Association for Information Systems
AIS Electronic Library (AISeL)

Research Papers

ECIS 2019 Proceedings

5-15-2019

FEAR AND LOATHING IN BOSTON: THE ROLES OF DIFFERENT EMOTIONS IN INFORMATION SHARING ON SOCIAL MEDIA FOLLOWING A TERROR ATTACK

Hissu Hyvärinen
IT University of Copenhagen, hissu@itu.dk

Roman Beck
IT University, romb@itu.dk

FEAR AND LOATHING IN BOSTON: THE ROLES OF DIFFERENT EMOTIONS IN INFORMATION SHARING ON SOCIAL MEDIA FOLLOWING A TERROR ATTACK

Research paper

Hyvärinen, Hissu, IT University of Copenhagen, Copenhagen, Denmark, hissu@itu.dk

Beck, Roman, IT University of Copenhagen, Copenhagen, Denmark, romb@itu.dk

Abstract

Emotions are essential to how we communicate, and online discussions are no exception. As most of the analysis on emotion so far has looked at polarity rather than specific emotions, we do not yet have a full understanding of how different emotions spark different behaviours. This study examines how five different emotions are associated with information sharing in the context of a terror attack both on a large scale and when including geolocation information in the analysis. Contrary to what previous findings suggest, increased fear and contempt levels have a negative relation with increased levels of retweeting. Positive emotion in tweets meant a decrease in retweet rates in the geolocation specific data, but an increase when all tweets were considered.

Keywords: social media, emotion, sentiment analysis, terror attacks.

1 Introduction

Social media has since its emergence quite drastically changed the way we communicate, not only affecting how and how easily it can be done, but also altering with whom we can connect. Some of the most common user reported motivations for using social media platforms are creating and maintaining connections with other users, sharing and obtaining information, and personal enjoyment (Dickinger et al., 2008; Ellison et al., 2007; Lin and Lu, 2011). Compared to traditional media, social media enables information to travel faster and reach a wider audience, resulting to phenomena such as allowing the development of collective situation awareness in crisis (Mukkamala and Beck, 2016), or – on the darker side of things – various types of rapidly escalating firestorms (Pfeffer et al., 2014).

Some social media messages go viral, but we do not yet fully understand what compels people to share them. Emotions certainly play a role in the decision to share information or opinions, but the findings regarding how, exactly, vary from one case to another (Gruzd et al., 2011; Hansen et al., 2011; Stieglitz and Dang-Xuan, 2013). The context of the discussion may be part of the explanation: different types of events spark very different types of online conversation (Ferrara and Yang, 2015). In *anticipatory discussions*, most of the discussion happens before the peak, which typically occurs around some real-world event. The converse is true for *unexpected events*, which quickly spark a peak of conversation that fades away. *Symmetric discussions* have a less distinct peak, and the discussion goes on for a longer period of time, whereas in the case of *transient events*, the peaks in conversation activity are sharp and bursty, and the activity fades quickly.

Emotions are contagious within a social network (Fowler and Christakis, 2008), and it has been shown to also apply to online environments (Hancock et al., 2008; Kramer et al., 2014; Kwon and Gruzd, 2017). It is therefore no wonder that emotions affect our information sharing behaviour (Gruzd, 2013; Hansen et al., 2011; Oh et al., 2013; Stieglitz and Dang-Xuan, 2013); when a social media user sees something emotional on social media, sometimes the emotion is passed on to the user, which makes the user more likely to share the emotion-evoking message onwards. We hypothesize that this also applies to the conversations occurring during and after a crisis event such as a terrorist attack.

In addition to context affecting the impact of emotions on information sharing, a potential explanation to the differences between previous findings is that some emotions may drive sharing behaviour more strongly than others: for example, anger is more likely to spark the action of sharing news than sadness in a one-to-one communication relation (Berger and Milkman, 2012). While previous research has yielded plenty of valuable information about the relationship of emotions and information sharing online, it seems that most of the analysis so far has focused on measuring the positivity and negativity levels without looking at specific emotions separately. Therefore, to expand our understanding of the role of specific emotions, the main goal of this study is to investigate *how different emotions are related to the degree of information sharing on social media in the context of a terror attack*.

In addition to hypothesising that context and the type of emotion are connected to how information is shared, we consider whether geographic location is related to how and whose information is shared online. Research on the after-effects of 9/11 found that in the aftermath of a terror attack, people in the affected regions suffer from more elevated stress and anxiety than people farther away. This may mean the emotional intensity of online communication may depend on the location of the actor. Proximity to the attack location might also enable an actor to provide timely, relevant information, which might compel others to share the information onwards more actively.

Geolocation data can be a valuable tool for investigating local phenomena, not the least in crisis situations such as natural disasters (Mukkamala and Beck, 2016). However, disclosing location information is typically voluntary on social media platforms, which means that the proportion of users who actively decide to do so may be small enough to noticeably limit the size of analysable data, and potentially introduce an unknown degree of self-selection bias in the dataset. To examine the effect of the location, we analyse a geotagged subset of the data used for this research and report the results for both the full dataset and the location specific subset side by side to examine the impact of including location information into the analysis of online information sharing.

The remainder of the paper is organised in the following way: First, we lay out the basis of our study by going through existing research on the topic and formulating our hypotheses. In the section Methods and Data, we explain our data collection and processing, the reasoning behind our approach, and the methodology applied in this study. In the subsequent section, we report our results, after which we discuss our findings in the context of existing knowledge. The last section reports our conclusions, limitations, and suggestions for future research in the area.

2 Theoretical Background

2.1 The Role of Emotions in Information Sharing Online

Where there is conversation, there is also emotion. Sentiment analysis enables analysing the presence and extent of emotions in an automated fashion, and has been used for purposes such as customer relation management (Risius and Beck, 2015), mining for electronic word-of-mouth (Chen et al., 2017; Kim et al., 2017; Relling et al., 2016), and predicting changes in the stock market (Risius and Beck, 2015) as well as outcomes of sports matches (Schumaker et al., 2016). There are two main approaches to sentiment analysis: lexicon-based approaches utilise dictionaries containing information about the emotional loadings of words, and machine learning based approaches use training data sets and/or features to build classifiers to sort text into emotion categories (Pang and Lee, 2008). Most sentiment analysis looks at the polarity (positive and negative sentiment) of the text analysed, but there are some studies that adopt a more fine-grained approach (Hyvärinen and Beck, 2018).

