

COLLECTIVE RUMOR CORRECTION ON THE DEATH HOAX OF A POLITICAL FIGURE IN SOCIAL MEDIA

Alton Y.K. Chua, Wee Kim School of Communication & Information, Nanyang Technological University, Singapore, AltonChua@ntu.edu.sg

Sin-Mei Cheah, Wee Kim School of Communication & Information, Nanyang Technological University, Singapore, sin_mei@yahoo.com

Dion H Goh, Wee Kim School of Communication & Information, Nanyang Technological University, Singapore, ASHLGoh@ntu.edu.sg

Ee-Peng Lim, Singapore Management University, Singapore, eplim@smu.edu.sg

Abstract

Conversations on social media networks that discuss a crisis incident as it unfolds have become a norm in recent years. Left to its own devices, such conversations could quickly degenerate into rumor mills. Little research has thus far examined the correction of rumors on social media. Using the third-person effect as a theoretical underpinning, we developed a model of collective rumor correction on social media based on an incident surrounding the death hoax of a political figure. Tweets from Twitter were collected and analyzed for the period when a spike of circulating rumors speculating the demise of Singapore's first prime minister was detected. Corrections of the rumor also went viral on the same day. Our study reveals that corrective behavior during a death hoax situation on Twitter is characterized by affirmative and rational rebuttals verifiable by credible sources. While the inclusion of credible sources is essential for both rumor diffusion and corrections, correcting a rumor differs from its diffusion in that unambiguity and low emotional levels are crucial. Key characteristics of collective rumor correction identified by this study have implications for both theory and practice. We discussed these implications together with the study's limitations and suggestions for future research.

Keywords: Rumor correction, Death hoax, Political figure, Twitter, Third-person effect.

1 INTRODUCTION

While social media remains a popular channel for the online community to convene every time a major crisis erupts, online chatter is at best viewed as citizen journalism and at worst “collective rumor mills” (Oh, et al., 2013, p. 408). When timely official information is absent at times of crises, citizen journalism can become a primary source of real-time news. However, the uncertainty and anxiety surrounding crises situations often trigger rumors, or what Shibutani (1966) calls “improvised news” based on hearsay, which go viral online. The ubiquity of mobile devices coupled with the ease of use, speed and global reach of social networking sites have helped the proliferation of rumors.

Online rumors if left uncorrected may lead to dire consequences. Those of a political nature could lead to hostile attitudes and aggressive behavior (Garrett 2011). Rumor correction is therefore critical so as to minimize the damaging effects from the propagation of false information. Mechanisms to control rumors have long been in place, with governments and traditional news media serving as the primary gatekeepers (Donovan 2007). Recent studies suggest that the online community is capable of self-correction and self-policing when presented with doubtful information (Shklovski et al. 2008; Starbird et al. 2014).

The research on rumor correction remains nascent as most rumor studies continue to emphasize largely on the transmission and retransmission stages, while ignoring the later stages in the rumor lifecycle, in particular rumor correction (Liu et al. 2014). The few studies on rumor correction tend to adopt a network modelling approach (e.g. Bernard et al. 2014; Takayasu et al. 2015) or to investigate corrections as a part of the work on rumor diffusion (e.g. Maddock et al. 2015; Nadamoto et al. 2013; Starbird et al. 2014). The lack of attention on rumor correction in social media presents a knowledge gap that motivates this study. The objective of this study is, therefore, to examine the characteristics of rumor corrections occurring in Twitter, set in the context of the death hoax of a famous politician, Lee Kuan Yew, who was Singapore’s first prime minister. In addition to drawing from areas of socio-psychology and communication, this study takes an Information Systems perspective by framing the antecedents of the phenomenon in terms of information ambiguity and problems related to information diffusion.

The rest of the paper is organized as follows. Literature on rumor correction and the theoretical underpinning of third-person effect is reviewed. A research model is then developed pertaining to rumor corrections surrounding the death hoax of a politician. Data collection and analysis methods are described in the methodology section, and the results of the regression analysis are presented in the findings section. Thereafter the findings are discussed, followed by the paper’s contributions and limitations.

2 LITERATURE REVIEW

2.1 Rumor Corrections

2.1.1 *Conceptualization of Rumor Correction*

Rumor corrections cannot occur in the absence of rumors. We first seek to understand multiple definitions of rumor in different streams of literature before conceptualizing rumor correction. In the socio-psychology domain, rumor is defined as “unverified and instrumentally relevant information statements in circulation that arise in contexts of ambiguity, danger or potential threat, and that function to help people make sense and manage risk” (DiFonzo & Bordia 2007, p. 13). Sociologists described rumors as improvised news (Shibutani 1966) and rumor-mongering as a collective problem solving process for people to cope with perceived threats during a crisis. Synthesizing these definitions, recent work from an IS perspective defined rumoring as “a collective and collaborative transaction in which community members offer, evaluate and interpret information to reach a common

understanding of uncertain situations, to alleviate social tension and to solve collective crisis problems” (Oh et al. 2013, p. 409).

Rumor corrections are often conceptualized as control strategies [e.g. [Bernard et al. \(2014\)](#); [Kimmel \(2004\)](#); [Tripathy et al. 2010](#)]. Neutralizing rumors is one way to rein them in ([DiFonzo et al. 1994](#)). Implemented at strategic points in the rumor process – during generation, evaluation or dissemination – neutralization can take multiple forms, such as ignoring, confirmation of the truth, and denial. Ignoring a rumor is deemed the weakest tactic of all and is used only if the rumor is highly implausible. As a rumor seldom dies on its own, an explicit control tactic is still preferred. Rumors often carry some truth and by confirming that part of the rumor that is true may suffice to neutralize its impact. Denial is a popular mechanism used to refute rumors ([Rosnow 1974](#)) but its effectiveness has been questioned ([DiFonzo et al. 1994](#)). Other rumor coping tactics include providing the information that is in demand and enhancing trust and credibility by engaging in public relations ([Kimmel 2004](#); [Kimmel & Audrain-Pontevia 2010](#)).

