

Postprint Version	1.0
Journal website	<a href="http://www.sciencedirect.com/science/article/pii/S074756321630084X">http://www.sciencedirect.com/science/article/pii/S074756321630084X</a>
Pubmed link	
DOI	

This is a NIVEL certified Post Print, more info at <http://www.nivel.eu>

# Self-correcting mechanisms and echo-effects in social media: An analysis of the “gunman in the newsroom” crisis

WOUTER JONG <sup>A, B, \*</sup>, MICHEL L.A. DÜCKERS <sup>C, D</sup>

a INTERVICT, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands

b Dutch Association of Mayors (Nederlands Genootschap van Burgemeesters), Nassaulaan 12, 2514 JS The Hague, The Netherlands

c Impact - National Knowledge and Advice Centre for Psychosocial Care Concerning Critical Incidents, Partner in Arq Psychotrauma Expert Group, Nienoord 5, 1112 XE Diemen, The Netherlands

d NIVEL - Netherlands Institute for Health Services Research, PO Box 1568, 3500BN Utrecht, The Netherlands

## ABSTRACT

The positive and negative effects of social media in crises are currently receiving an increased amount of scholarly attention. This study focuses on Twitter users in the context of a crisis in the Netherlands on January 29, 2015. After having made a bomb threat, an armed man managed to get access to the national news broadcasting station around 8 pm, where he demanded airplay to share “an important message” with Dutch citizens. Three weeks after the terrorist attack on Charlie Hebdo in Paris, approximately 1.5 million viewers were anxious that a similar attack was taking place in the television studio. The crisis, also followed by social media users, reached a climax when armed policemen arrested the man, which was later shown on national TV. We analyzed 58,931 tweets, posted in the six hours after the incident. By examining shared facts and rumors during the gunman crisis, we identified an “echo-effect”: the dissemination of older tweets continued after the posting of new facts by the same source. Moreover, we found that two rumors were based on misinterpreted humor in Twitter messages. The study adds insight into the self-correcting mechanism of social media communities when verifying and dispelling online rumors during crises.

## 1. INTRODUCTION

The emergence of social media over the last decade has enormous potential to influence information-sharing within societies worldwide. The amount of information that is being shared through social media has increased and is being

shared at a pace and a scale that until recently was difficult to imagine. Little is known about the function of social media in crises (Binder, 2012).

According to Alexander social media are used in seven ways: listening to public debate; monitoring situations; extending emergency response and management; crowd-sourcing and collaborative development; creating cohesion; furthering causes (including charitable donations); enhancing research (Alexander, 2014). Alexander has also highlighted two potential negative side effects of social media. They can, on the one hand, undermine authority and promote terrorist acts, and, on the other hand, disseminate rumors. It is the dissemination of rumors that plays a key role in the current study based on social media traffic during the “gunman in the newsroom crisis” in the Netherlands (hereafter “gunman crisis”).

So far, much of the evidence-based work on crisis communications is based on Coombs' crisis response strategies which protect an organization's reputation during a crisis, represented by what it says and does when a crisis has occurred (Coombs, 2004). This line of reasoning has also been adopted with regard to social media (Ki & Nekmat, 2014; Snoeijers, Poels, & Nicolay, 2014). However, where communication efforts can be described in terms of “webcare” and brand reputation, some crisis situations can be regarded as crises where authorities have other communication goals than image repair alone (Jong, Dückers, & van der Velden, 2016). Research about how social media evolve in the lifecycle of a crisis is scarce (Pang & Ng, 2016). In order to gain a better understanding of the extent to which authorities can exert more or less influence on social media communication, it is necessary to gain insight into the communication mechanisms of online communities during particular events.

### **1.1. The gunman crisis**

In the Netherlands there are currently (2015) approximately 1 million daily Twitter users. In this study, we analyzed the tweets in messages sent on the Twitter network about a crisis situation in the Netherlands on the evening of January 29, 2015. On that day, a man wielding a fake gun entered the studio of the Dutch public news broadcasting station NOS. The situation caused the main evening bulletins to be canceled. The man, smartly dressed and carrying what appeared to be a long pistol, paced around a studio, demanding airtime. Armed police stormed into the studio and overpowered him.

During the disruption, the TV channel did not broadcast the regular evening news. As such, Twitter users were wondering what had caused the disruption, as the TV news had not been disrupted since 1961. Three weeks after the attack on Charlie Hebdo in Paris, viewers were anxious that a similar attack was taking place in the television studio. The situation in the television studio was intensively discussed on Twitter, which functioned as a “second screen” for many viewers in the Netherlands. In the hours after the incident, two rumors were persistent on the social network. The first rumor was that an identical situation took place in Belgium, involving the evening news broadcast by the Flemish Radio- and Television Network Organization (VRT). When the identity of the gunman became known on Twitter, a second rumor started to develop, suggesting that both his parents were allegedly killed in the MH17 tragedy; an airplane crash in the Ukraine in July 2014 with 298 casualties. Both rumors turned out not to be true.

We will explore the use of Twitter as a social medium during the disruption in The Netherlands. This type of analysis is welcome as the development of theoretical

considerations is still meager in this area of research (Takahashi, Tandoc Jr, & Carmichael, 2015) and in order to validate earlier research (Groenendaal, De Bas, & Helsloot, 2013). Internationally, Twitter is a widely used social media platform and is a popular second screen (Yu & Wang, 2015). Users share their real-time reactions and emotions in short tweets (Ji & Raney, 2015; Wang, 2013). Especially during crises, when the appetite for information can be regarded as high, the channel grows in use for informational purposes (Westerman, Spence, & Van Der Heide, 2014). That said, Graham, Avery, and Park (2015) found that the use of social media depends on the type of crises, and that social media were used significantly more for crisis communication during public health crises than for natural disaster, transportation, political, social, or criminal crises. The information spreads fast because it is easily shared by people tend to share the information about crises easily. On the other hand, Twitter is not only known to enable the effective broadcasting of valid news, but also of baseless rumors (Mendoza, Poblete, & Castillo, 2010). Research from the Fukushima incident in Japan tends to show that within this flow of information, false rumors that began to circulate appear to be easily corrected by knowledgeable people (Alexander, 2014). An exploration of the online behavior and rumors during the incident of January 29 might provide us with more insight into this “self-correcting”- mechanism. This will be helpful as it will broaden our scope in terms of crisis communications and the involvement of the general public during crises.

