter—speculated that Sunil resembled the one of the suspects. Searching our event data for ("sunil" or "tripathi"), there are 27,934 tweets related to this rumor.

The first few tweets referring to Sunil Tripathi as a possible suspect seem to refer to conversations occurring elsewhere, for example:

@userH (Apr 18 7:38pm): Sunil Tripathi - Some might think he looks like the kid in Boston. But the FBI photos are too grainy to say for sure. http://t.co/9y5CivjlCX

The above tweet affirming the rumor has several kinds of uncertainty, including milling/doubting, an unnamed attribution shield (*some might think*), and a personal plausibility shield (*but […] to say for sure*).

An Early Period of Persistent Low Volume Uncertainty

For the first several hours, volume of the Falsely Accused rumor was low, less than ten tweets per minute (TPM) as seen in Figure 2. Most tweets affirmed the rumor, and most contained uncertainty. During this time, uncertainty was primarily "building" milling behavior (63% of uncertainty tweets and 42% of the total)—i.e. speculating and theorizing like the tweet below, which includes a leading question and the repeating /(\w+)(..|…)/ pattern.

@userI (Apr 18 11:39pm): Hmm..Sunil Tripathi..Possible Suspect #2 Brown University student missing since last month? #bostonsuspects #Boston <link>

Plausibility shields (both personal and impersonal) were also common. 51% of the uncertainty tweets had impersonal plausibility shields, using phrases like *might be*, *could be, possible suspect,* and *unconfirmed report*:

@userJ (2013-04-18 9:17pm): One of the suspects in the Boston Marathon bombing might be the missing Brown University student Sunil Tripathi - <link>

A Rapid Shift in Expressed Certainty

The dynamics of this rumor changed drastically at 2:50am, immediately after this tweet, an affirm with no uncertainty:

@userK (Apr 19 2:50am): BPD scanner has identified the names: Suspect 1: Mike Mulugeta Suspect 2: Sunil Tripathi. #Boston

The information contained in this tweet quickly went viral (through retweets and original tweets containing the same claim), corresponding to a massive spike in the rumor, which peaked at 653 TPM at 3:02am (Figure 2, Point C). Most of these tweets contained no uncertainty—i.e. this information spread instead as a (false) factual claim.

Early on, this rumor contained high percentages of uncertainty, but as total volume peaked around the (false) scanner report, the signal shifted to mostly affirms, and though absolute volume of uncertainty rose with the increase in overall tweets related to the rumor, uncertainty as a percentage dropped drastically at this point (and stayed low throughout the remaining lifecycle of the rumor). Prior to the above scanner tweet, 72% of rumor-related tweets had uncertainty. After the scanner tweet, only 13% did.

The types of uncertainty expressed in the tweets changed as well. Attribution shields become more frequent, especially implied attributions (from 1.7% to 18%):

@userL (Apr 19 3:02am): apparently one of the boston bombing suspects is sunil tripathi, the brown univ student who disappeared w/o a trace about a month ago

In addition to *apparently*, words like *reportedly* and *alleged* were used to imply that the author had learned of the information elsewhere and was passing it along.

During the rapid spreading phase, plausibility hedges, both personal and impersonal, dropped. Milling/building behavior (as a percentage of uncertainty) also dropped, while milling/doubting behavior rose significantly (from 1.5% to 10% of uncertainty tweets). Statements of incredulity, like this one, increased significantly as well:

@userM (Apr 19 3:40am): Omg Sunil Tripathi is the bombing suspect in Boston?!?! ...wtf

Uncertainty Precedes the Denial

Though across the total rumor the volume of denials (4,570) is nearly equal to the volume of tweets with expressed uncertainty (4,173), the uncertainty signal appears much earlier than the denial signature. Prior to the scanner tweet, which likely brought this rumor to widespread attention, of 780 tweets only 18 were denials, but 529 had expressed uncertainty. This suggests that, at least for some rumors, expressed uncertainty precedes explicit denials.

Rumor #4: Ray Hadley Speaks to Hostages

Ray Hadley is a popular (and notorious) "shock-jock" radio host in Australia. He was on air during the Sydney Siege and participated in spreading several rumors, including the Lakemba Raids rumor described below. Around 12:30pm AEDT, Hadley claimed that he had been in contact via phone with one of the hostages. This claim was soon spread on Twitter (and elsewhere). Though the claim turned out to be true, the online crowd expressed some uncertainty about it, likely due to Hadley's reputation for spreading rumors.