Rumors could serve the function of collective sense-making to a community attempting to understand ambiguous or uncertain situations when official information is lacking ([DiFonzo & Bordia 2007](#)). Moreover, the contents of rumors usually contain instrumentally relevant information to its participants ([DiFonzo & Bordia 2007](#)). The contents of rumor corrections are similar to rumors, except for the key distinction, that is, verification. While rumors are unverified statements with no proof of evidence, corrections are based on established evidence supported by testimonies from credible sources ([DiFonzo & Bordia 2007](#)). In this paper, we define rumor correction as an affirmative statement declaring a rumor as false by providing plausible explanations, facts or evidence that is instrumentally relevant and verifiable with a credible source.

2.1.2 Correction: A Stage in Rumor Lifecycle

A rumor and its corrections are intricately related in the sense that the genesis of corrections occur within the life course of a rumor. Scholars have attempted to identify rumor life stages, although a clear distinction between the different stages may be difficult in some cases ([Kimmel 2004](#)). In [Rosnow’s \(1974\)](#) proposition of a three-stage rumor process model, rumor control is theorized as the final stage of the rumor lifecycle. After a rumor goes through parturition and diffusion, eventually it will either die a natural death as interest wanes, or perish as a result of deliberate effort in stopping its propagation. A common device to control rumor mongering is denial by official authorities that the rumor is untrue.

[Bordia and DiFonzo’s \(2004\)](#) four-stage rumor interaction process suggested that corrective actions occur at the third stage. After the first two stages of sharing and sense-making of the shared information, the group discussion is then centered on deciding the validity of the rumor. Collective corrections take the form of directive statements suggesting a course of action for the group, and typically include putting a stop to the discussion topic.

2.1.3 Correction as a Collective Behaviour

Rumor corrections made by an online community can be viewed as a type of collective action ([Oh et al. 2015](#)). Collective behaviour, defined as “an improvised, emergent, self-organized and adaptive group behaviour” ([Oh et al. 2015, p. 211](#)), occurs in the context of crisis when the usual communication channels and established social norms do not function as expected. In place of traditional news media, social media takes over to empower citizens in collective participation, particularly in processes that affect them ([Leong et al. 2015](#)).

Although social media can act as conduits of misinformation and rumor in times of crisis, studies indicate that these platforms can effectively be used to stop the spread of rumors as well ([Bordia & DiFonzo 2004](#)). The argument that collective action in social media platforms can help check the spread of rumors has gained traction in such cases.

Collective rumor corrections are characterised by advice sharing, which include coaxing the members to check facts before posting and to refrain from spreading rumors (Bruns et al. 2012). Grassroots websites and locally-operated news producers contribute towards collective correction of rumors aimed at reducing information and situational ambiguity. Based on mathematical modelling, studies have evaluated the metrics that are involved in the rebuttal of rumors online (e.g., Tripathy et al. 2013). The premise is that most people are inclined to refute rumors when they can see through the rumors' falsehood. However, studies that attempted to explain rumor correction from a socio-psychological perspective are rare. Thus, we utilized the theory of third-person effect, which helped explore the motivations behind collective action against the perceived scourge of rumor.

3 THEORETICAL BACKGROUND: THIRD-PERSON EFFECT AND CORRECTIVE BEHAVIOUR

Third-person effect describes an individual's belief that other people (i.e. the third person), not oneself, are more susceptible to the negative persuasion of the media. The individual is consequently motivated to react out of concern for others (Davison 1983; Gunther 1995). In this hypothesis, the behavioural effect is based on the individual's perception of public opinion being negatively affected by the media (Gunther 1995; Perloff 1993). Behavioural reactions to the third-person perception have been classified as preventive, accommodative and corrective (Gunther et al. 2006; Rojas 2010). Preventive actions serve to stop the exposure of a selected group of people to media content perceived having undesirable effects (Gunther et al. 2006). Accommodative actions, on the other hand, are observed when people accept presumably desirable impact on others (Gunther et al. 2006).

Corrective action represents people's effort to make their own views heard, therefore counteracting and minimizing the perceived media effects on others (Rojas 2010). Corrective or rectifying behaviour is reactive as it is triggered by the perception that others are unable to resist the harmful influence (Sun et al. 2008b). The primary driver of reaction is directed towards swaying public opinion away from the perceived negative effects of media. Studying the effects of media on political behaviors, Rojas (2010) found support for the third-person effect, in the form of posting to discussion forums, commenting on news articles and persuading others.

Studies have also demonstrated that the perceived desirability of a message play an important role in influencing behavior. Messages with ambiguous influence, in particular, are more likely to elicit corrective behaviours (Sun et al. 2008a, 2008b). In a similar vein, Hwang et al. (2008) asserted that an individual's perception of the media being biased can elicit negative emotions that in turn result in one's tendency to engage in discursive activities, such as voicing one's opinions publicly.

The corrective behavioral component of the third-person effect is applicable to the context of rumor corrections because it provides an explanation for the motivation of individuals to initiate actions in rectifying untrue information in public sphere. Drawing on the third-person effect on individuals' behaviour, we extrapolate the effect to an online community's collective behaviour and propose a model of rumor correction regarding the death hoax of a political leader on Twitter.

4 RESEARCH MODEL ON RUMOR CORRECTION

4.1 Background on Previous Models

Our proposed model is developed by drawing on previous work on rumor transmission and retransmission as there is a scarcity of research models on rumor correction. Much of the work on rumor research is grounded on Allport and Postman's (1947) seminal model of rumor transmission. Their model postulated the importance of the subject topic to an individual and the ambiguity of topical information as predictors of rumor transmission. The concept of anxiety was subsequently coined by Anthony (1973) to represent both topical importance as the affective factor and ambiguity as the cognitive factor (Chorus 1953) that influences the diffusion of rumor. As the corpus of rumor

literature grew, a meta-analysis revealed that anxiety, uncertainty (or ambiguity) and credulity (or critical sensibility) are key variables in rumor-mongering situations ([Rosnow 1991](#)).