## 1.2. Objective

Our objective is to gain a better understanding of the mechanisms within social media communities in an acute crisis context.

Based on this objective, we present the following research question, inspired by earlier research suggesting an important role for traditional media channels in times of uncertainty (Binder, 2012; Utz, Schultz, & Glocka, 2013).

RQ1. *How did “breaking news” broadcast by news agencies spread on Twitter?*

Shibutani noted that in ambiguous situations, people tend to act as pragmatic problem-solvers by combining their knowledge and information including data, bits of information from self and others, speculation, and interpretations through careful deliberation to construct a shared understanding of the situation in question and thus allowing considered concerted action (Miller, 2005; Shibutani, 1966). As the online problem-solving might work both ways (either construct a rumor or debunk a rumor), we also ask:

RQ2. *Regarding two rumors - concerning a similar course of events at the VRT-studio and the MH17-connection - shared on Twitter during and after the crisis, how were they constructed and were efforts taken within the Twitter community to correct and dispel myths?*

In a study of 573,000 lines of pager messages during the events of September 11, 2001 (Back, Küfner, & Egloff, 2010), the authors found that words which expressed anxiety dropped soon after the events. According to them, the immediate recovery from anxiety might be explained by the lessening of uncertainty shortly after each event, as a result of the spread of information. These findings lead us to our last research question:

RQ3. *Can we identify an online change in mood during the events on the evening of January 29?*

## 2. METHOD

We analyzed content shared on Twitter about this incident. We collected 58,931 tweets which were posted online between 7.40 pm on January 29, 2015 and 2.00 am on January 30, 2015.

Tweets were retrieved using Coosto software ([www.coosto.nl](http://www.coosto.nl)), a widely used tool suited to webcare and qualitative online data analysis. The software enabled us to export the tweets captured by using keywords in any given period. Tweets were searched and collected based on the following keywords, not necessarily hashtags: NOS (the name of the public broadcasting station), gijzeling (kidnapping), as well as keywords referring to rumors, the kidnapper, his arrest and/or his background. As the crisis can be considered a “developing story”, new keywords (such as the name, residence and education of the gunman) became relevant as new facts became known to the public. We could capture a massive set of 58,931 tweets from 29,165 individual users (a convenience sample) and used it to explore the Twitter stream during the gunman crisis. Regarding two of the rumors we found in our database, we contacted the Twitter users who posted influential messages. A complete list of all the identified tweets (in Dutch) can be obtained from the first author.

In the sample, we distinguish tweets, retweets and mentions. A tweet is an original remark posted online by a Twitter user. A retweet is similar to a tweet, but consists of a message from a person (user I), which has been copied and resent by another Twitter user (user II). From the 58,931 tweets we retrieved, almost 64,9% can be regarded as retweets (38,248). Both tweets and retweets can be read by all followers in a Twitter timeline. Mentions are messages which start with an “@” and are targeted at a specific user. When “user A” posts a Twitter message starting with “@userB”, only the followers of both user A and user B find this interaction in their timeline. In our database, 2,044 tweets are mentions (see Fig. 1).

### [FIGURE 1]

To independently check the timing of the developments in the crisis, we used an evaluation report by COT (2015) as a reference for our timelines. Based on this report, the developments are as follows: 7.40 pm The gunman enters the studio complex.

7.53 pm The gunman is taken to studio 10.

8.00 pm Time the evening news usually starts.

8.00 pm Message is brought to the viewers at home: “Due to circumstances, no broadcast is available at this time”.

8.14 pm the gunman is arrested by local police. The message is still shown on the television screens nationwide.

8.20 pm The arrest of the gunman becomes known to the outside world.

8.26 pm Commercial television channel RTL does a “breaking news” item, which is followed by regular updates.

8.38 pm First picture of the gunman and threatening letter appear online and are shown by other news stations as well.

9.05 pm NOS restarts its live broadcasts from alternative studio in The Hague.

### 3. RESULTS

#### 3.1. Sub-question 1: how did “breaking news” broadcast by news agencies spread on Twitter?

In the first few minutes, the cause of the disruption is unknown to the viewers at home. Few people (32) tweet about the disruption.

Twitter users start to speculate about the causes, oftentimes with humor (such as “message with highest viewing rates ever”).

At 8.07 pm, the Dutch newswire ANP brings the news “STUDIO COMPLEX NOS IN HILVERSUM BEING EVACUATED. MAN WITH GUN DEMANDS AIRTIME (SOURCE: NOS)”. Up to 8.14 pm, the time of the arrest, this message is retweeted 2.628 times. Of all the tweets connected to the disruption at NOS (4031 tweets) and posted between 7.40 pm and 8.14 pm, almost two thirds of the tweets are retweets of this breaking news.

At 8.14 pm the gunman is arrested. However, this fact does not become known to the outside world, till user @stevendalebout, working for NOS but not at the scene, tweets “Understand from editors that the man was overpowered and taken away” at 8.20 pm.<sup>1</sup> Shortly after, ANP gives a breaking news update of its news at 8.20 pm: “GUNMAN BEING ARRESTED AND TAKEN AWAY”.

This news update gets retweeted right away. While this update makes the older news obsolete, the old news (“gunman demands airtime”) is regularly retweeted until 10.15 pm. The last retweet in our dataset can be found at 11.56 pm by user @alien04268326.<sup>2</sup> Apparently, the old news echoes, even when it is updated by the same source (Fig. 2).

Influencing the sharing of obsolete news can, technically speaking, be done in three ways. First of all, the sender can delete his or her tweet, thereby automatically removing all potential retweets. When users want to correct other users' tweets, they can tweet the update with a general remark, which can be read by all a user's followers on Twitter. This can also be done by directly informing the Twitter user who spreads the old news with a so-called “mention”. This pinpoints the message to the one who spreads the news. In this case, we find only five tweets in which a follower informs a Twitter user directly (with a mention) that the message he or she was spreading is obsolete. The majority of users seem to ignore the people who resend obsolete information and do not look for direct interaction with these senders. We also looked into the spread of information with respect to the threatening letter the gunman passed to Dutch TV staff (Fig. 3).