Over the last decade, Twitter has become a massively popular social media platform with celebrities and laymen alike expressing their views, and often retweeting messages authored by other people. Retweeting was a convention that organically emerged among the users during the early years, after which it has steadily become more and more commonplace (Liu et al., 2014). By 2014, around 25–

30% of all messages posted on Twitter were retweets, which makes it clear that sharing information and opinions is an integral part of the conversation culture (Liu et al., 2014).

It seems the tone varies from one conversation to another. A study on tweets about the 2010 Winter Olympics found that there are more positive than negative tweets, and that positive tweets were three times more likely to be retweeted (Gruzd et al., 2011). On the other hand, a study on German political tweets found that emotional tweets are more likely to get retweeted regardless of whether they are positive or negative (Stieglitz and Dang-Xuan, 2013). A third study posits that news related content propagates better when the sentiment is negative, whereas the opposite holds for non-news content (Hansen et al., 2011). The studies agree on elevated emotions being related to increased information sharing, but the descriptions of the exact nature of that relationship vary. One explanation for the differences between the findings could be the type of the event examined, as that has been found to influence the emotion levels of online conversations (Ferrara and Yang, 2015). Another possible contributing factor could be how emotional contagion works. Emotions have been shown to be contagious online (Hancock et al., 2008; Kramer et al., 2014; Kwon and Gruzd, 2017), but how contagion works may depend on the type of conversation, the relationship between participants, and the type of emotion in question.

A study examining the role of emotions in forwarding information via email finds that high activation emotions (such as anger or fear, sometimes also referred to as high arousal emotions) are associated with an increased tendency to share information, which means that valence (positivity or negativity) alone is not sufficient in explaining information sharing behaviour (Berger and Milkman, 2012). Although the study focused on a dyadic context, it is possible that the finding also applies to social media sharing, which is more of a broadcasting, one-to-many type of communication setting. We were therefore interested in looking at different emotions beyond polarity analysis in order to find out whether they have unique effects on information sharing on social media, which is a question the previous studies in the area have not yet covered, to the best of our knowledge.

Psychology literature contains various categorisations and definitions for emotions; some divide emotions into distinct states such as enjoyment, sadness, anger, fear, and disgust (Ekman, 1992), others define them as points in the dimensions of valence (or pleasure) and arousal (or activation) (Russell, 2003). The hierarchical domain of emotions developed by Ekkekakis (Ekkekakis, 2013) combines the distinct state and dimension approaches into a comprehensive framework while drawing from the previously existing wisdom (see Table 3). This study uses a sentiment analysis approach developed based on that framework, developed for the purpose of analysing differential emotions on social media (Risius et al., 2015). The positive emotion categories are *affection, happiness*, and *satisfaction*, and the negative ones are *anger, fear, depression*, and *contempt*. Our initial plan was to treat each emotion separately, but – upon finding correlations between the positive emotions in the analysis phase notable enough to potentially cause trouble – decided to merge the three positive emotion categories into a single one. Our observation regarding the proximity of the positive emotions could be due to negative emotions being more distinct from each other than positive ones (Fredrickson, 1998).

Negative emotions have been associated with elevated levels of information sharing in several past studies (Berger and Milkman, 2012; Hansen et al., 2011; Stieglitz and Dang-Xuan, 2013). We expect that effect to be particularly pronounced in our dataset, given the context of a terror attack evoking several types of negative emotions. This assumption is further supported by a study on information propagation on Twitter following the Woolwich terrorist attack in 2013 concluding that the presence of (any) emotion in tweets is related to a higher level of retweeting (Burnap et al., 2014), which is in accordance with the findings concluding that sentiment increases information diffusion on social media in other contexts (Gruzd et al., 2011; Hansen et al., 2011; Stieglitz and Dang-Xuan, 2013).

2.2 Information Sharing Online in the Context of a Terror Attack

The primary function of using social media in a disaster situation is sharing and obtaining information, which also allows actors to make sense of the events (Eismann et al., 2016). In the context of human caused disasters – such as terror attacks – social media are used for expressing emotions and memorialising victims, establishing connections between geographically distant members of the community, and coordinating response and recovery efforts (Huang et al., 2010; Kaufmann, 2015; Mazer et al., 2015; Neubaum et al., 2014). Twitter is used across disaster categories by all types of social units to share warnings and situational updates, but also in a more interactional fashion, such as confirmations on others' wellbeing and conversations on events and their consequences (Eismann et al., 2016). In crisis situations such as natural disasters, anxiety has been linked to the behaviour of spreading rumours (Oh et al., 2013), which means anxiety may correlate with an increased urge to share information also in the context of an act of terror.

In the wake of a terrorist attack, people are driven to seek information, but also talk about the attacks on social media in order to defend their cultural world views and maintain their self-esteem (Fischer et al., 2016). An act of terrorism will increase the levels of fear, uncertainty, and anger in people's minds, which affects their behaviour also online (Boyle et al., 2004). People close to areas where terror attacks occur report stress and anxiety after the incident, which leads us to include geographic proximity as a variable in our analysis (Morrison et al., 2001).

Anger was the dominant reaction to the 9/11 attacks, and was particularly intense in the New York area where the levels of negative emotions in general were higher than in the rest of the country (Smith et al., 2001). We therefore expect the levels of anger to be high, and be likely to be actively passed on due to emotional contagion in the aftermath of the Boston Marathon bombing.

Hypothesis 1a: The higher the level of anger in a terror-related tweet, the more it is retweeted

The second most prevalent emotion following a terrorist attack is fear, which is typically related to questions such as how the crisis affects one's own life or whether anyone is safe (Smith et al., 2001). Like anger, fear is a high activation emotion, and is therefore likely to be associated to increased information sharing (Berger, 2011).