@userN (Dec 15 12:31pm): Hostage apparently called ray Hadley demanding police leave the area #MartinPlaceSeige

The above tweet, the third we collected related to this rumor, affirms the rumor but passes it along with some uncertainty in the form of an implied attribution shield: *apparently* (as in, information heard from somewhere).

We scoped this rumor by searching for "Hadley or "radio host". It had very few denials (33 total). However, it did have a significant amount of uncertainty—23% of tweets have expressed uncertainty (608 tweets). Figure 3 shows how uncertainty accompanies the early peaks in affirmations and provides a much stronger signal of rumoring than the denial pattern.

Expressed Uncertainty in a True Rumor

The first small peak (10 TPM) occurred at 12:45pm AEDT. That volume was wholly constituted by retweets of this popular tweet propagating at the time:

@NewsOnTheMin (Dec 15 12:42pm): Distressed caller claiming to be hostage has contacted radio broadcaster Ray Hadley advising police move away. #sydneysiege

This tweet contains uncertainty around the identity of the caller, expressed as an unnamed attribution shield (*claiming to be a hostage*).

The second peak (82 TPM at 1:30pm) punctuates a one-hour period (1:20pm to 2:20pm) that contained the bulk of tweets related to this rumor. 1,586 tweets were sent during this time. Nearly 20% of them had expressed uncertainty, almost all in the form of named or unnamed attribution shields. There is almost no denial during this period. However, there were 38 neutral tweets, and all but one had expressed uncertainty. Most were retweets of the following tweet, sent by a journalist, which has an attribution shield (*@nswpolice spokeswoman said*) and an explicit reference to an uncertain information space (*she won't confirm*):

@MarkDiStef (Dec 15 1:49pm): .@nswpolice spokeswoman said she won't confirm if a hostage has been speaking to Hadley, "he can claim it if he wants"

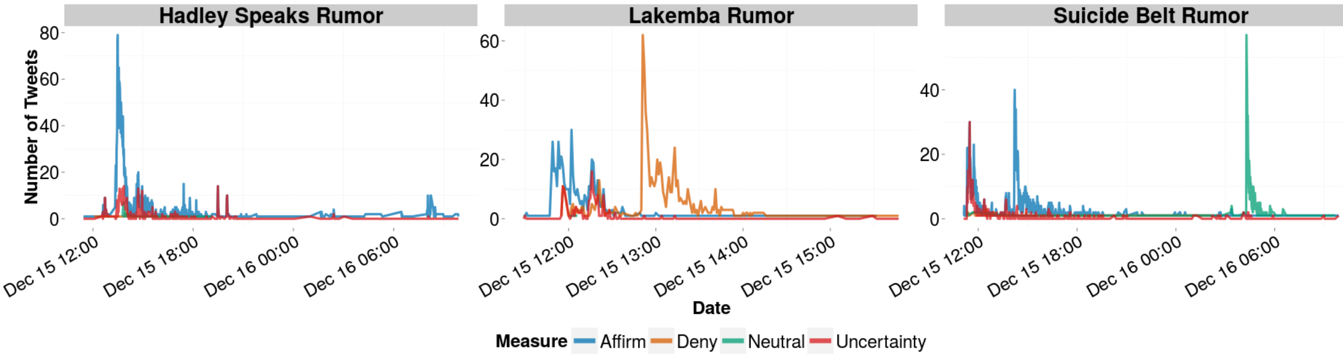

Figure 3: Tweet volume by rumor stance and expressed uncertainty for Sydney Siege rumors

Uncertainty in affirming tweets in this rumor was largely associated with shields, as many Twitter users seemed reluctant to wholly trust Hadley's claims. Denial tweets however, though few, exhibited mostly milling/doubting behavior, again primarily in the form of leading questions:

@userQ (Dec 15 1:45pm): Why the <expletive> would a hostage (or anyone, for that matter) want to speak to Ray Hadley specifically? I don't <expletive> think so.

@userR (Dec 15 2:02pm): so who confirmed that ray hadley is talking to a hostage? or was it ray hadley?