User @ibhaarsma was the first one who posted the letter online.

Shortly afterwards, it was retweeted by several journalists. At 8:53 pm an improved and more readable version of the letter was posted online. The tweet with the letter reached a maximum of 356 retweets per minute (8.57 pm). All in all, nearly 8.9% of all tweets in our database (5228 out of 58,931) referred to the threatening letter.

Regarding the tweets from the authorities, the Dutch police use Twitter account @politieDGV as its communications channel to inform the public and to confirm breaking news. Its first tweet (8.49 pm) confirms that the gunman has been arrested. Even though the tweet is posted 29 min after the breaking news update from the ANP news wire, it gets retweeted 204 times. The mayor of Hilversum's office (@gemeenteHsum) refers to @politieDGV for updates (9.24 pm). On its own account, it sends an announcement about a press conference (10.22 pm) and a request to stay away from the studio complex of NOS (10.27 pm). The mayors' own Twitter

account (@pieterbroertjes) is not used in the aftermath of the events. All in all, we find 81 tweets referring to the mayor; ranging from announcements about a press conference to questions about his responsibility for security measures on the premises.

### **3.2. Regarding two rumors shared on Twitter during and after the crisis, how were they constructed and were efforts taken within the Twitter community to correct and dispel myths?**

We also looked into two rumors that were posted online during and after the events in the studio. The first rumor appearing online mentioned the suggestion that a gunman entered the studios of Flemish VRT at the same time. It was suggested that the situation in the Dutch and Flemish studios were connected to each other. We refer to this as the VRT-rumor. The second rumor consisted of the claim that both the gunman's parents died in the MH17 airplane crash. We refer to this as the MH17-rumor. We found 599 tweets relevant to the VRT-rumor and 247 relevant to the MH17 rumor.

[FIGURE 2]

[FIGURE 3]

The VRT-rumor starts with a tweet by a Dutch journalist @lavieJanRoos (approx. 30,000 followers). Around 8.04 pm he tweets sarcastically that this broadcast is the most independent evening news ever, as it is not broadcasted at all. At 8.07 pm, it becomes known that a gunman claims airtime. Unaware of this “breaking news”, Flemish Twitter user @koendelaey replies to @lavieJanRoos with the remark “The same goes for Belgian VRT”<sup>3</sup> at 8.11 pm.<sup>4</sup> He includes - with similar intentional sarcasm - a picture of a blank VRT screen, retrieved through Google Images.<sup>5</sup> Seconds afterwards, @lavieJanRoos retweets @koendelaey's tweet, thereby implying that “the same goes for Belgian VRT”. As Twitter users already know about the gunman in the Dutch studio, the retweet from @lavieJanRoos is interpreted immediately as confirmation that a similar situation is going on in Brussels' VRT studios. Tweets, both with stills from NOS and VRT are tweeted online, as “proof” that both the Dutch as well as the Flemish evening news have been disrupted by an armed man. As soon as the VRT rumor starts to spread, user @koendelaey tries to explain what happens. His tweets have hardly any impact, as he only sends “mentions” and his general remarks are not retweeted by @lavieJanRoos nor others. How did the rumor turn out to be untrue? First, we selected all the tweets which mentioned VRT (either with or without hashtag), resulting in a total set of 726 tweets. Next, we divided the tweets into four categories. When VRT was mentioned without a link to the rumor, for example “#elinedemunck appears in TV program #reyerslaat. Showcase for low quality at #vrt.”, tweets were labeled as irrelevant (n ¼ 34). In other cases people questioned whether or not the rumor was true. These tweets were labeled “neutral” (n ¼ 27). Tweets confirming the rumor were labeled as “true” (n ¼ 588).

Tweets which denied the rumor were labeled as “false” (n ¼ 77).

Tweets in this last category mention several reasons why the rumor is “false”. Either the time of the broadcast (VRT usually broadcasts at 7.00 pm), the image @koendelaey posted which turned out to be an existing and not original picture, the

outdated house style of the image (the logo on the image was a former VRT logo) and the fact that people were still able to watch VRT. The rumor is initially dispelled around 8.18 pm, when a link is posted to tineye.com, which proves the uploaded image is fake. The most influential discrediting of the rumor comes online ten minutes later (8.30 pm), when former NOS-anchorwoman @SachadeBoer asks fellow Twitter users to stop spreading the rumor.<sup>6</sup> The rumor gets a final impulse at 8.36 pm, when weather woman @helgavanleur (approx. 90,000 followers) from commercial RTL television station retweets the rumor. She receives several mentions from Twitter users and deletes her tweet at 8.41 pm, thereby making it impossible for her followers to further retweet it. Also, this action removes all earlier retweets. In Fig. 4, we show the tweets mentioning the rumor “true” versus tweets mentioning the rumor is “false” through time.

The second rumor online stated that the gunman's parents died in the MH17 crash, an airplane crash over Ukraine in which 193 Dutch nationals lost their lives. The first tweets by @Ganzendijker and @TommyGunner state that the gunman's parents are said to be dead (the cause of their death is not mentioned by them). According to Twitter user @essiel309, the origin of the rumor that both his parents died can be found on the Facebook page of a local newspaper, “Telstar”, from the gunman's hometown of Pijnacker. She added the “MH17” link herself after reading the rumor on the Facebook page. At 11.34 pm, her tweet appears which states clearly that the gunman lost both his parents in the MH17 disaster. According to @essiel309, her tweet was meant ironically.<sup>7</sup> The irony is taken literally and her tweet gets retweeted. She knew both parents and knew they were still alive.

The impact of the MH17-rumor is lower compared to the VRT rumor; at its maximum, we found only 5e6 tweets per minute for either true or false, where the overall number of tweets after 11.30 pm is on average around 30 tweets per minute. All in all, the rumor lasts until 2.00 am when our database stops.