Oh et al.'s (2013) model of rumor transmission in social media found significant effects of source ambiguity, content ambiguity and anxiety on rumor mongering, whereas social ties, which refers to the interpersonal relationships between community members, did not influence the spreading of rumor.

Following Oh et al.'s model (2013), a rumor retransmission model was proposed by Liu et al. (2014), with cues adopted from the Elaboration Likelihood Model. Their model included the following variables: anxiety, personal involvement, content ambiguity, sender's credibility and attractiveness. Sender's credibility refers to the perceived trustworthiness, competency and reputation of an information source to publicize a message ([Pornpitakpan 2004](#)). Attractiveness of a message is assessed based on the use of visual aids, such as videos or pictures, to make the message more appealing. The results indicate that content ambiguity, sender's credibility and attractiveness of a message lead to rumor retransmission.

Empirical evidence of previous models on rumor-mongering provided the underpinning on which we selected variables to include in our model. We chose variables that produced varied results: variables which had positive results (i.e. source ambiguity, sender credibility and content attractiveness) and variables which had mixed results (i.e. anxiety and content ambiguity). In the subsections that follow, we introduce the variables used in our model and explicate the possible effect they have on rumor correction.

4.2 Emotions

Emotions, including anxiety, intensify as an unexpected crisis first occurs. In the context of the rumored death of a political figure, the lack of timely information from official sources or mainstream media leads the public to social media platforms ([Oh et al. 2013](#)). These online conversations include emotional statements that carry both positive and negative feelings ([Oh et al. 2010](#)). Therefore in our model, we measure the effect of emotions in general, and not anxiety alone. Emotional expressions, including feelings of anxiety, sorrow or compassion, are dominant at the beginning of a rumor lifecycle but diminish rapidly as time progresses ([Oh et al. 2010](#)). Given that rumor corrections generally occur at the final stages of its lifecycle, we predict that these corrections are less likely to be emotional. Presence of emotions in Twitter messages also implies an absence of critical evaluation on the part of the sender. Conversely, a disposition for objective assessment of message precludes an emotional engagement with it. Hence, we propose the first hypothesis as follows:

H1: The expression of emotions in a message that corrects rumors about the demise of a political figure is negatively associated with rumor correction.

4.3 Source Credibility

Source credibility refers to the extent of perceiving an information source as trustworthy, competent ([Wilson 1983](#)), believable ([Wathen & Burkell 2002](#)), reliable, objective and accurate ([Hilligoss & Rieh 2008](#)). A key concept related to source credibility is cognitive authority, that is, a source that exerts influence on our thoughts, and is considered to be "thought credible, worthy of belief" ([Wilson 1983](#)). Cognitive authorities can be both institutional (such as academic and government institutions) and individuals who demonstrate trustworthiness, competence and expertise ([Pornpitakpan 2004](#); [Rieh & Danielson 2007](#); [Wilson 1983](#)). Messages sent by cognitive authorities are deemed more credible, convincing and persuasive ([Pornpitakpan 2004](#)). A credible and authoritative source, such as the mainstream media or an established organization, lends support to rumor correcting messages ([Oh et al. 2010](#)). Under the third-person effect, credible information could also influence judgments. Hence, we propose the second hypothesis as follows:

H2: The source credibility of a message that corrects rumors about the demise of a political figure is positively associated with rumor correction.

4.4 Content Ambiguity

Content ambiguity refers to the lack of interpretative clarity of a message's meaning (Oh et al. 2013). As long as the information content remains ambiguous, members within a community will engage in active information seeking, sharing and discussion in an attempt to reduce cognitive ambiguity (Oh et al. 2013; Rosnow 1991). One effective way to combat rumors is to confirm the truth (DiFonzo et al. 1994) and disseminate unambiguous information (Oh et al. 2010). As a rumor progresses, messages carrying unambiguous information increase over time, until the rumor is suppressed by a sufficient amount of certainty (Oh et al. 2010). Reducing uncertainty is necessary, as suggested by the third-person effect. We contend that as a rumor approaches its final stages, more affirmative statements that refute the rumor are anticipated, and thus we present the next hypothesis as follows:

H3: The content ambiguity in a message that corrects rumors about the demise of a political figure is negatively associated with rumor correction.

4.5 Visual Content

Visual imagery adds to the effectiveness of a message's argument by illustrating a point or reinforcing the content, thereby amplifying its persuasive effects (Seo et al. 2013). Exemplification theory explains how images can elicit affective reactions leading to persuasive outcomes (Zillmann 2002). Rumor messages with images or videos were found to be more persuasive due to the added meanings and emotions conveyed as compared to a plain text message (Liu et al. 2014). Multimedia resources, such as photos and videos, constitute a substantial amount of information tweeted by the online community, especially during situations of social crisis (Bruns et al. 2012). In a similar vein, correction messages that are accompanied by visual content that pertains to the death hoax of a political figure are expected to reinforce the message's meaning. Hence, we propose the final hypothesis as follows:

H4: Visual content in a message that corrects the rumors about the demise of a political figure is positively associated with rumor correction.