There are more of these doubting leading questions than there are straight denials in this rumor. Like the last tweet shown here, 9% of uncertainty tweets question the source of the information—i.e. Ray Hadley.

Rumor #5: Lakemba Raids

This false rumor claimed that the Australian Federal Police were raiding homes in the primarily Muslim Lakemba neighborhood during the Sydney Siege. In actuality (and coincidentally), there were 20 officers touring Lakemba Mosque at the time as part of a police induction day. This rumor first appeared on Twitter with tweets relaying the claim, attributing its source to Ray Hadley's radio show:

@userS (Dec 15 11:29am): Ray Hadley reporting on @2GB873 homes in Lakemba are being raided by police at present #siege

@userT (Dec 15 11:30am): BREAKING: UNCONFIRMED reports of police raids taking place across Lakemba in Sydney's west, which has a large Muslim community

The first tweet above was not coded as containing expressed uncertainty, though (like many tweets) it does attribute the information source to Hadley, which could be seen as an attribution shield. The second tweet contains an unnamed attribution shield (*reports*) and an impersonal plausibility shield (*unconfirmed*).

We scoped this rumor to tweets containing "Lakemba". The temporal signature of this rumor (Figure 3) shows an initial blip of a few tweets around 11:30am on Dec 15, followed by a series of affirming peaks of 25 to 30 TPM between 11:45am and 12:05pm. Following a trend seen in other rumors (Navy Seals, Falsely Accused, and Hadley), during the early part of the rumor, denials occur at lower volume and lag behind both affirmations and uncertainty. For the Lakemba rumor, after a small period of very low overall rumoring activity, there is a strong, clear denial signal. This was catalyzed by the following tweet, sent by the official account of the Australian Federal Police:

@AFPMedia (Dec 15 12:50pm): Reports that the AFP is conducting search warrants in the Sydney suburb of Lakemba are incorrect.

This tweet was retweeted 475 times and it effectively ended the spread of the rumor. After this point, there were very few affirming (39) or uncertainty (10) tweets.

In this rumor, a small number of tweets (and retweets of those tweets) are responsible for a large portion of the uncertainty. For example, of 77 affirming tweets that have uncertainty (15% of the total affirmation signal), 70 are retweets of the following tweet from a media organization:

@9NewsSyd (Dec 15 12:14pm): JUST IN: Raids occurring at Lakemba homes in south west Sydney. It's unknown if raids are related to siege underway in Sydney…

This tweet's uncertainty occurs in the phrase *"It's unknown if"* which can be viewed as an impersonal plausibility shield or an explicit mention of uncertainty in the information space. Within the denial signal, 58 tweets with uncertainty are retweets of the text below, which has an unnamed attribution shield (*My sources … are saying)*:

@safimichael (Dec 15 11:55am): My sources in Lakemba are saying there are no raids underway in that suburb. Police are at Lakemba mosque as part of induction day #siege

Rumor #6: Suicide Belt or Other Explosive Device

This false rumor asserted that the hostage taker was wearing an explosive device of some kind. We scoped this rumor using ("suicide" or "belt" or "vest" or "backpack"). Although this rumor began spreading before our collection, we used retweet records—which point back to the original tweet and provide the number of times retweeted—to get some sense of earlier volume. The first widely visible tweet we can identify was posted by a journalist on Dec 15 at 11:04pm AEDT (three minutes before collection started):

@turnerscope (Dec 15 11:04am): We believe there are 13 hostages inside Lindt Cafe in Martin Place. Woman saw man with backpack and possibly gun walk in at 944am.

Mainstream Media, Uncertainty, and Rumor Spreading

This factual tweet became fodder for speculation—among professional journalists and others—around the purpose of the backpack and the intentions of the hostage taker. A first wave of tweets connecting the backpack (and a vest) to explosive devices occurred between 11:20am and 12:20 pm, averaging about 9 TPM and peaking at 30 TPM at 11:30am. More than half (61%) of tweets sent during this hour-long period had uncertainty. Again, the uncertainty signal was stronger and earlier than the denial signal, which comprised only 2.5% of tweets.