### **3.3. Can we identify an online change in mood during the events on the evening on January 29?**

Until the moment it becomes clear why there is no broadcast of the evening news, users of Twitter are posting jokes about the possible reasons, starting with the earlier mentioned sarcastic tweet by @lavieJanRoos that this broadcast is the most independent evening news ever, to “testing image with the highest viewing ratings ever”. As soon as it becomes clear that a gunman entered the TV studio, the mood changes. As shown in Fig. 4, we looked for tweets which used explicit emotional language, such as the acronyms WTF and OMG and Dutch equivalents. Tweets with explicit emotional expressions started at 8.07 pm and dropped substantially after 8.20 pm, when the gunman was arrested (Fig. 5). The later increase, around 8.41 pm, occurred when cartoonist @fokkesukke (55,000 followers) posts a tweet, containing a cartoon with explicit language.

Comedian @freekjonge posts nine jokes online between 8.07 pm and 8.20 pm. He gets several replies from other users who tell him his jokes are inappropriate. The mood seems to change again after 8.20 pm, when joking seems to be “allowed” again. The disruption continues and users who looked forward to the scheduled broadcast of “The Mole” e a popular reality television game show - start to ask ironically if the gunman might be one of the participants in this TV show.

Around 8.38 pm, a picture of the gunman and a picture of his threatening letter are published on Twitter (with a follow-up around 8.52 pm with pictures of higher

quality). It is the start of a collective act of fact-checking. In contrast to the initial emotional response, the behavior changes in an online hunt for information about the gunman. At 9.31 pm, the gunman is identified as “Tarik Z.” (full name shared online) for the first time. One minute later, at 9.32 pm, his hometown “Pijnacker” appears online for the first time. News updates follow quickly. Images and background information from his Facebook page is shared on Twitter at 9.32 pm as well. More information (study, town of birth, age) is shared in the hour afterwards. In terms of sense-making, we also looked for tweets which referred explicitly to the events at Charlie Hebdo. In their press conference (10.30 pm), the mayor and public prosecutor referred to the earlier events at Charlie Hebdo in Paris, just three weeks before to emphasize the impact of the situation. References in tweets to “Charlie Hebdo” and/or “#Jesuischarlie” might have been an indication of the impact and online perception of the events. Only 24 out of the 58,931 tweets refer to Charlie Hebdo. Most of these tweets relate a supposed lack of security measures at the television studio premises to the events at Charlie Hebdo. It is not used to describe the impact of the event as such.

#### 4. DISCUSSION

We retrieved 58,931 tweets over a relatively short period of time. At certain points, we found instances with 900 tweets per minute (or 15 tweets per second). Apparently the gunman crisis attracted a considerable amount of attention from social media users. It remains difficult to quantify the relative impact. What we do know is that the tweets-time-ratio is higher than reported in another study. Lachlan, Spence, Lin, Najarian, and Del Greco (2016) found rates between 2 and 5 tweets per second (or 120 to 300 tweets per minute) sent during a severe winter storm. Studies like these contribute to the knowledge about the active participation of people who want to share news and insights with their followers, even when they are remote onlookers, similar to the findings of Pang and Ng (2016). In our case Twitter users behaved as a watchman for the crisis, giving an insight into the developments, the gunman's background and attempts to fact-check online rumors.

Below, we highlight and discuss our main findings, implications and limitations.

##### 4.1. Authority and trust

[FIGURE 4]

[FIGURE 5]

Looking back we see that messages from senders with a certain level of authority and trust play an important role in the sharing of social media messages. Twitter users with many followers have a leading status within the network. It is a status they earned, for instance, based on particular expertise (e.g. as a journalist) or for being engaged in public debate in an authentic or intelligent way.

The tweets by respected journalists, experts, politicians, and even pop artists, actors and comedians are likely to be picked up by followers, as, for example, tweets from @lavieJanRoos (journalist, approx. 30,000 followers), @freakjonge (comedian, approx. 40,000 followers), @SachadeBoer (former anchor woman, approx. 50,000 followers) and @helgavanleur (weather woman, approx. 90,000 followers) showed. They are extra careful with sharing obsolete information, which is a safety valve

against the sharing of incorrect messages within social media communities. Indeed, apart from news being tweeted by media organizations, tweets from individual journalists were rapidly retweeted. Users @jiscacohen (breaking news about gunman), @stevendalebout (breaking news about arrest) and @ibhaarsma (breaking news with picture of threatening letter) were all heavily retweeted since they were regarded as trusted news sources.

#### **4.2. Echo-effect of news on Twitter**

In the shared news updates we observed a certain echo-effect: obsolete news is retweeted regardless of the availability of new facts. From 8.20 pm, tweets about the gunman claiming airtime as well as tweets about the arrest of the gunman are shared online.

Even though these messages contradict each other, the older news gets even more retweets after 8.34 pm. However, both news items fade out as more updates from other angles appear online. Tweets with a picture of the letter the gunman brought to the studio come online from 8.38 pm (“This is the letter the gunman showed to personnel”). It is retweeted 3862 times during the first half an hour.

The echo-effect diminishes as time evolves. The authorities were not really involved online. Police (@politiedgv) confirmed the arrest of the gunman 29 minutes after it was made public by newswire ANP and shared online. Their tweet does not seem to have had a substantial effect on diminishing the echo-effect.

Although it might be surprising, the tweet by the police at 8.49 pm did not seem to have an effect on the sharing of older facts (Nu.nl: “gunman claims airtime”). It is in line with an earlier finding: government information gets lost in citizen information (Groenendaal et al., 2013). The impact of the police tweet would probably have been higher if it had been tweeted directly after the arrest. On the other hand, when the newswire uses police information as a recognizable source for its news items, the added value of a distinct police tweet is relatively low anyway.

Users on Twitter kept retweeting the obsolete news afterwards.

The reason for retweeting obsolete news can be twofold. Either the user who retweets it is not aware that the news has been updated in the meantime. It can also be retweeted because some users believe it is still worth spreading as it is part of the news story, even if it is known that the gunman has been arrested and taken into custody. However, if the reason behind a retweet is not explained, other users might get the wrong end of the stick if they cannot put the tweet in the right perspective. Regardless of the motivation of users to retweet outdated news when new facts are known, it does not matter if they do this because they consider a topic interesting or for other reasons. After all, the effect is still that they contribute to an echo-effect regardless of their motivation.