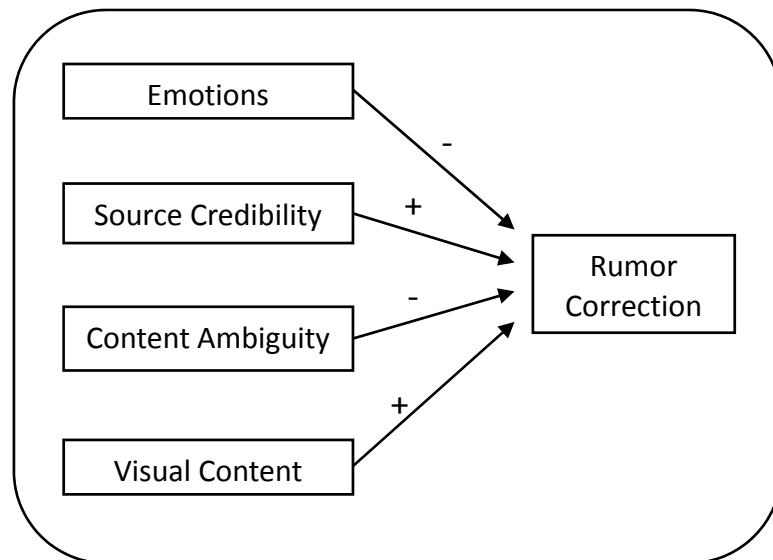


Figure 1. The proposed model on rumor correction

5 METHODOLOGY

5.1 Background: Premature Reporting of Lee Kuan Yew's Death

The death of a key political leader is a significant event in a country's history, leaving a pervasive effect on the society and the economy. A death hoax, that is, the misreporting of the passing of a person, while having less damaging effects, can mislead people into believing the death is real until the rumor is corrected. Leaders like former British Prime Minister Margaret Thatcher, U.S. President Barack Obama and North Korea's leader Kim Jong-un have been targets of death hoaxes in recent history.

The recent case of the rumored death of Singapore's first Prime Minister, Lee Kuan Yew, provides an opportunity to test the study's hypotheses empirically. Rumors about his passing made its rounds on social media after news of his admission in a local hospital's intensive care unit for severe pneumonia was released. Subsequently, the worsening of his condition again brought him into the limelight. On 18 March 2015, a doctored screen capture of an official announcement of his death, purportedly issued from the Prime Minister Office, went viral on social media. The fake statement allegedly declared that Lee, aged 91, had passed away at the Singapore General Hospital on 5.30pm that day. Closely resembling the official press releases, the forged image consequently misled netizens and foreign news media into prematurely reporting his demise.

5.2 Data Collection and Selection

The data for this study is drawn from the Twittersphere. Twitter, the popular microblogging website, has increasingly gained prominence as an influential news channel (Murthy 2011) and as a platform for scholarly inquiry into communicative behaviour (Kwak et al. 2010), including rumor research. As news of Lee Kuan Yew's worsening health condition was publicised, many netizens took to the Twittersphere to express their concerns and well-wishes for him. The outpouring of support for the former statesman reached a distinctive peak on 18 March (Lin et al. 2015). On the same day, a fake announcement of his death started to spread on Twitter at around 20:00 hours. It was not long before a local news channel (Channel NewsAsia) promptly announced that it had verified the circulating image was fake and debunked the rumor. This was followed shortly by correction tweets sent out by the local newspaper (The Straits Times), which were retweeted widely within a few hours. The rumor messages eventually subsided around 23:00 that day. Therefore we selected 18 March as the day to collect tweet messages for the study.

For data collection, a web scraper was used to gather tweets tagged with the hashtags #LeeKuanYew and #LKY on 18 March 2015. This yielded over 5,800 tweet messages from 3,995 unique tweet authors. To identify corrective messages, we referred to the definition of rumor correction in this paper, which is "an affirmative statement declaring a rumor as false by providing plausible explanations, facts or evidence that is instrumentally relevant and verifiable with a credible source". The corrective messages were coded manually using content analysis (Krippendorff 2012). An example of a tweet coded as a corrective message is:

(18 March 2015, 22:29) Kindly do not spread rumors about Mr #LeeKuanYew. The image that is spreading is edited from that of Mrs #LKY.

Upon analysis of the tweets, 30% were found to be correction messages. The non-corrections consisted of a broad range of conversation topics including well-wishes (e.g. "rest in peace") and gratitude (e.g. "thank you"). For the final dataset admitted for further coding and analysis, we randomly selected 500 corrections and an equal number of non-corrections. The sampling was based on a subjective assessment of tweets deemed to be the most representative. This approach not only enabled us to work with a manageable sample size of 1,000 tweets, and helped minimize the problem of redundancy in tweet content.

5.3 Further Coding and Analysis

Each tweet is a unit of analysis and is coded to measure the effect of the independent variables on rumor correction, the dependent variable in this study. Consistent with previous work on rumor prediction models (e.g. Liu et al. 2014; Oh et al. 2013), content analytic coding is used to determine the presence or absence of characteristics of a variable in our research model and all variables are operationalized as dichotomous. To code the independent variables, we adapted the definitions used by Liu et al. (2014) and Oh et al. (2013). The coding scheme for independent variables is described in Table 1.

For the coding process, a pilot dataset of 100 tweet samples which was randomly selected from the initial dataset was used. Two local researchers based in Singapore coded the sample set separately. As the first round of coding result did not reach the acceptable level of agreement at 0.70 (Lombard 2002) for emotions and source credibility, the researchers discussed the disagreed results and conducted a second round of coding. The final Cohen kappa's value for inter-coder reliability was 0.92 for rumor correction, 0.71 for emotions, 0.82 for source credibility, 0.81 for content ambiguity and 0.76 for visual content. After confirming the reliability of the pilot dataset as acceptable (above 0.70), each researcher coded an equal amount of the remaining tweets.


Independent Variable	Definition
Emotion	<p>A Twitter message that contains emotional feelings (e.g. anxiety, anger, sadness) towards the rumored death of a political figure.</p> <p>Example: “Don’t be so idiot! He’s not gone!!! Who is so evil to spread that stupid fake picture?!! #LeeKuanYew” is coded as 1 (i.e. presence of emotions).</p>
Source credibility	<p>A Twitter message that either originates from the local mainstream media or government authorities, or is forwarded from the mainstream media or government authorities’ message.</p> <p>Example: “*PMO lodging police report about fake website announcing death of Mr Lee Kuan Yew” is coded as 1 (Originated from the local mainstream media: The Straits Times). *PMO stands for Prime Minister Office.</p>
Content ambiguity	<p>A Twitter message that ask questions or expresses doubt towards the rumored death of a political figure.</p> <p>Example: “Is this true about #LeeKuanYew?” is coded as 1 (i.e. content is ambiguous).</p>
Visual Content	<p>A Twitter message containing a graphic that is related to the rumored death of a political figure.</p> <p>Example: A photo of Lee Kuan Yew or the doctored screen capture announcing his death.</p>  <p>The image shows a screenshot of a tweet from the account 'sg-cafe @media_sgcafe'. The tweet text reads: 'Kindly do not spread rumours about Mr #LeeKuanYew. The image that is spreading is edited from that of Mrs #JKY.' Below the text is a screenshot of a website, which is a doctored version of the Prime Minister's Office (PMO) website. The doctored image shows a red banner at the top with the PMO logo and the text 'PRIME MINISTER'S OFFICE' and 'PRIME MINISTER'S OFFICE: PASSING OF MR LEE KUAN YEW'. The main content area of the website is also visible, showing a statement from the PMO.</p>