The vast majority (329 of 345) of these early uncertainty tweets contained impersonal plausibility shields:

@NewsOnTheMin (Dec 15 11:24am): MORE: One of the terrorist inside the coffee shop is wearing backpack and vest, likely a bomb. #Sydney http://t.co/88FHhLw3qo

This tweet functions to pass along the existing rumor, but includes language that shields the author from making false claims. Notably, these impersonal plausibility hedges were often utilized by news media accounts, which played a major role in propagating this rumor. Other affirming tweets with uncertainty spread the rumor with attribution shields pointing back to media sources.

For a few hours, volume decreased to less than 5 TPM, then a second wave of affirmations occurred, peaking at 39 TPM at 2:14pm AEDT. Much of that volume was generated by a few highly-retweeted tweets from "breaking news" accounts [2]. Very few tweets (38 of 378) sent during this wave had uncertainty. The final wave of neutral tweets was generated almost entirely by retweets of a @cnnbrk tweet (with no uncertainty) linking to an after-action report, which stated both that the gunman had been wearing a backpack and that it had been checked for explosives.

DISCUSSION AND CONCLUSION

A major contribution of this study is a theoretically grounded framework for identifying and classifying types of expressed uncertainty, building upon existing work on linguistic shields in medical settings [23] and milling behavior during crisis events [36]. By applying this coding scheme to a large corpus of tweets we identified specific words and linguistic patterns that are characteristic of different types of rumoring behavior.

Linguistic Shields in Rumoring

In each rumor we studied, more than half of tweets that contained expressed uncertainty were coded as employing linguistic shields—i.e. mechanisms that protect the author if a rumor turns out to be false. For four of these rumors (False Flag, Falsely Accused, Lakemba, and Explosive Devices), these were primarily impersonal plausibility shields, often communicated through words like *possibly* or phrases like *could be*. Attribution shields, which deflect the responsibility of a rumor onto another source, were heavily employed in the Proposal and Hadley rumors and appeared in relatively high volume in the Lakemba rumor as well (see Table 2). Specific dynamics of those rumors may help explain the different types of attribution shields used—e.g. implied attribution shields for an Internet Meme rumor that had no real source (Proposal), and named attribution shields for a true rumor that originated from a less-than-credible source (Hadley).

	Prpsl	False Flag	Flsly Accsd	Hdly Spks	Lake-mba	Suicde Belts
Total Uncrtnty	74	1533	3412	608	143	594
Shields						
Attribution Shields						
Named	0%	6%	16%	70%	2%	6%
Unnamed	27%	3%	6%	16%	43%	23%
Implied	62%	3%	16%	5%	3%	8%
Plausibility Shields						
Impersonal	7%	84%	51%	18%	56%	67%
Personal	11%	4%	18%	1%	0%	6%
Milling						
Building	14%	83%	44%	3%	1%	8%
Leading Q	7%	79%	21%	0%	<1%	5%
Doubting	51%	3%	9%	7%	1%	1%
Leading Q	18%	1%	3%	6%	1%	1%
Open Q	10%	2%	2%	2%	2%	4%
Uncertain Space	3%	1%	4%	10%	0	5%
Q Source	3%	<1%	1%	9%	2%	1%
Incredulity	18%	2%	6%	1%	1%	1%

Table 2. Each Uncertainty Code as a Percentage of Overall Uncertainty Across Rumors

Expressed Uncertainty in the Rumor Mill

Previous research describes verbal milling behavior during crisis events as a process of collective sensemaking through which people provide and discuss possible explanations of what has occurred [26,36]. Among tweets with expressed uncertainty in the crisis-related rumors we studied, 60% demonstrated milling behavior. Though present in each rumor, milling behaviors were far more prominent in the rumors related to the Boston Marathon event. This is likely indicative of the nature of the specific rumors we selected for analysis—a conspiracy theory (False Flag Rumor) and crowd-sleuthing activity to find the suspects (Falsely Accused Rumor) show the most milling activity. It is important to note that not all milling behavior includes expressed uncertainty, so the tweets examined here (and the categories derived from those tweets) represent a subset of overall milling behavior.