Hypothesis 1b: The higher the level of fear in a terror-related tweet, the more it is retweeted

The level of depression was also found to be elevated following an act of terror (Lerner et al., 2003; Smith et al., 2001). Crying was reported as one of the most common physical and emotional symptoms following the 9/11 attack (Smith et al., 2001). Due to sadness – for which this study uses the term depression – being a low activation emotion, we expect the rate of sharing to be lower than in the case of anger and fear. However, we still expect the relation to be positive. People tend to feel the need to find a shared space for mourning a crisis event, leading to online convergence (Hughes et al., 2008). Consequently, we hypothesize that feeling sadness, and feeling the need to share that will also mean users are more likely to relate to content matching their emotions and thus more likely to share it.

Hypothesis 1c: The higher the level of depression in terror-related a tweet, the more it is retweeted

In this context of usage, contempt is defined as a negative emotion related to socially offensive or inappropriate actions (e.g. "deceitful", "despicable"), and personal reactions to them (e.g. "shame", "guilt", "condemn"). This is perhaps the most unpredictable of the negative emotions with respect to information sharing. In general, we know that experiencing an emotion will make people want to share it with people around them; however, guilt and shame seem to be exceptions to this rule (Rimé, 2009). A study comparing emotional reactions to the shooting of John F. Kennedy and to the 9/11 attack found that the shooting evoked more shame related emotions than 9/11, and that people were less willing to discuss the event with others than after 9/11 (Smith et al., 2001). Although the study does not establish causality, the findings are in line with shame in general being associated with lower willingness to share emotions. We therefore hypothesize that contempt is the only negative emotion associated with decreased retweeting.

Hypothesis 1d: The higher the level of contempt in a terror-related tweet, the less it is retweeted

Hypothesising how positive emotions are related to information sharing in the context of a terror attack is less straight-forward as with negative ones. It may be that in a context abundant with negative emotions sparked by immediate negative events, positive messages would feel less relevant and thus be shared less. On the other hand, in the case of non-news content, positive emotion is associated with increased information sharing; perhaps gratitude towards helpers, sharing experiences and thoughts, or relief might prompt retweeting under such circumstances. Based on previous findings, positive emotion tends to rather increase than decrease retweeting, which leads us to formulate the hypothesis:

Hypothesis 1e: The higher the level of positive emotions in a terror-related tweet, the more it is retweeted

As one might expect, proximity to the affected area of a terror attack increases the intensity of the emotions people experience during the aftermath (Smith et al., 2001). Initial inspection of the tweets in our data set containing location information showed that – contrary to what one might expect – tweets in the directly affected area (Boston and Massachusetts) are less emotional than tweets originating farther away. It could be that Bostonians are focusing on sharing valuable situational information rather than expressing how they feel (Mukkamala and Beck, 2016). Although in general lower emotion content is associated to lower information sharing (see previous section), we suspect that in the case of situationally relevant information, location plays a role equal of or bigger than emotions, leading us to formulate an additional hypothesis:

Hypothesis 2: Tweets from the affected area retweeted more than other tweets

3 Methods and Data

3.1 Data

The data in this study consists of Boston Marathon Bombing related tweets from during and after the event 15th – 23rd of April 2013. Non-English and other non-relevant tweets (e.g. related to Boston but not the bombing) were removed during the pre-processing. After careful consideration, we decided to exclude retweets from the data. Initial analysis revealed that retweets get significantly fewer – if any – retweets compared to the original post even when the original one has high retweet rates and the content is identical. This has the potential to severely confound the analysis of the relationship between the emotional content of a message and its probability of being shared onwards. Without a further look into the reasoning of users sharing retweets versus the originals upon encountering a retweet (which is out of scope here, but perhaps worthy of its own study), including the retweets' retweet rates in the data set might introduce a bias in the results, which is why we decided it to be prudent to exclude them from the analysis.

Our final data set thus consists of 4.4 million original tweets for which we extracted relevant metadata and counted the number of retweets. That dataset contained 93 000 tweets with geolocation information, which is around 2% of the full dataset. For those tweets, we extracted the coordinates, based on which we grouped them into four location categories:

(1) *"within the same city"* (within 30 km of the location of the bombing, which covers Boston as well as nearby areas such as Cambridge and Brookline), N=5 525

(2) *"not within the same city but within the same state"* (any coordinates outside of the first category but inside the state of Massachusetts), N=1 968

(3) *"not within the same state but within the same country"* (any coordinates outside of Massachusetts within the US), N=55 265

(4) *"abroad"* (everything outside of the US), N=30 338

We started out by getting an intuition of what our data set contains by examining it: the levels of each emotion it contains, how the retweet rates – our main interest – vary, and what kinds of messages are high in emotional intensity in general. In addition to statistical analysis, manual inspection of subsets was frequently used to confirm the observations. Tables 1, 2a, and 2b contain basic information of and emotion levels in the data. As could be expected, some of the high outliers in retweet numbers in the full data are not included in the geolocation set, which is also reflected in the variance of the retweets.

The overall tweeting density related to the bombing was markedly higher within the Boston and Massachusetts area with proportion to the 2013 population counts (0.11 tweets per citizen for Massachusetts including Boston, 0.02 tweets per citizen elsewhere in the US).

	Full set	Geo set
The number of retweets, mean	1,34	0,72
The number of retweets, variance	3 308	1 106
The number of retweets, maximum	65 294	8 762
Data set size	4 442 261	93 096

Table 1. The mean, variance, and maximum value for the number of retweets in each data set.

	Anger			Fear			Depression			Contempt		
	none	low	high	none	low	high	none	low	high	none	low	high
Boston	61.83	36.71	1.47	89.00	7.73	3.28	83.96	13.35	2.68	88.72	10.70	0.58
MA	60.42	37.35	2.24	90.50	6.66	2.84	83.11	15.15	2.74	89.74	9.56	0.71
US	50.75	47.02	2.23	88.26	7.97	3.77	80.20	15.91	3.89	86.03	12.93	1.04
Abroad	47.84	49.54	2.62	82.10	12.33	5.57	79.47	16.24	4.30	85.94	12.02	2.03
Full set	44.84	53.24	1.92	86.78	9.73	3.49	81.01	15.86	3.13	85.17	13.84	0.99

Table 2a. The percentages of levels of negative emotion in the full dataset of 4,4M tweets, and in the geolocation dataset of 93 000 tweets for each of the four location categories. SentiStrength scores 4-5 are combined in "high", scores 2-3 are combined in low, and score 1 is "none". Each tweet in the dataset is scored separately for each sentiment, which means it can simultaneously have a higher than 1 score on more than one emotion.