Table 1. Coding Scheme

Given the dichotomous nature of the dependent variable (rumor correction), logistic regression is used to analyze the data (Hair et al. 2013). A logistic regression model allows us to determine the relationship between a binary outcome variable (in this case, rumor correction) and a set of predictor variables (that is, emotions, source credibility, content ambiguity and visual content). The result of the logit model, in the form of odds ratios, can be interpreted to represent the effect a certain predictor variable has on the likelihood of rumor correction being detected (i.e. having the value of 1).

6 FINDINGS

6.1 Descriptive Statistics

Of the four independent variables, source credibility appears to be the most prevalent in correction tweets (59%) while content ambiguity (96%) and emotions (63%) occur more frequently in non-correction tweets. Visual content related to the rumored demise (178 tweets, 17.8%) appears only occasionally in the sample of 1,000 tweets posted.

Correction messages are characterized by unambiguous content (reversed coded, 435/500 tweets, 87%), followed by a lack of emotions and credibility of sender or sources cited. Comparatively, the non-correction messages tend to be more ambiguous, more emotional, citing less of credible sources and are accompanied by less visual content. The descriptive statistics of the sample are shown in Table 2.

Independent Variable	No. of correction tweets (%)	No. of non-correction tweets (%)
Emotions	83 (17%)	315 (63%)
Source Credibility	293 (59%)	16 (3%)
Content Ambiguity	65 (13%)	479 (96%)
Visual Content	160 (32%)	18 (4%)
Total	500	500
Sample Size	1,000	

Table 2. General Descriptive Statistics

We observed the diffusion of the correction and non-correction tweets over time as illustrated in Figure 1. A sharp spike of correction messages was detected around 10pm after the local mainstream newspaper, The Straits Times (@STcom), released two tweets declaring the death news as fake. Their tweets were quickly retweeted and became the top two most frequently shared messages. Table 3 summarizes the top five Twitter messages (including tweets and retweets) that corrected the rumored death. As for non-corrections, they comprised of varied themes, with the most frequently posted messages containing references to “rest in peace” (69/500, 14%), “thank you” (58/500, 12%), “pray” (41/500, 8%), “SG50¹” (31/500, 6%) and “Singapore General Hospital” (30/500, 6%). The volume of tweets began to drop drastically as the day came to an end.

¹ SG50 is the abbreviation for Singapore’s 50th birthday

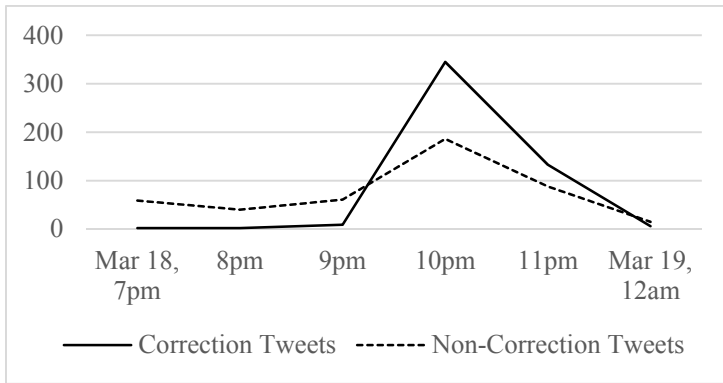


Figure 1. The diffusion of correction versus non-correction tweets

Correction Message	No. of tweets (%)
“Prime Minister’s Office is lodging a police report about fake website announcing death of Mr Lee Kuan Yew. #LeeKuanYew” (Originated from the media: The Straits Times)	196 (20%)
“PMO lodging police report about fake website announcing death of Mr Lee Kuan Yew #LeeKuanYew” (Originated from the media: The Straits Times)	78 (8%)
“Kindly do not spread rumors about Mr #LeeKuanYew. The image that is spreading is edited from that of Mrs #LKY.” (Originated from an author: media_sgcafe)	49 (5%)
“1. LKY is not dead yet. 2. Stop saying he is dead. 3. If you have nothing better to say about him, don't say. #LeeKuanYew” (Originated from an author: pjunkiat)	21 (2%)
“BREAKING & EXCLUSIVE: @cctvnews joins @cnnbrk in misreporting death of #LeeKuanYew. #LKY” (Originated from an author: firdianshah1)	10 (1%)
Total	500

Table 3. Top 5 correction messages

6.2 Inferential Statistics

The Spearman correlation test results indicate no multicollinearity problems as all correlations are below 0.7 (Dormann et al. 2013) (Table 4).

	Emotions	Source Credibility	Content Ambiguity	Visual Content
Emotions	1.000	-	-	-
Source Credibility	-0.522**	1.000	-	-
Content Ambiguity	-0.466**	0.648**	1.000	-
Visual Content	-0.240**	0.147**	0.430**	1.000

**Significant at the 0.01 level (2-tailed).

Table 4. Spearman correlation results

Using logistic regression to estimate the probability of rumor correction, we present the results in Table 5.