The echo-effect can be studied in the context of Rogers' adoption curve (Rogers, 2003). Besides an early group of “innovators”, there seem to be groups that are quite late in sharing information on social media. The echo-effect might be caused by “laggards” with particular characteristics, which make it possible to identify them based on their social media behavior. The same thing applies to the innovators, assuming they are influential and have many followers.

#### **4.3. Self-correcting mechanism**

There appears to be a self-correcting mechanism. Twitter users try to dispel rumors after they appear online. We found two rumors, although they accounted for a

modest amount of the Twitter stream. They both originated from the misinterpretation of a sarcastic or ironic tweet. As soon as the rumor started, other Twitter users tried to fact-check the rumors. The VRT and MH17 rumors were not corrected by official or governmental bodies but by individual Twitter users. The VRT-rumor was relatively easy to verify by means of retracing the original picture, the logo and the fact that the broadcast was still going on at VRT. The MH17 rumor looked quite persistent, although it made less impact than the VRT-rumor. All in all, it took more than two hours before the rumor finally started to fade out. Most likely, the rumor itself was more difficult to verify since it was mostly based on hearsay from people who were familiar with the gunman.

The self-correcting mechanism we identified is reflected in the fact that meaningful information is distinguished from rumors within social media communities. It might even be part of the game, and definitely a topic that warrants further study. The mechanism works as a safety valve. In the first place senders of incorrect or misinterpreted tweets do their best to correct the stream themselves. In the second place other Twitter users actively try to make sense of what is happening; facts are constantly validated.

In the third place, new information that influences discussions finds its way into social media, helping and at the same time challenging the self-correcting mechanism. There appears to be a certain “wisdom of crowds”, even though it should be pointed out that this wisdom as well as the crowds are fluid and highly dynamic. Further research is needed to deepen our understanding of the self-correcting mechanism of social media and whether or not the self-correcting mechanism is stronger when fact-checking becomes relatively easy, as was the case in the VRT-rumor compared to the rumor connected to flight MH17.

#### **4.4. Emotional response**

We believe to have identified online changes in mood, consisting of several alternating stages. The interpretation of changes in mood should of course be handled with care, since tweets with jokes and emotionally expressive language accounted for only a fairly small number compared to the retweets from breaking news updates. However, we have found an indication of an early stage with jokes during the first few minutes of the disruption of the news broadcast. It was soon followed by expressions of anxiety when it became clear that a gunman had entered the studio. Joking was no longer appreciated. As soon as the gunman was arrested, explicit emotional expressions on Twitter declined significantly. Together with the collective relief, the room for humor returned, finalizing in a more rational and collective factfinding stage, which started as soon as pictures of the gunman and his threatening letter came online. This lasted until the end of our data.

A common refrain in academic studies of crisis management holds that when a society is hit by a crisis, members of that society will look to their leaders to hear how they interpret the situation and what they plan to do to restore a state of normality (Boin, Kuipers, & Overdijk, 2013). In the online discussions, we did not find any call upon this kind of “meaning making” by public leaders or the director of the TV station. Apparently, online users felt no need for public interpretation as soon as the arrest was made. The online community started to interpret the situation on its own.

During his press conference at 10.30 pm, the mayor of Hilversum referred to the impact of the events at Charlie Hebdo. In the tweets we examined, the Charlie Hebdo connection did not seem to concern the Twitter users at all.

Fact-checking information was not possible while there was no other information than “a gunman” available. However, shortly after his arrest, users of Twitter started to share information about the gunman, his letter and his background. Instead of emotional responses, the Twitter community focused on collective factchecking the shared information for the rest of the night.

#### **4.5. Implications**

It would be interesting to replicate and further examine our main findings in the context of other crises. The current study provides a practical chance for crisis managers to gain more control over social media traffic. Social media have been confirmed to be useful as an efficient disaster backchannel communication mechanism when it comes to “governmental and public collaboration in times of disaster” (Tseng, Chen, & Chi, 2011). In recent crises we could see examples of this collaboration in a specific way, as Twitter is not only used for the spread of government information, but for operational purposes as well. The shooting in Ottawa (22 October 2014), the Paris attacks (13 November 2015) and the Brussels lockdown (22 November 2015) show that police tweets were retweeted extensively. The tweet from the Ottawa police department, in which the public was asked to refrain from tweeting pictures and information concerning the ongoing incident, was retweeted over 2000 times.<sup>8</sup> A similar tweet from Police Nationale after the Paris attacks was retweeted over 26,000 times.<sup>9</sup> A similar request by Brussels police became trending topic under hashtag #Brusselslockdown. It suggests that emergency services seek to utilize the online social exchange to encourage the public not to hinder police actions with live streaming. We know little about the potential of the social media cooperation between government and society, nor about strategies for its optimization. The same applies to the practical possibilities linked to the echo-effect and the exploitation of self-correcting mechanisms as reported in the current study. We can learn from the tweet by @helgavanleur, who initially retweeted the VRT-rumor. As she realized the story about VRT was a hoax, she decided to delete her tweet after five minutes.

As a result, all retweets by her followers and her followers' followers were removed as well. The practical implication is valuable from a risk and crisis communications perspective: news agencies that have an update of their news can prevent the retweeting of a tweet with obsolete news by simply deleting the source tweet. Future research can add new insights into these and other aspects of social media use in crises.

#### **4.6. Limitations**

The findings in this single case should be validated by other studies on other types of crises since this study is not without limitations. The collection and analysis of 58,931 tweets is challenging and subject to interpretation. This counts especially for tweets with an ironic or sarcastic tone, which are multiinterpretable.

Moreover, we are aware that we have not collected all the tweets in the given period, although we are unaware of the number of tweets we missed or which were posted and deleted by users afterwards. As we focused on the initial hours of the crisis, we cannot draw conclusions on the aftermath after 2.00 am.