Using Leading Questions to Build/Challenge Rumors
One especially salient pattern in the rumors we analyzed was the use of leading questions. Leading questions were a significant category (more than 5% of uncertainty tweets) in five of six rumors we studied (see Table 2). Across the six rumors, 2148 tweets (and 34% of all uncertainty tweets) were in the form of leading questions. In leading question tweets, it is clear which side of the rumor story (affirm/deny) the author is leaning towards; the question is phrased in a biased way, showing a lack of neutrality that we would see in an open question. Often, these questions are indicated by the use of question marks. Other times they are indicated by interrogative words or the use of the /(…)/ pattern. When leading questions occur, it is often not just one question but rather a series of questions, often within one tweet, for example:

```
@DavidVitter can you find out why the Navy Seals were in Boston? And why are they lying about saying there was a bombing drill?
```

Researchers have explored how leading questions can be used to guide witness testimony, with significant implications in legal contexts [e.g. 11,16,32]. In experimental work, Loftus [16] demonstrated that leading questions, including true and false presuppositions, can affect how eyewitnesses recall event details, and Swann et al. [32] showed how leading questions by serving as conjectural evidence to guide listeners towards making specific kinds of inferences about the information presented. However, leading questions have not explicitly been examined in regards to their role in online rumoring, where our data suggest they serve multiple purposes. Aligned with Swann et al.'s [32] findings, leading questions in rumoring tweets may enact a rhetorical strategy (intended or not) that effectively spreads unsupported claims.

Leading Questions as Impersonal Plausibility Shields
We also see leading questions employed as a shielding strategy—i.e. as a method for spreading information (or doubting information) without fully committing to those claims. In our data, 73% of milling leading questions were also coded as impersonal plausibility shields, reflecting both an overlap in those behaviors/strategies and a difficulty in distinguishing between the two, for example:

```
Sunil Tripathi: one of the marathon bombing suspects? <link>
```

We hypothesize that people phrase claims in a question format both as an expression of some doubt and as a mechanism for avoiding blame if the theory they are putting forward is later proven false. Often, tweets of this type contain a statement with no interrogatory marker other than the question mark at the end, but this question mark is extremely important, because, as the above tweet shows, with limited space in a 140-character tweet, that single addition (or an added …) can significantly change the function of a linguistic rumoring act.

Temporal Patterns in Uncertainty
Aligning with Zhao et al.'s [39] claims that skepticism may precede explicit corrections, for most rumors in this set, expressed uncertainty occurs earlier than denials. For some, the volume of uncertainty is also greater than that of denial. The relative volume of uncertainty, as well as when that uncertainty occurs within the rumor lifecycle, can provide insight into a rumor's type. For example, Internet meme-type rumors [17] like the Proposal rumor propagate with very little uncertainty and likely represent a different kind of "problem" for crisis communications than other more speculative rumors—such as Falsely Accused. In the Navy Seals, Falsely Accused and Explosive Devices rumors we can see a distinct shift from an early speculative phase to a later phase when the rumor propagates "factually" as misinformation. In Falsely Accused, this phase shift can be traced to a specific tweet regarding information shared on the police scanner. Shifts in the type of uncertainty, for example from milling/building to attribution shields, may also align with significant moments in a rumor's evolution and characterize certain types of rumors.

Implications and Future Work

Using Expressed Uncertainty to Detect Rumors
Recent work explores leveraging collective sensemaking processes, specifically crowd corrections [9,19] and skepticism [39] to build automated rumor detection systems. Our work suggests that expressed uncertainty may be an earlier indicator of rumors than denials or corrections, which could improve the speed of detection, offering a promising new direction for future work. Moreover, specific kinds of expressed uncertainty, especially those that accompany the affirming phases of some rumors—e.g. attribution and plausibility shields and milling/building—could be powerful, early predictors of rumoring.

Using Expressed Uncertainty to Understand Sensemaking
The larger goal of this research project is to better understand collective sensemaking online. In this paper, we discuss how uncertainty is expressed in rumors, and describe a coding scheme derived to categorize distinct types of uncertainty. Though we focus on the crisis context, the coding scheme we have developed for expressed uncertainty could be applied to other kinds of online discussions—wherever uncertainty is detected. Posts with expressed uncertainty represent only a subset of online sensemaking behaviors, but we intend to expand this coding scheme in future work to address a more complete range of sensemaking activities, including posts without uncertainty.

ACKNOWLEDGMENTS
This research is a collaboration between the emCOMP Lab, DataLab and SoMe Lab at the University of Washington, supported by NSF Grants 1420255, 1541688, 1243170 and PAC121028. We wish to thank students who provided significant assistance to this project, including John Robinson, Stephanie Stanek, Logan Walls, Coulter L'Heureux, Chris Lee, Annie Tao, and Hai Nguyen.