	Positive		
	none	low	high
Boston	81.92	18.06	0.02
MA	79.37	20.63	0.00
US	82.91	17.09	0.00
Abroad	84.72	15.28	0.00
Full set	87.42	12.57	0.00

Table 2b. The percentages of levels of positive emotion in the full dataset and the geolocation.

3.2 Sentiment Analysis

The sentiment was analysed using SentiStrength (Thelwall et al., 2010), a lexicon-based tool especially suited for short, informal texts. It assigns each unit of text – in this case tweet – a positivity (from 1 to 5) and negativity (from –1 to –5). Because we wanted to focus on differentiated emotions rather than polarity, we applied customized lexicons to detect affection, happiness, satisfaction, anger, fear, depression and contempt, each on a five-point scale. These lexicons were developed specifically to analyse different emotions in social media posts, and have been used and evaluated in previous research (Risius et al., 2015; Risius and Akolk, 2015). The emotional categories are based on Ekkekakis' hierarchical structure of the affective domain (Ekkekakis, 2013) (see Table 3).

We chose to use the custom lexicons rather than a more established approach because they cover differentiated emotions on an intensity scale in a way the existing and available sentiment analysis tools could not. However, just to be cautious, we decided to sanity check the quality against an established tool to the degree that is possible. Out of the established sentiment analysis tools, LIWC offers most insight beyond polarity, detecting positive sentiment, anger, anxiety, and sadness. We ran our geolocation data set through both LIWC and the custom lexicons to establish how often the two approaches agree on the presence of those four emotions.

Ekkekakis	**Risius et al.**	**Description**
Joy	Happiness	Amplified enthusiasm and excitement about attaining something desired or desirable
Love	Affection	Genuine fondness and liking attributed to a person or object
Pride	Satisfaction	Proud acknowledgement of and contentment with reaching a predetermined goal
Sadness	Depression	Impeding sadness evoked by an aversive event that may hinder activity
Anger	Anger	Animated animosity towards malice that can motivate rectification
Fear	Fear	Anticipatory horror or anxiety in unpredictable or potentially harmful situations
Shame	Contempt	Revulsion to something considered socially offensive or unpleasant

Table 3. Emotions in the hierarchical affective domain, their adaptation, and explanations for each emotion.

For positive sentiment, the agreement rate is 84%, anger is at 61%, fear/anxiety at 93%, and depression/sadness at 87%. As agreement on anger is clearly lower than the others, we looked at the cases where the anger lexicon and LIWC disagree. In 30 698 out of those 34 584 cases, our anger lexicon detected the mildest level of anger while LIWC detected none. A manual inspection of these "false positives" showed that some of them were false negatives for LIWC (*"They finally got the Boston Bomber! Now Execute Him!!!"*), and some of them false positives for the anger lexicon (*"After watching hours of CNN they caught the second bomber in the Boston Marathon #success #caughtinaboat"*). The anger lexicon is clearly more sensitive in detecting anger than LIWC, which is likely due to the fact that the anger lexicon is bigger than the other emotion lexicons (Risius et al., 2015). This is good to remember when interpreting the results, but all in all the tool comparison does not give us reason to suspect the custom lexicons are unsuitable for our purpose.

Preliminary examination of the data revealed that the levels of each of the negative emotions were lower in the proximity of the location of the terror attack than farther away, which is interesting considering that previous findings establish that after a terror attack, people living in the area exhibit stress and anxiety on a higher level than people with greater distance (see Table 2a).

3.3 Regression Analysis

We used regression analysis in order to examine the relationship between emotions and retweeting. The dependent variable is the *number of retweets* for each tweet in the data. Upon inspecting the correlation matrix, we established that the *positive emotions* were all highly correlated with each other, which lead us to decide to represent them using their mean as one variable instead of including them separately. No significant correlation was found between the negative emotions, so *anger, fear, depression*, and *sadness* were included in the model as separate variables. For the geolocation data set, the *location* is represented by a categorical variable denoting whether the tweet originated from Boston, elsewhere in Massachusetts, elsewhere in the US, or abroad. To account for other known effects in the data set, we use four control variables chosen based on their relevance in previous research: the *number of followers* of the author of the message, the activity of the author represented by the *number of messages* the author has previously posted, the number of *hashtags* in the tweet, and a binary variable for whether the tweet contains a *URL* (Stieglitz and Dang-Xuan, 2013).

The dependent variable – the number of retweets – in our dataset is count data consisting of non-negative integers, and is over-dispersed (the mean being significantly smaller than the variance for both the full and location specific datasets), which *generalized linear models* tend to be able to handle better than simpler models would. After examining model fits for different types of models (including quasi-Poisson estimation and negative binomial models), we determined the *negative binomial model* to be the most accurate model for our data set. In addition to the fixed independent variables, our final model contains a random variable to account for multiple tweets from the same user probably being more similar than tweets between users. Comparative tests including and excluding the random variable confirmed its inclusion to be a clear improvement to the model. The inclusion of a random variable meant using a mixed model, so our final choice was to go with a *generalized linear mixed model with a negative binomial distribution* with the following equation:

$\log(E(rt|*)) = \beta_0 + \beta_1 positive + \beta_2 anger + \beta_3 fear + \beta_4 depression + \beta_5 contempt + \beta_6 \log(followers) + \beta_7 \log(posts) + \beta_8 hashtags + \beta_9 url + \beta_{10}(1|userid)$

The equation for the geodata subset is the same except for the addition of a categorical variable for location information:

$\log(E(rt|*)) = \beta_0 + \beta_1 positive + \beta_2 anger + \beta_3 fear + \beta_4 depression + \beta_5 contempt + \beta_6 location + \beta_7 \log(followers) + \beta_8 \log(posts) + \beta_9 hashtags + \beta_{10} url + \beta_{11}(1|userid)$

where E(rt|*) is the expectation of the number of retweets given the right-hand side variables, location is a categorical variable for geolocation, and (1|userid) a random variable based on user ID numbers. The analysis of the model was run using the R package glmmTMB.