## 5. CONCLUSION

Little is known in scientific literature about how social media evolve in the lifecycle of a crisis. The present paper reports on facts and rumors shared on Twitter during and after an incident at the Dutch national news studio. The gunman crisis we examined is a unique case because, despite a publicly communicated vague risk, it was unclear what was actually happening in the studio complex.

The study sheds light on the way Twitter users behaved as a watchman for the crisis, tracking and commenting on developments, the background of the gunman, and undertook attempts to fact-check online rumors. Humor can cause misunderstandings when it is interpreted incorrectly. In the early hours of the gunman-crisis we identified an echo-effect and a selfcorrecting mechanism that warrant further consideration. Also, in our view the study points to interesting options to enforce the cooperation between authorities and public services with social media users, which might enhance the utilization options of social media for crisis management and communication.

## Acknowledgments

We would like to thank Harro Ranter for making available his dataset of tweets. We thank Tim Reefman from HowAboutYou for collecting specific tweets on the VRT-rumor and Rens Dietz from Coosto for his general feedback. And we would also like to thank Twitter users @koendelaey, @jiscacohen, @telstaronline, @helgavanleur, @stevendalebout, @ibhaarsma, @jan\_homble and @essie1309 for their feedback.

## REFERENCES

- Alexander, D. E. (2014). Social media in disaster risk reduction and crisis management. *Science and Engineering Ethics*, 20(3), 717-733.
- Back, M. D., Küfner, A. C. P., & Egloff, B. (2010). The emotional timeline of September 11, 2001. *Psychological Science*, 21(10), 1417-1419.
- Binder, A. R. (2012). Figuring out #fukushima: an initial look at functions and content of US twitter commentary about nuclear risk. *Environmental Communication*, 6(2), 268-277.
- Boin, A., Kuipers, S., & Overdijk, W. (2013). Leadership in times of crisis: a framework for assessment. *International Review of Public Administration*, 18(1), 79-91.
- Coombs, W. T. (2004). Impact of past crises on current crisis communication: insights from situational crisis communication theory. *Journal of Business Communication*, 41(3), 265-289.
- COT Instituut voor Veiligheid en Crisismanagement. (2015). Het nieuws gegijzeld. Ervaringen en lessen naar aanleiding van het veiligheidsincident bij NOS en NPO op 29 januari 2015.
- Graham, M. W., Avery, E. J., & Park, S. (2015). The role of social media in local government crisis communications. *Public Relations Review*, 41(3), 386-394.
- Groenendaal, J., De Bas, M., & Helsloot, I. (2013). Twitter: an underutilized potential during sudden crises? *Journal of Contingencies and Crisis Management*, 21(3), 178-183.
- Ji, Q., & Raney, A. A. (2015). Morally judging entertainment: a case study of live tweeting during Downton Abbey. *Media Psychology*, 18(2), 221-242.
- Jong, W., Dückers, M. L. A., & van der Velden, P. G. (2016). Leadership of mayors and governors during crises: a systematic review on tasks and effectiveness. *Journal of Contingencies and Crisis Management*, 24(1), 46-58.
- Ki, E. J., & Nekmat, E. (2014). Situational crisis communication and interactivity: usage and effectiveness of Facebook for crisis management by fortune 500 companies. *Computers in Human Behavior*, 35, 140-147.
- Lachlan, K. A., Spence, P. R., Lin, X., Najarian, K., & Del Greco, M. (2016). Social media and crisis management: CERC, search strategies, and Twitter content. *Computers in Human Behavior*, 54, 647-652.

- Mendoza, M., Poblete, B., & Castillo, C. (2010). Twitter under crisis: can we trust what we RT?. In *Proceedings of the first workshop on social Media analytics* (pp. 71-79). Washington D.C., District of Columbia: ACM.
- Miller, D. E. (2005). Rumor: an examination of some stereotypes. *Symbolic Interaction*, 28(4), 505-519.
- Pang, N., & Ng, J. (2016). Twittering the Little India Riot: audience responses, information behavior and the use of emotive cues. *Computers in Human Behavior*, 54, 607-619.
- Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Simon and Schuster.
- Shibutani, T. (1966). *Improvised News*: Bobbs-Merrill.
- Snoeiijers, E. M., Poels, K., & Nicolay, C. (2014). #universitycrisis: the impact of social media type, source, and information on student responses toward a university crisis. *Social Science Computer Review*, 32(5), 647-661.
- Takahashi, B., Tandoc, E. C., Jr., & Carmichael, C. (2015). Communicating on Twitter during a disaster: an analysis of tweets during typhoon Haiyan in the Philippines. *Computers in Human Behavior*, 50, 392-398.
- Tseng, S.-F., Chen, W.-C., & Chi, C.-L. (2011). Online social media in a disaster event: network and public participation. In H. Cherifi, J. Zain, & E. El-Qawasmeh (Eds.), *Digital information and communication technology and its applications* (Vol. 167, pp. 256-264). Heidelberg: Springer Berlin.
- Utz, S., Schultz, F., & Glocka, S. (2013). Crisis communication online: how medium, crisis type and emotions affected public reactions in the Fukushima Daiichi nuclear disaster. *Public Relations Review*, 39(1), 40-46.
- Wang, X. (2013). Applying the integrative model of behavioral prediction and attitude functions in the context of social media use while viewing mediated sports. *Computers in Human Behavior*, 29(4), 1538-1545.
- Westerman, D., Spence, P. R., & Van Der Heide, B. (2014). Social media as information source: recency of updates and credibility of information. *Journal of Computer-Mediated Communication*, 19(2), 171-183.
- Yu, Y., & Wang, X. (2015). World cup 2014 in the Twitter world: a big data analysis of sentiments in U.S. sports fans' tweets. *Computers in Human Behavior*, 48, 392-400.