REFERENCES

1. Gordon W. Allport and Leo Postman. 1947. *The Psychology of Rumor*. Henry Holt and Company.
2. Cynthia Andrews, Elodie Fichet, Stella Ding, Emma S. Spiro, and Kate Starbird. 2016. Keeping Up with the Tweet-Dashians: The Impact of 'Official' Accounts on Online Rumoring. *In Proceedings of the ACM 2016 Computer Supported Cooperative Work (CSCW 2016).*
3. Susan Anthony. 1973. Anxiety and rumor. *The Journal of Social Psychology,* 89, 1: 91-98
4. Ahmer Arif, Kelley Shanahan, Fang-Ju Chou, Yoanna Dosouto, Kate Starbird, Emma S. Spiro. 2016. How Information Snowballs: Exploring the Role of Exposure in Online Rumor Propagation. *In Proceedings of the ACM 2016 Computer Supported Cooperative Work (CSCW 2016).*
5. Austin S. Babrow, Chris R. Kasch, and Leigh A. Ford. 1998. The many meanings of uncertainty in illness: Toward a systematic accounting. *Health Communication,* 10(1), 1-23.
6. Austin S. Babrow, Stephen C. Hines, and Chris R. Kasch. 2000. Managing uncertainty in illness explanation: An application of problematic integration theory. *Explaining illness: Research, theory, and strategies,* 41-67.
7. Prashant Bordia, and Nicholas DiFonzo. 2004. Problem solving in social interactions on the Internet: Rumor as social cognition. *Social Psychology Quarterly* 67, no. 1 (2004): 33-49.
8. Dale E. Brashers. 2001. Communication and uncertainty management. *Journal of Communication*, *51*(3), 477-497.
9. Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. 2013. Predicting information credibility in time-sensitive social media. *Internet Research* 23, no. 5: 560-588.
10. Nicholas DiFonzo and Prashant Bordia. 2007. *Rumor psychology: Social and organizational approaches.* American Psychological Association.
11. Edward R. Geiselman, Ronald P. Fisher, Gina Cohen, and Heidi Holland. 1986. Eyewitness responses to leading and misleading questions under the cognitive interview. *Journal of Police Science & Administration.*
12. Amanda L. Hughes, and Leysia Palen. 2012. The evolving role of the public information officer: An examination of social media in emergency management. *Journal of Homeland Security and Emergency Management,* 9(1).
13. Starr Roxanne Hiltz, Jane Kushma, and Linda Plotnick. 2014. Use of social media by US public sector emergency managers: Barriers and wish lists. *In Proceedings of ISCRAM (2014).*
14. Amanda L. Hughes, Lise A. St Denis, Leysia Palen, and Kenneth M. Anderson. 2014. Online public communications by police & fire services during the 2012 Hurricane Sandy. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, pp. 1505-1514.
15. Amanda L. Hughes, and Leysia Palen (2012). The Evolving Role of the Public Information Officer: An Examination of Social Media in Emergency Management. Journal of Homeland Security and Emergency Management, 9 (1).
16. Elizabeth F. Loftus, (1975). Leading questions and the eyewitness report. *Cognitive psychology*, *7*(4), 560-572.
17. Jim Maddock, Kate Starbird, Haneen J. Al-Hassani, Daniel E. Sandoval, Mania Orand, and Robert M. Mason. Characterizing online rumoring behavior using multi-dimensional signatures. *In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (ACM 2015)* 228-241.
18. Jim Maddock, Kate Starbird and Robert M. Mason. (2015). Using Historical Twitter Data for Research: Ethical Challenges of Tweet Deletions**.** Presented at *CSCW '15 Workshop on Ethics at the 2015 Conference on Computer Supported Cooperative Work (CSCW 2015)*, Vancouver, Canada.
19. Mendoza, Marcelo, Barbara Poblete, and Carlos Castillo. 2010. Twitter Under Crisis: Can we trust what we RT?. In *Proceedings of the first workshop on social media analytics*, pp. 71-79. ACM.
20. Onook Oh, Kyounghee Hazel Kwon, and H. Raghav Rao. 2010. An Exploration of Social Media in Extreme Events: Rumor Theory and Twitter during the Haiti Earthquake 2010. In *ICIS*, p. 231.
21. Onook Oh, Manish Agrawal and H. Raghav Rao. Community Intelligence and Social Media Services: A Rumor Theoretical Analysis of Tweets during Social Crises. *Mis. Quarterly 37*(2), 407-426.
22. Susan Coppess Pendleton. 1998. Rumor research revisited and expanded. *Language & Communication,* 18,1: 69-86.
23. Ellen F. Prince, Joel Frader, and Charles Bosk. 1982. On hedging in physician-physician discourse. *Linguistics and the Professions,* 8: 83-97.
24. Ralph L. Rosnow. 1980. Psychology of rumor reconsidered. *Psychological Bulletin.* 87, 3: 578-591.
25. Ralph L. Rosnow, 1991. Inside rumor: A personal journey. *American Psychologist,* 46, 5: 484.
26. Tamotsu Shibutani. 1966. *Improvised News: A Sociological Study of Rumor.* New York: Bobbs-Merrill.