4 Results

The correlation matrices of the datasets confirm that correlation between independent variables is not an issue (see Tables 4 and 5). In both datasets, the highest correlation is <0.25 between the URL binary variable and posting history. The correlations between the variables of primary interest, the emotions, are all on a very low level.

	positive	anger	fear	depression	contempt	followers	posts	hashtags	url
positive	1								
anger	-0.04***	1							
fear	0.01***	0.04***	1						
depression	0.02***	-0.02***	0.04***	1					
contempt	0.00***	0.08***	0.03***	0.03***	1				
followers	-0.01***	0.00	0.00***	0.00***	0.00	1			
posts	-0.07***	0.04***	-0.01***	-0.04***	0.02***	0.06***	1		
hashtags	0.01***	-0.10***	-0.02***	-0.01***	-0.02***	0.00***	-0.01***	1	
url	-0.16***	0.09***	-0.05***	-0.06***	0.00**	0.03***	0.22***	-0.03***	1

*Table 4. Correlation matrix for the independent variables when analysing the full dataset. '***' < 0.001, '**' < 0.01, '*' < 0.05, '.' < 0.1*

	positive	anger	fear	depression	contempt	followers	posts	hashtags	url
positive	1								
anger	-0.03***	1							
fear	0.00	0.06***	1						
depression	-0.01*	0.02***	0.04***	1					
contempt	0.01*	0.06***	0.04***	0.04***	1				
followers	-0.01.	0.00	0.01**	0.00	0.00	1			
posts	-0.06***	0.06***	0.00	-0.01***	0.01***	0.04***	1		
hashtags	0.00	-0.13***	-0.03***	-0.03***	-0.03***	0.00	-0.10***	1	
url	-0.09***	0.02***	-0.05***	-0.04***	-0.01***	0.02***	0.24***	0.06***	1

*Table 5. Correlation matrix for the independent variables when analysing the geolocation dataset. '***' < 0.001, '**' < 0.01, '*' < 0.05, '.' < 0.1*

The results of the regression analyses are reported in Table 6. As is the case typically with regression on large datasets, the standard errors and p-values for the full data are all very small, and the results should not be overinterpreted. Column *exp(b)* in Table 6 lists the exponentiated versions of the coefficients (*b*) for the sake of legibility, as coefficients from a negative binomial model are in relation to the logarithm or the dependent variable rather than the actual values. This means that, for instance, for a one unit increase in positive emotions for the full dataset, the number of retweets is expected to increase by 1.03 times, i.e. 3% (b=0.03, exp(b)=1.03), assuming all other variables remain constant.

For each model, the table discloses two types of pseudo R^2 measures: the marginal pseudo R^2 describing the proportion of variance explained by the fixed effects, and the conditional pseudo R^2 describing the proportion of variance explained by both the fixed and random effects.

There are some differences in the levels of control variables, but the tendencies are similar. The impact of a user's follower count on the expected retweet rate is large and significant in both data sets. This stands to reason, as the number of followers directly impacts how many people are likely to see the tweet, which is a necessary precondition for sharing it onwards. The largest difference in the control variables between the datasets concerns the URL variable (0.84 in the geolocation data, 0.55 in the full data), which – against expectations based on previous literature – has a negative correlation with retweet rates.

The correlations for fear and contempt were significant and negative in both datasets, although the effect was stronger in the geolocation dataset. The negativity of the correlation leads us to reject hypothesis H1b for fear as it suggested a positive relation, and confirm H1d suggesting a negative relation. For the other negative emotions, anger and depression, the results were significant only in the full data set. The coefficients are rather small, suggesting that with the presence of anger or depression in a tweet, the number of retweets is expected to increase by 1%.

Results of the regression analysis for each dataset						
	Geo			**Full**		
Independent Variables	**b**	**SE**	**exp(b)**	**b**	**SE**	**exp(b)**
positive	-0.12***	0.03	0.87	0.03***	0.00	1.03
anger	0.02	0.01	1.02	0.01***	0.00	1.01
fear	-0.04**	0.01	0.96	-0.02***	0.00	0.98
depression	0.02	0.01	1.02	0.01**	0.00	1.01
contempt	-0.07***	0.02	0.93	-0.05***	0.00	0.95
log(followers)	0.68***	0.01	1.98	0.76***	0.00	2.14
log(posts)	-0.10***	0.01	0.99	-0.17***	0.00	0.84
hashtags	0.07***	0.01	1.07	0.11***	0.00	1.12
url	-0.18***	0.03	0.84	-0.59***	0.00	0.55
Constant	-4.85***	0.07		-4.82***	0.01	
Geo: Boston	0.45***	0.05	1.57			
Geo: MA	0.47***	0.07	1.60			
Geo: US	0.26***	0.02	1.30			
Pseudo R^2: marginal		0.18			0.26	
Pseudo R^2: conditional		0.42			0.53	
Number of observations		93 096			4 442 261	
p-values: '***' < 0.001, '**' < 0.01, '*' < 0.05, '.' < 0.1						

Table 6. The regression results for both the full dataset and the geolocation dataset. "b" is the coefficient resulting from the negative binomial model, "SE" is the standard error, "exp(b)" is the exponentiated coefficient allowing for easier interpretation

Hypotheses	Full data	Geodata
H1a: The higher the level of anger in a terror-related tweet, the more it is re-tweeted	confirmed	inconclusive
H1b: The higher the level of fear in a terror-related tweet, the more it is retweeted	rejected	rejected
H1c: The higher the level of depression in terror-related a tweet, the more it is retweeted	confirmed	inconclusive
H1d: The higher the level of contempt in a terror-related tweet, the less it is re-tweeted	confirmed	confirmed
H1e: The higher the level of positive emotions in a terror-related tweet, the more it is retweeted	confirmed	rejected
H2: Tweets from the affected area retweeted more than other tweets	N/A	confirmed

Table 7. The list of hypotheses outlined in the Theoretical Background section. Cases where the results did not have sufficient statistical significance are marked as inconclusive.