FIGURES

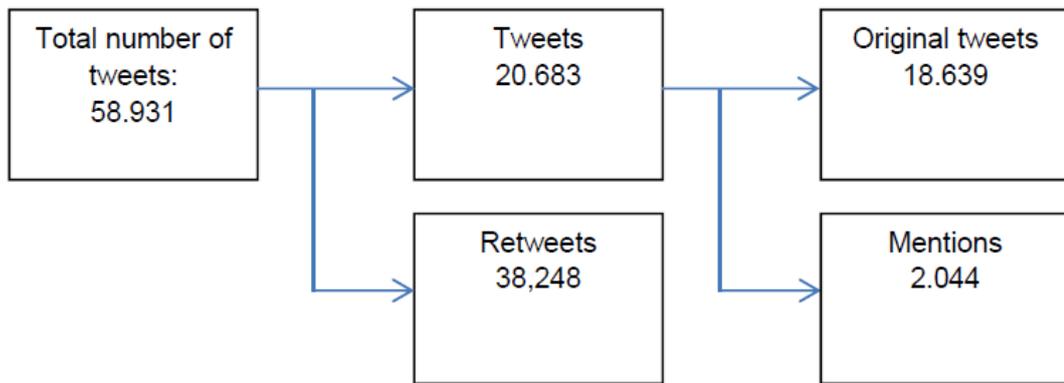


Fig. 1. Total number of tweets (divided into retweets, original tweets and mentions).

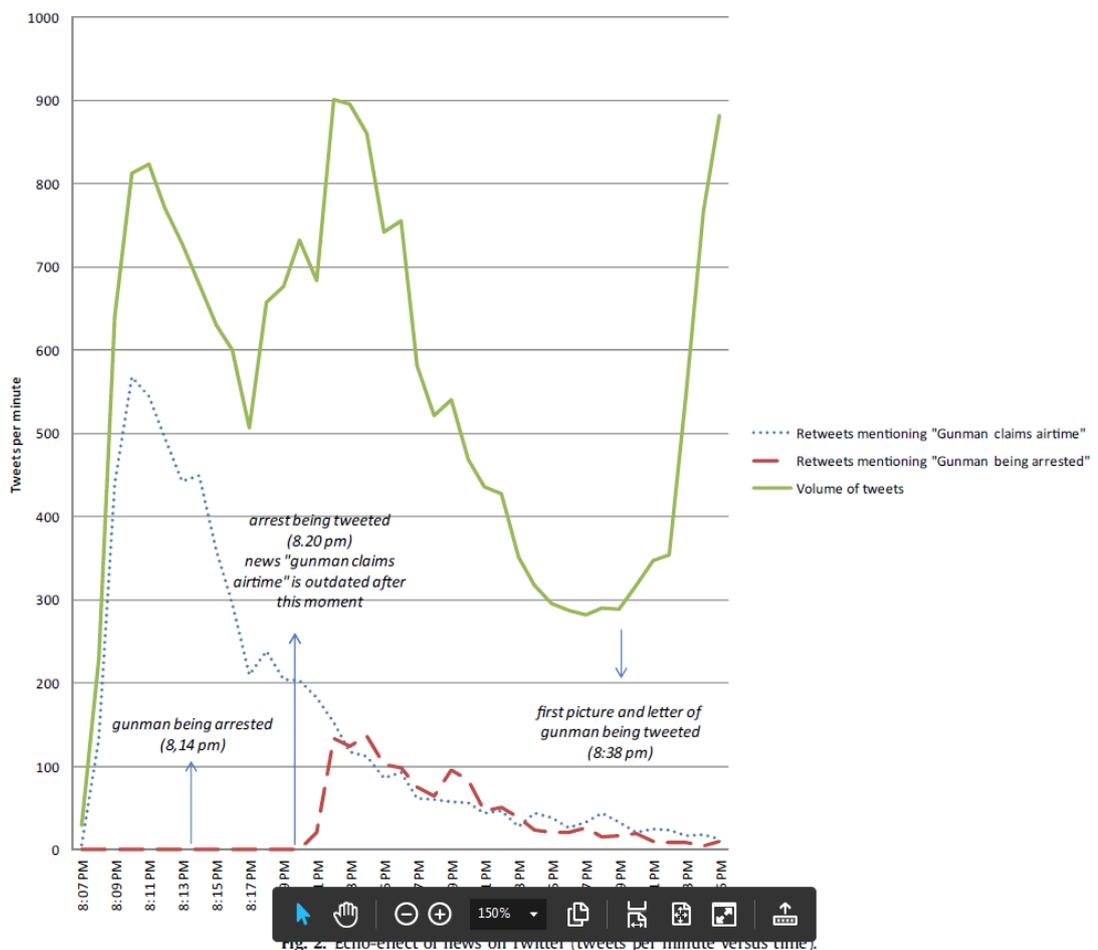


Fig. 2. Echo-effect of news on Twitter (tweets per minute versus time).

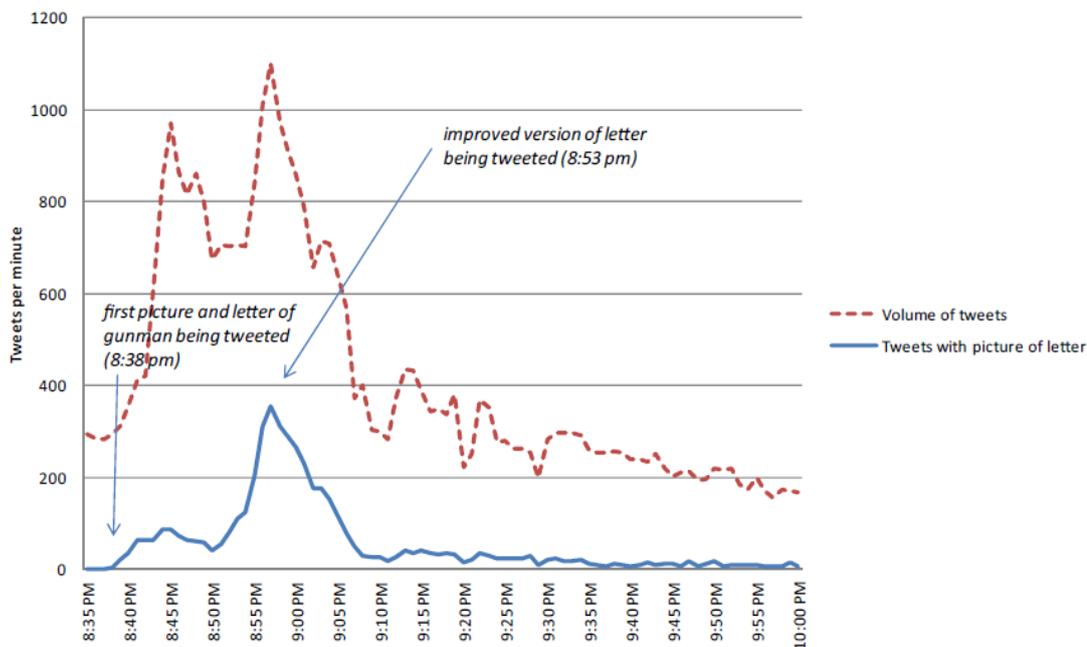


Fig. 3. Tweets concerning threatening letter (tweets per minute versus time).