27. Emma S. Spiro, Sean Fitzhugh, Jeannette Sutton, Nicole Pierski, Matt Greczek, and Carter T. Butts. 2012. Rumoring during extreme events: A case study of Deepwater Horizon 2010. *In Proceedings of the 4th Annual ACM Web Science Conference (ACM 2012)*, 275-283.

28. Kate Starbird, Dharma Dailey, Ann Hayward Walker, Thomas M. Leschine, Robert Pavia, and Ann Bostrom. 2015. Social Media, Public Participation, and the 2010 BP Deepwater Horizon Oil Spill. *Human and Ecological Risk Assessment: An International Journal,* 21, 3: 605-630.

29. Kate Starbird, Jim Maddock, Mania Orand, Peg Achterman, and Robert M. Mason, (2014). Rumors, False Flags, and Digital Vigilantes: Misinformation on Twitter after the 2013 Boston Marathon Bombings. Short paper. *iConference 2014*, Berlin, Germany.

30. Kate Starbird, and Leysia Palen. 2010. Pass it on?: Retweeting in mass emergency. In Proceedings of the International Conference on Information Systems for Crisis Response and Management (ISCRAM 2010).

31. Kate Starbird, and Leysia Palen. 2011. Voluntweeters: Self-organizing by digital volunteers in times of crisis. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (ACM 2011)*, 1071-1080.

32. William B. Swann, Toni Giuliano, and Daniel M. Wegner. 1982. Where leading questions can lead: The power of conjecture in social interaction. *Journal of Personality and Social Psychology* 42(6): 1025.

33. Jeannette Sutton. 2010. "Twittering Tennessee: Distributed Networks and Collaboration following a Technological Disaster. *In Proceedings of the International Conference on Information Systems for Crisis Response and Management (ISCRAM 2010)*.

34. Jeannette Sutton, Leysia Palen, and Irina Shklovski. 2008. Backchannels on the Front Lines: Emergent Uses of Social Media in the 2007 Southern California Wildfires. *In Proceedings of the 5th International ISCRAM Conference*, 1–9.

35. Jeannette C. Sutton, Ben Gibson, Emma S. Spiro, Cedar League, Sean M. Fitzhugh, and Carter T. Butts. 2015. What it takes to get passed on: message content, style, and structure as predictors of retransmission in the Boston Marathon bombing response. *PLoS one* 10(8): e0134452.

36. Ralph H. Turner, and Lewis M. Killian. 1957. *Collective Behavior.* Prentice-Hall, Englewood Cliffs, NJ.

37. Charles J. Walker & Carol A. Beckerle. 1987. The effect of anxiety on rumor transmission. *Journal of Social Behavior and Personality*, 2, 353–360.

38. Karl E. Weick, and Kathleen M. Sutcliffe. 2011. Managing the unexpected: Resilient performance in an age of uncertainty. John Wiley & Sons..

39. Zhe Zhao, Paul Resnick, and Qiaozhu Mei. 2015. Enquiring Minds: Early Detection of Rumors in Social Media from Enquiry Posts. In *Proceedings of the 24th International Conference on World Wide Web*, pp. 1395-1405. International World Wide Web Conferences Steering Committee.