The effect of the geolocation variable is almost as large as the effect of the number of followers. The strong positive relation confirms H2. The regression analysis uses the fourth category *abroad* as a baseline, and the exponentiated coefficients report how much higher we expect the number of retweets to be if the tweet originates from another category. Tweeting from the Boston and Massachusetts area increase the expected retweet rate by a factor of 1.57 and 1.60 times respectively compared to tweets from abroad, and tweets from the US are also more likely to get retweeted than from outside the US.

Perhaps the most surprising result is that in the geolocation data, positive emotions are negatively correlated with retweet rates, while in the full dataset there is a small positive correlation. The negative relation in the geo set is the strongest correlation detected for any emotions in both datasets.

5 Discussion

Increasing our understanding on emotional drivers in user behaviour online is not only relevant from an academic standpoint, but also has some practical implications. In particular in the aftermath of a traumatic event, people seek out other people to exchange information, receive support, and make sense of what has happened. However, there are other motivations for using social media in a crisis context; several types of conversations unfold simultaneously on the same platform with different goals. Some aim to feel connected, others search for information and news either out of general curiosity or out of the need to ensure the wellbeing of others, and yet other actors monitor social media feeds to make sure they are updated on information relevant to their efforts. A better understanding of these conversations would enable more efficient real-time filtering for instance for emergency services or various authorities, but it would also allow organising the communication on social media to be better tailored for the users who are primarily seeking for connection and support.

Based on previous literature, we assumed that anger, fear, and depression would increase the number of retweets. Analyses on both datasets proved our assumptions wrong regarding fear, leading us to reject hypothesis H1b. For anger and depression, the correlations were either inconclusive (in the case of the geolocation dataset) or – in the case of the full dataset – very slightly positive, meaning that for the full data, hypotheses H1a and H1c are confirmed. Contempt was confirmed to be connected to retweeting, confirming hypothesis H1d. This allows us to conclude that analysing different emotions separately gives us better insight than treating all negative emotions as one feature.

One of the interesting questions arising from these unexpected results is why the positive relationship between elevated negative emotions and retweeting is not present in this case. It would seem reasonable to assume that when people are experiencing negative emotions, they also share them online, relate to other users' messages, and pass along what likeminded users have commented. On the other hand, if messages in the affected area were lower in emotion intensity and were shared more, it might suggest that retweeting is associated to passing on situational information and facts rather than engaging in conversations, making elevated negative emotions a distraction. Examples from the data illustrate that the informational value is relatively low in tweets with high levels of fear:

> *"Glued to the news. I honestly hope everyone I know in Boston is safe. This is absolutely horrific...."*
>
> *"Absolutely horrendous scenes in Boston! Dunno how people can be so evil! High alerts in London now Hope England not next crist! #Pray4Boston"*
>
> *"Watching this Boston explosions coverage. So fucking scary. Hope nothing happens at the London marathon"*

Tweets containing high levels of contempt are also typically expressions of personal feelings:

> *".@NateBell4AR Using the tragedy in Boston to deliver tasteless commentary on guns is horrible and very cruel to the victims. Shame on you!"*
>
> *"Our first thoughts (are) with the victims... This was a heinous and cowardly act. FBI investigating as "act of terrorism" - Obama #Boston"*
>
> *"Please pray for the people of Boston, we MUST protect our homeland, and FIND THOSE GUILTY!!!! Terror will NOT stand!"*

Conveying and obtaining factual information is a known motivational factor for people using social media in a crisis context; after 9/11 people attempted to reduce feelings of uncertainty by seeking information through various media (Boyle et al., 2004). If there is an increased interest in obtaining information, it could lead to increased information sharing behaviour, which might contribute to explaining why the results of this study differ from what has been reported previously. It would be interesting to establish whether increased information seeking and sharing is a general property of the unexpected conversation type, characterized by spiking rapidly after a real-world event, containing negative emotions, and quieting down quickly afterwards (Ferrara and Yang, 2015).

Geolocation has quite a strong impact on retweet rates; Boston and Massachusetts get way more retweets in spite of their on average lower emotion levels, and US gets more retweets than posts originating elsewhere. It could be that retweeting messages from users close to the location of the bombings is motivated by an urge to share and obtain timely information on events, casualties, and the possible arrest of the bomber. Tweeters close to the events may be considered to have more important things to say than people farther away. It is also noteworthy that the people most affected by the bombing are the ones exhibiting the least extreme levels of emotion, which is surprising considering that they should be experiencing emotion levels higher than those with more distance. This means we cannot draw a direct parallel between the intensity of emotion an individual's experiences and the degree of emotion in their online communication. Perhaps the strongest emotional reactions remain outside of social media but manifest in offline personal communications, or perhaps proximity to a crisis event means there is less time or opportunity to focus on one's own emotions soon after the event.

The area specific differences suggest that when researching online user behaviour, there is a benefit in considering a smaller dataset in order to include more detailed information of user behaviour based on location, especially when dealing with a real-life event where proximity to the event location may have a large influence on people's emotions, behaviour and interests.

Positive emotions provide the biggest source of surprise in the findings of this study. The results for the two data sets differed from each other rather essentially: with only the geotagged data included, positive emotions were associated with a decrease in the retweet rates against our hypothesis H1e, while in the full dataset positive emotions meant more retweets, which is in favour of H1e. This could mean that there is some degree of self-selection bias among the Twitter users who choose to disclose their location. Are the location disclosers somehow different from other users? Are their average followers expecting specific types of tweets? Is there a reason for their positive tweets to be considered less important to retweet on average? It could be that disclosing location is a tendency specific to certain areas more strongly than others, which could mean there is a specific cultural emphasis in the geotagged data compared to the full dataset.

In order to examine the relationship between positive emotion and retweeting more closely, perhaps further research should be conducted on a dataset richer in positive expression. Due to the topic in the dataset, even the tweets scoring high on positivity are typically lined with worry:

> *"Sending love to the victims of the Boston Marathon bombing"*
>
> *"Showing respect for my daddy's hometown. Thoughts and prayers go out to everyone in Boston."*
>
> *"Hope everyone in Boston is all good and safe now."*

This is, in part, a limitation of the lexical sentiment analysis tools used in this study; as long as positive words, such as love or respect are detected, the sentence gets a higher positive score regardless of the larger topical context.