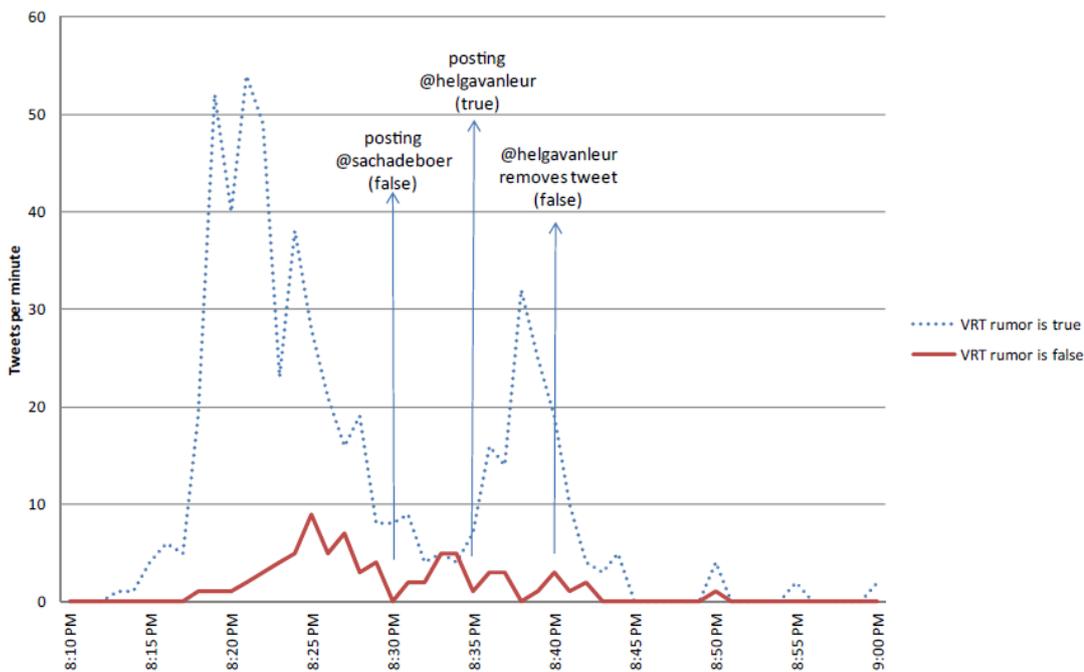


Fig. 4. Tweets mentioning the VRT-rumor 'true' and 'false' (tweets per minute versus time).

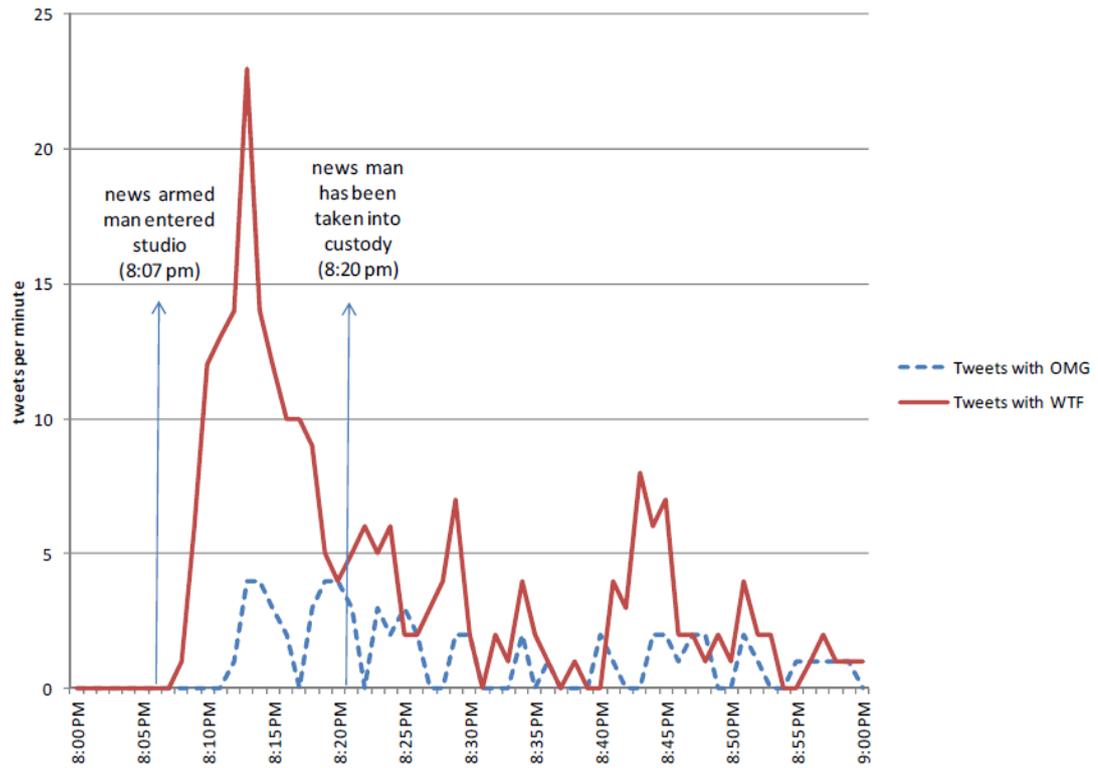


Fig. 5. Explicit emotional expressions in tweets (tweets per minute versus time).