Variable	B	SE	Sig.	Exp(B)	Hypothesis
Emotions	0.627	0.261	0.016**	1.873	H1 supported
Source Credibility	-0.874	0.489	0.074*	0.417	H2 supported
Content Ambiguity	-6.800	0.785	0.000***	0.001	H3 supported
Visual Content	0.948	0.826	0.251	2.580	H4 rejected
Constant	4.502	0.796	0.000	90.181	-

***Significant at the 0.01 level. **Significant at the 0.05 level. *Significant at the 0.1 level.

Table 5. Regression results

The results indicate a good model fit, as $\chi^2 = 943.207$, $df=4$ ($p<0.000$). H1 is supported indicating that the probability of an unemotional Twitter message correcting a rumor is 1.873 times higher than an emotional one. H2 is supported at the significance level of $p<0.1$ (0.074), implying that a corrective message is more likely to be sent from a credible source or cites credible sources. The strong support for H3 ($p<0.001$) indicates significant negative effects of content ambiguity on the correction of rumor, that is, the probability of a corrective message being unambiguous is highly likely compared to an ambiguous one. Visual content does not have any effect on rumor correction as shown in the regression results, and therefore H4 is rejected. The strongest predictor is emotions, followed by source credibility and content ambiguity.

7 DISCUSSION

Overall the results provide empirical evidence to demonstrate that three key elements, namely source credibility, a lack of emotions and content unambiguity, support collective rumor correction on social media. Four main findings can be gleaned from the results.

First, the results suggest that source credibility is a critical component of a persuasive message when its purpose is to curtail rumors. Further, messages sent from credible and authoritative sources appear to shape netizens' perception of the truth and drives corrective behaviour based on the third-person effect. The rumor in this study started when a doctored screenshot reporting Lee's death, which bore close resemblance to a typical official announcement, was shared extensively in the Twittersphere. This was also the case when subsequent corrections issued from a local mainstream newspaper clarified his death report was fake. Both were believed to have originated from authoritative sources. The Elaboration Likelihood Model suggests that authority is a peripheral cue that convinces people to believe a message (Cialdini 1993). Persuasiveness of both the rumor and the corrective messages in the study was enhanced with their attribution to authoritative sources. For instance, the top two most frequently circulated correction tweets in the study made references to the Prime Minister Office (or PMO in short). The citation of authoritative sources is therefore a clear indication of veracity in the correction tweets (Oh et al. 2010).

Second, content ambiguity and emotions are both negatively associated with collective rumor correction. Contrary to the reason why rumors flourish due to heightened anxiety or emotions (Allport & Postman 1947), the results of our model suggest that the sharing of corrections by the online community is not driven by high emotional tension. Rather, as explained by the third-person effect, corrective behavior emerges out of a process of sensing and perception. As netizens sense a potential bias in the circulating rumor and perceive possible undesirable consequences, they take to social media and act by either urging others to refrain from spreading rumors or by sharing the correct information disseminated from official sources. Unlike the frequent use of emotional words depicting grief (e.g. "rest in peace"), gratitude (e.g., "thank you") and hope (e.g., "pray") in non-correction tweets in this study, correction messages contains factual and unambiguous information (e.g. "LKY is

not dead yet”) and statements suggesting a course of action (e.g., “Kindly do not spread rumors about Mr #LeeKuanYew” and “Stop saying he is dead”). Therefore, communication that reduces information ambiguity and conveyed in a less emotional tone seems to be essential in correcting rumors.

Third, visual imagery does not appear to be a feature of rumor corrections on social media. Consistent with studies that analyzed image sharing, e.g. [Bruns et al. \(2012\)](#), a substantial amount of tweets were accompanied by images in our data. One in every four tweets shared images of either the doctored screen capture, stock pictures of Lee Kuan Yew, memes or other photographs. Although much visual content is being shared online, our results did not indicate that visual imagery could be an attribute of rumor corrections, which is in contrast to [Liu et al.’s \(2014\)](#) findings for rumor retransmission in a disaster situation. Our findings thus demonstrate the limitation of the third-person effect through the use of visual imagery on the discerning online community in this context. One possible explanation is the limited variability of visual content in a death hoax incident, as compared to the diverse variety of images such as first-hand photographs that can be taken of the aftermath at disaster sites.

Finally, it was observed that collective rumor corrections seem to follow a rise-and-fall diffusion pattern over time and shows low content diversity. Using signatures or information patterns, we observed that rumor corrections and non-corrections both surge and plunge in tandem, but differ in magnitude, with corrections reaching a higher peak than non-corrections (see Figure 1). [Starbird et al. \(2014\)](#) found similar behavioral patterns in their study on rumors and corrections. Furthermore non-corrections, particularly rumor messages, eventually subsided after corrections were shared widely. This suggests that repeated exposure to rumor corrections may reduce rumor-mongering, which is consistent with [Garrett’s \(2011\)](#) findings in a study on online political rumoring. Next, when it comes to content diversity, we found a considerable amount of retweeted corrections that echoed the tweets of a local mainstream newspaper without modification, thereby indicating an amplification effect of repetitive content. As observed by [Allport and Lepkin \(1945\)](#) in their classic study on diffusion of wartime rumors, repetition of the same statement from different sources makes it more believable. Again, this suggests collective rumor corrections by multiple parties may impede the spread of rumors. Non-correction messages, on the other hand, are made up of more diversified content comprising of the sender’s own composed messages. Contrary to [Maddock et al.’s \(2015\)](#) observation that correction tweets contain more original content than rumors, the local Twitter community in the study circulated the same corrective messages, which curtailed the scope for originality in such messages.

7.1 Theoretical implications

The study contributes to the field of information systems by framing the antecedents of online rumors in terms of information ambiguity and problems related to information diffusion. Additionally, it tested the validity of constructs traditionally used in the socio-psychological studies on rumor, such as emotions and source credibility, and found these to be pertinent to rumor correction messages. The proposed model along with the hypotheses have helped deepen our understanding of rumor research on social media.