These questions may have implications for researchers with respect to choosing between the completeness of a dataset versus narrowing it down in order to be able to include potentially relevant factors such as geolocation. This means scholars should be mindful of how to choose their data based on what compromises they are willing to make, and what their primary interest is.

6 Conclusions, Limitations, and Future Work

Previous work unanimously states that emotions play an important role in what we say and share online, and our study extends the understanding of how by examining the emotional content of retweets in the wake of a terrorist attack, focusing on five categories of emotions: positive emotions, anger, fear, depression, and contempt.

We found that different emotions are associated with behaviour in different ways; elevated levels of fear and contempt in a tweet make it less likely to be retweeted, while other negative emotions have a small positive correlation. When focusing the analysis only on tweets of users who disclose geolocation information, positive emotions in the tweet are associated with a decrease in the retweet rates, but when examined on a larger level, the effect is opposite – positive messages get more retweets. Considering geolocation data in analysing social media content has the potential to provide interesting additional insight, but there is a chance it may mean compromising some of the generalizability of the results. We also found that tweets originating in the affected area of a terror attack are clearly more retweeted than tweets from farther away.

Our theoretical contributions are adding to the understanding of the role of emotions in online information sharing in the context of a terror attack, and discovering that proximity to the location of a terror attack influences online behaviour both on the part of the person providing information online and the people assessing the relevance of said information. Findings from previous research may not generalise well in all contexts, and it seems like sometimes a neutral message carries farther than an emotional one.

Our findings also have more practical implications. Including geolocation information in social media analysis is potentially useful, but as narrowing down the dataset may impact the results, we recommend exploring and comparing the datasets in order to be able to make an informed decision while aware of the trade-offs.

Better understanding of what types of conversations unfold online in the wake of crises allows for more efficient filtering and searching real-time social media streams, which is helpful from a crisis management point of view. Considering that the tweets from the Boston and Massachusetts area were low on emotion and highly retweeted, it might be a feasible approach to access timely local information by filtering out high emotion messages which may be less likely to be passed on as useful by others, and more likely to be written by someone far from the location.

The results from the geotagged dataset should be interpreted carefully, keeping in mind that it is a 2% subset of the complete data. The dataset could be biased due to user self-selection in disclosing location, as it is voluntary, and it is likely that specific types of users go through that explicit effort.

The geographic categories in this study are used for examining the relationship between Twitter users' geographic proximity to a terror attack and their emotional expression. In particular the category "abroad" covers a heterogenous group of users around the world and it is fairly likely to contain a wide range of cultural diversity which may have an essential impact on emotional self-expression online. We decided that accounting for that falls outside of the scope of this study, but it would be interesting to look into in the future on a more fine-grained level.

The results of this study, as usual, raise further questions relevant for future research. The motivations for using social media vary depending on the context, and several types of motivations are likely to exist in any given context. Different motivations lead to different information sharing behaviour, and being able to account for more than one of them at a time when analysing online discussions would enable a deeper understanding of them. Potential future avenues of research could include investigating the levels of different emotions over time, and looking more closely into dominant topics of discussion in order to better understand the collective online dynamics that follow a traumatic event.

References

Berger, J. (2011) 'Arousal Increases Social Transmission of Information', *Psychological Science* 22(7): 891–893.

Berger, J. and K.L. Milkman (2012) 'What Makes Online Content Viral?', *Journal of Marketing Research* 49(2): 192–205.

Boyle, M.P., M. Schmierbach, C.L. Armstrong, D.M. McLeod, D.V. Shah, and Z. Pan (2004) 'Information Seeking and Emotional Reactions to the September 11 Terrorist Attacks', *Journalism & Mass Communication Quarterly* 81(1): 155–67.

Burnap, P. et al. (2014) 'Tweeting the Terror: Modelling the Social Media Reaction to the Woolwich Terrorist Attack', *Social Network Analysis and Mining* 4(1): 206.

Chen, K., P. Luo, and H. Wang (2017) 'An Influence Framework on Product Word-of-Mouth (WoM) Measurement', *Information & Management* 54(2): 228–240.

Dickinger, A., M. Arami, and D. Meyer (2008) 'The Role of Perceived Enjoyment and Social Norm in the Adoption of Technology with Network Externalities', *European Journal of Information Systems* 17(1): 4–11.

Eismann, K., O. Posegga, and K. Fischbach (2016) 'Collective Behaviour, Social Media, and Disasters: A Systematic Literature Review', *ECIS*.

Ekkekakis, P. (2013) *The Measurement of Affect, Mood, and Emotion: A Guide for Health-Behavioral Research*. Cambridge University Press.

Ekman, P. (1992) 'An Argument for Basic Emotions', *Cognition & Emotion* 6(3–4): 169–200.

Ellison, N.B., C. Steinfield, and C. Lampe (2007) 'The Benefits of Facebook "Friends:" Social Capital and College Students' Use of Online Social Network Sites', *Journal of Computer-Mediated Communication* 12(4): 1143–1168.

Ferrara, E. and Z. Yang (2015) 'Quantifying the Effect of Sentiment on Information Diffusion in Social Media', *PeerJ Computer Science* 1: e26.

Fischer, D., K. Eismann, and K. Fischbach (2016) 'Usage Behavior of Social Network Sites in the Aftermath of Terrorist Attacks'.

Fowler, J.H. and N.A. Christakis (2008) 'Dynamic Spread of Happiness in a Large Social Network: Longitudinal Analysis over 20 Years in the Framingham Heart Study', *Bmj* 337: a2338.

Fredrickson, B.L. (1998) 'What Good Are Positive Emotions?', *Review of General Psychology* 2(3): 300–319.

Gruzd, A. (2013) 'Emotions in the Twitterverse and Implications for User Interface Design', *AIS Transactions on Human-Computer Interaction* 5(1): 42–56.

Gruzd, A., S. Doiron, and P. Mai (2011) 'Is Happiness Contagious Online? A Case of Twitter and the 2010 Winter Olympics', *System Sciences (HICSS), 2011 44th Hawaii International Conference On*. IEEE.