7.2 Practical implications

The results of our study have implications for practice as well. Our study identified the content characteristics of corrections that are useful for any organization’s public relations or communication managers when planning a rumor rebuttal for public dissemination on social media. Unlike other online platforms for user-generated content where the message length is seldom a limitation, correction tweets on Twitter pose a challenge for practitioners to pack information succinctly into 140 characters. The most straightforward way of maximizing a message’s corrective effect, as demonstrated by our findings, is to avoid the use of emotive language and include specificity of information, reliability of source and supporting facts. Well-crafted messages that exploit the optimal characteristics of a correction tweet is more likely to constitute a stronger rebuttal. In times of social

crisis, effective rumor management through the execution of corrections on social media can help to provide reassurance, reduce tensions and restore public confidence.

8 CONCLUSION

Drawing on the third-person effect framework and earlier models of rumor diffusion and retransmission, this study developed and empirically tested a model of rumor correction relating to a death hoax of a politician on social media. The results indicated that the collective corrections by netizens and news media of a rumored death are characterized by low emotional levels, high source credibility and low content ambiguity. It appears that netizens who posted correction tweets were not driven by emotion. Instead, the content of the posts appear to be rational, unambiguous and is centred on a single purpose, that is, rumor rebuttal. A high volume of correction tweets were retweets from highly credible sources, such as the local mainstream newspaper, thereby suggesting the mainstream news media as the preferred conduit of rumor corrections. The results also revealed that visual imagery is, however, not a characteristic of rumor corrections.

The study contributes to research by identifying the key characteristics of rumor corrections in social media through a developed and empirically tested model. Building on previous research which focuses on rumor transmission and retransmission, this study provides fresh insights on reducing the spread of rumors. To the best of our knowledge, ours is the first predictive model developed to explain the phenomenon of online rumor correction. Another point of novelty in our work is the application of the third-person effect in explaining the motivation behind the collective corrective behavior of netizens. The analysis of this motivation behind corrective messages was not based on rumor theories, as they are usually restricted to explaining rumor transmission and retransmission alone. Whereas, an analysis of the content of the correction tweets revealed that a third-person effect was at work in such messages. Unlike many rumor studies that are set in the situational context of natural disasters, we based our study on the death hoax of a political figure, and from which we derived fruitful findings.

The findings here are, however constrained by the data that is openly available on Twitter and without any direct clarifications with the tweet authors. While the case study of a politician's death hoax presents a unique genre of social crisis, we could only focus on a single case as such incidents are rare to come by. Nevertheless this study presents opportunities for future research to improve the generalizability on predictions regarding rumor correction. Other researchers could extend our study by analysing the communication patterns in different types of social phenomena, such as natural disasters, health crisis, commercial crisis, and compare the differences in characteristics of each type. Moreover, extending this study to other types of social media platforms (e.g. Facebook) would be instructive to obtain a better understanding of rumor correction.

Acknowledgement

This work was supported by the Ministry of Education Research Grant AcRF Tier 2 (MOE2014-T2-2-020).

References

- Allport, F., & Lepkin, M. (1945). Wartime rumors of waste and special privilege: Why some people believe them. *Journal of Abnormal and Social Psychology*, 40 (1), 3-36.
- Allport, G., & Postman, L. (1947). The psychology of rumor. Henry Holt and Company, Oxford, England.
- Anthony, S. (1973). Anxiety and rumor. *Journal of Social Psychology*, 89 (1), 91-98.
- Bernard, S., Bouza, G., & Piétrus, A. (2014). An optimal control approach for E-rumor. *Revista Investigacion Operacional*, 36 (2), 108-114.

- [Bordia, P., & DiFonzo, N. \(2004\). Problem solving in social interactions on the Internet: Rumor as social cognition. *Social Psychology Quarterly*, 67 \(1\), 33-49.](#)
- [Bruns, A., Burgess, J. E., Crawford, K., & Shaw, F. \(2012\). # qldfloods and@ QPSMedia: Crisis Communication on Twitter in the 2011 South East Queensland Floods. Queensland University of Technology, Australia. Retrieved on 26 April 2016 from <http://eprints.qut.edu.au/48241/1/floodsreport.pdf>](#)
- [Chorus, A. \(1953\). The basic law of rumor. *Journal of Abnormal and Social Psychology*, 48 \(2\), 313-314.](#)
- [Cialdini, R.B. \(1993\). *Influence: The psychology of persuasion*. Morrow, New York.](#)
- [Davison, W.P. \(1983\). The third-person effect in communication. *Public Opinion Quarterly*, 47 \(1\), 1-15.](#)
- [DiFonzo, N., & Bordia, P. \(2007\). *Rumor psychology: Social and organizational approaches*. American Psychological Association, Washington, D.C.](#)
- [DiFonzo, N., Bordia, P., & Rosnow, R. L. \(1994\). Reining in rumors. *Organizational Dynamics*, 23 \(1\), 47-62.](#)
- [Donovan, P. \(2007\). How idle is idle talk? One hundred years of rumor research. *Diogenes*, 54 \(1\), 59-82.](#)
- [Dormann, C., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., . . . Lautenbach, S. \(2013\). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36 \(1\), 27-46.](#)
- [Garrett, R.K. \(2011\). Troubling consequences of online political rumoring. *Human Communication Research*, 37 \(2\), 255-274.](#)
- [Gunther, A.C. \(1995\). Overrating the X-rating: The third-person perception and support for censorship of pornography. *Journal of Communication*, 45 \(1\), 27-38.](#)
- [Gunther, A.C., Bolt, D., Borzekowski, D. L. G., Liebhart, J. L., & Dillard, J. P. \(2006\). Presumed influence on peer norms: How mass media indirectly affect adolescent smoking. *Journal of Communication*, 56 \(1\), 52-68.](#)
- [Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. \(2013\). *Multivariate data analysis*. Pearson Prentice Hall, Upper Saddle River, N.J.](#)
- [Hilligoss, B., & Rieh, S. Y. \(2008\). Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context. *Information Processing & Management*, 44\(4\), 1467-1484.](#)
- [Hwang, H., Pan, Z., & Sun, Y. \(2008\). Influence of hostile media perception on willingness to engage in discursive activities: An examination of mediating role of media indignation. *Media Psychology*, 11 \(1\), 76-97.](#)
- [Kimmel, A. J. \(2004\). *Rumors and rumor control*. L. Erlbaum Associates, Mahwah.](#)
- [Kimmel, A.J., & Audrain-Pontevia, A. \(2010\). Analysis of commercial rumors from the perspective of marketing managers: Rumor prevalence, effects, and control tactics. *Journal of Marketing Communications*, 16 \(4\), 239-253.](#)
- [Krippendorff, K. \(2012\). *Content analysis: An introduction to its methodology*. Sage, Thousand Oaks, CA.](#)
- [Kwak, H., Lee, C., Park, H., & Moon, S. \(2010\). What is Twitter, a social network or a news media? In *Proceedings of the 19th International Conference on World Wide Web*, Raleigh, NC.](#)
- [Leong, M.L.C., Pan, S. L., Ractham, P., & Kaewkitipong, L. \(2015\). ICT-enabled community empowerment in crisis response: Social media in Thailand flooding 2011. *Journal of the Association of Information Systems*, 16 \(3\), 174-212.](#)
- [Lin, Z., Pazos, R., & Benites, Y. \(2015\). 1923-2015 Lee Kuan Yew: How the Twittersphere reacted to the news. Retrieved August 5, 2015, from <http://leekuaneyew.straitstimes.com/ST/recap/index.html>](#)
- [Liu, F., Burton-Jones, A., & Xu, D. \(2014\). Rumors on social media in disasters: Extending transmission to retransmission. In *Proceedings of 18th Pacific Asia Conference on Information Systems*, Chengdu, China.](#)
- [Lombard, M. \(2002\). Content analysis in mass communication. *Human Communication Research*, 28 \(4\), 587-604.](#)

- Maddock, J., Starbird, K., Al-Hassani, H., Sandoval, D. E., Orand, M., & Mason, R. M. (2015). Characterizing Online Rumoring Behavior Using Multi-Dimensional Signatures. Paper presented at the 18th ACM Conference on Computer-Supported Cooperative Work and Social Computing, Vancouver, Canada.
- Murthy, D. (2011). Twitter: Microphone for the masses? *Media Culture and Society*, 33 (5), 779-789.
- Nadamoto, A., Miyabe, M., & Aramaki, E. (2013). Analysis of microblog rumors and correction texts for disaster situations. Paper presented at the 15th International Conference on Information Integration and Web-Based Applications and Services, iiWAS 2013, Vienna, Austria.
- Oh, O., Agrawal, M., & Rao, H.R. (2013). Community intelligence and social media services: A rumor theoretic analysis of tweets during social crises. *MIS Quarterly*, 37 (2), 407-426.
- Oh, O., Eom, C., & Rao, H. R. (2015). Role of social media in social change: An analysis of collective sense making during the 2011 Egypt Revolution. *Information Systems Research*, 26 (1), 210-223.
- Oh, O., Kwon, K.H., & Rao, H.R. (2010). An exploration of social media in extreme events: Rumour theory and Twitter during the Haiti earthquake 2010. Paper presented at the International Conference in Information Systems, St Louis, Missouri.
- Perloff, R.M. (1993). Third-person effect research 1983–1992: A review and synthesis. *International Journal of Public Opinion Research*, 5 (2), 167-184.
- Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology*, 34 (2), 243-281.
- Rieh, S., & Danielson, D. (2007). Credibility: A multidisciplinary framework. *Annual Review of Information Science and Technology*, 41 (1), 307-364. doi:10.1002/aris.2007.1440410114
- Rojas, H. (2010). 'Corrective' actions in the public sphere: How perceptions of media and media effects shape political behaviors. *International Journal of Public Opinion Research*, 22 (3), 343-363.
- Rosnow, R.L. (1974). Communications as cultural science. *Journal of Communication*, 24 (3), 26-38.
- Rosnow, R.L. (1991). Inside rumor: A personal journey. *American Psychologist*, 46 (5), 484-496.
- Seo, K.k.p.e., Dillard, J.P., & Shen, F. (2013). The effects of message framing and visual image on persuasion. *Communication Quarterly*, 61 (5), 564-583.
- Shibutani, T. (1966). Improved news. Bobbs-Merrill Company, Indianapolis, Indiana.
- Shklovski, I., Palen, L., & Sutton, J. (2008). Finding community through information and communication technology in disaster response. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, San Diego, CA.
- Starbird, K., Maddock, J., Orand, M., Achterman, P., & Mason, R.M. (2014). Rumors, false flags, and digital vigilantes: Misinformation on Twitter after the 2013 Boston Marathon Bombing. In *Proceedings of iConference 2014*, Berlin, Germany.
- Sun, Y., Pan, Z., & Shen, L. (2008a). Understanding the third-person perception: Evidence from a meta-analysis. *Journal of communication*, 58 (2), 280-300.
- Sun, Y., Shen, L., & Pan, Z. (2008b). On the behavioral component of the third-person effect. *Communication Research*, 35 (2), 257-278.
- Takayasu, M., Sato, K., Sano, Y., Yamada, K., Miura, W., & Takayasu, H. (2015). Rumor diffusion and convergence during the 3.11 Earthquake: A Twitter case study. *PLoS ONE*, 10 (4), 1-18.
- Tripathy, R. M., Bagchi, A., & Mehta, S. (2013). Towards combating rumors in social networks: Models and metrics. *Intelligent Data Analysis*, 17 (1), 149-175.
- Tripathy, R.M., Bagchi, A., & Mehta, S. (2010). A study of rumor control strategies on social networks. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, Toronto, Canada.
- Wathen, C.N., & Burkell, J. (2002). Believe it or not: Factors influencing credibility on the Web. *Journal of the American Society for Information Science and Technology*, 53 (2), 134-144.
- Wilson, P. (1983). Second-hand knowledge: An inquiry into cognitive authority. Greenwood Press, Westport, Connecticut.
- Zillmann, D. (2002). Exemplification theory of media influence. In *Media effects: Advances in theory and research* (Jennings, B. Ed.), pp. 19-41. 2nd Edition. L. Elbaum Associates, Mahwah, N.J.