Hancock, J.T., K. Gee, K. Ciaccio, and J.M.-H. Lin (2008) 'I'm Sad You're Sad: Emotional Contagion in CMC', *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*. ACM.

Hansen, L.K., A. Arvidsson, F. \AArup Nielsen, E. Colleoni, and M. Etter (2011) 'Good Friends, Bad News-Affect and Virality in Twitter', *Future Information Technology*, pp. 34–43. Springer.

Huang, C.-M., E. Chan, and A.A. Hyder (2010) 'Web 20 and Internet Social Networking: A New Tool for Disaster Management?-Lessons from Taiwan', *BMC Medical Informatics and Decision Making* 10(1): 57.

Hughes, A.L., L. Palen, J. Sutton, S.B. Liu, and S. Vieweg (2008) 'Site-Seeing in Disaster: An Examination of on-Line Social Convergence', *Proceedings of the 5th International ISCRAM Conference*. Citeseer.

Hyvärinen, H. and R. Beck (2018) 'Emotions Trump Facts: The Role of Emotions in on Social Media: A Literature Review', *Proceedings of the 51st Hawaii International Conference on System Sciences*.

Kaufmann, M. (2015) 'Resilience 20: Social Media Use and (Self-) Care during the 2011 Norway Attacks', *Media, Culture & Society* 37(7): 972–987.

Kim, K., O. Park, S. Yun, and H. Yun (2017) 'What Makes Tourists Feel Negatively about Tourism Destinations? Application of Hybrid Text Mining Methodology to Smart Destination Management', *Technological Forecasting and Social Change* 123: 362–369.

Kramer, A.D., J.E. Guillory, and J.T. Hancock (2014) 'Experimental Evidence of Massive-Scale Emotional Contagion through Social Networks', *Proceedings of the National Academy of Sciences*: 201320040.

Kwon, K.H. and A. Gruzd (2017) 'Is Aggression Contagious Online? A Case of Swearing on Donald Trump's Campaign Videos on Youtube'.

Lerner, J.S., R.M. Gonzalez, D.A. Small, and B. Fischhoff (2003) 'Effects of Fear and Anger on Perceived Risks of Terrorism: A National Field Experiment', *Psychological Science* 14(2): 144–150.

Lin, K.-Y. and H.-P. Lu (2011) 'Why People Use Social Networking Sites: An Empirical Study Integrating Network Externalities and Motivation Theory', *Computers in Human Behavior* 27(3): 1152–1161.

Liu, Y., C. Kliman-Silver, and A. Mislove (2014) 'The Tweets They Are A-Changin: Evolution of Twitter Users and Behavior', *ICWSM* 30: 5–314.

Mazer, J.P., B. Thompson, J. Cherry, M. Russell, H.J. Payne, E.G. Kirby, and W. Pfohl (2015) 'Communication in the Face of a School Crisis: Examining the Volume and Content of Social Media Mentions during Active Shooter Incidents', *Computers in Human Behavior* 53: 238–248.

Morrison, J.L., Sandra, and H. Berry (2001) *A National Survey of Stress Reactions After the September 11, 2001, Terrorist Attacks*.

Mukkamala, A. and R. Beck (2016) 'Enhancing Disaster Management through Social Media Analytics to Develop Situation Awareness What Can Be Learned from Twitter Messages about Hurricane Sandy?', *PACIS*.

Neubaum, G., L. Rösner, A.M. Rosenthal-von der Pütten, and N.C. Krämer (2014) 'Psychosocial Functions of Social Media Usage in a Disaster Situation: A Multi-Methodological Approach', *Computers in Human Behavior* 34: 28–38.

Oh, O., M. Agrawal, and H.R. Rao (2013) 'Community Intelligence and Social Media Services: A Rumor Theoretic Analysis of Tweets during Social Crises', *Mis Quarterly* 37(2).

Pang, B. and L. Lee (2008) 'Opinion Mining and Sentiment Analysis', *Foundations and Trends® in Information Retrieval* 2(1–2): 1–135.

Pfeffer, J., T. Zorbach, and K.M. Carley (2014) 'Understanding Online Firestorms: Negative Word-of-Mouth Dynamics in Social Media Networks', *Journal of Marketing Communications* 20(1–2): 117–128.

Relling, M., O. Schnittka, H. Sattler, and M. Johnen (2016) 'Each Can Help or Hurt: Negative and Positive Word of Mouth in Social Network Brand Communities', *International Journal of Research in Marketing* 33(1): 42–58.

Rimé, B. (2009) 'Emotion Elicits the Social Sharing of Emotion: Theory and Empirical Review', *Emotion Review* 1(1): 60–85.

Risius, M. and F. Akolk (2015) 'Differentiated Sentiment Analysis of Corporate Social Media Accounts', *AMCIS*.

Risius, M., F. Akolk, and R. Beck (2015) 'Differential Emotions and the Stock Market-The Case of Company-Specific Trading', *ECIS*.

Risius, M. and R. Beck (2015) 'Effectiveness of Corporate Social Media Activities in Increasing Relational Outcomes', *Information & Management* 52(7): 824–839.

Russell, J.A. (2003) 'Core Affect and the Psychological Construction of Emotion', *Psychological Review* 110(1): 145.

Schumaker, R.P., A.T. Jarmoszko, and C.S. Labedz Jr (2016) 'Predicting Wins and Spread in the Premier League Using a Sentiment Analysis of Twitter', *Decision Support Systems* 88: 76–84.

Smith, T.W., K.A. Rasinski, and M. Toce (2001) 'America Rebounds: A National Study of Public Response to the September 11th Terrorist Attacks', *Chicago, IL: National Opinion Research Center at the University of Chicago*.

Stieglitz, S. and L. Dang-Xuan (2013) 'Emotions and Information Diffusion in Social Media—Sentiment of Microblogs and Sharing Behavior', *Journal of Management Information Systems* 29(4): 217–248.

Thelwall, M., K. Buckley, G. Paltoglou, D. Cai, and A. Kappas (2010) 'Sentiment Strength Detection in Short Informal Text', *Journal of the American Society for Information Science and Technology* 61(12): 2544–2558.