Computers in Human Behavior 121 (2021) 106800

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: http://www.elsevier.com/locate/comphumbeh

Full length article

What matters most in the responses to political campaign posts on social media: The candidate, message frame, or message format?

Shu Yan Benson Lam [a], Mei Fung Meily Cheung [b], Wai Han Lo [c,*]

[a] Department of Mathematics, Statistics and Insurance, The Hang Seng University of Hong Kong, Hong Kong
[b] AI and Media Research Lab, School of Communication, Hong Kong Baptist University, Hong Kong
[c] Department of Journalism, Hong Kong Baptist University, Hong Kong

ARTICLE INFO

Keywords:
Elections
Social media
Power structures
Political campaigns
Social media data

ABSTRACT

The responses of social media users to politicians' different content strategies have tremendous implications for election campaigns. However, the identification of the factors that influence social media users' responses remains a challenge. In this paper, we combine partial least squares (PLS) with a cross-validation technique to automatically identify the type of social media posts that are associated with strong responses from social media users. This approach is different from other methods that require manual identification of these associations. This study examines 534 posts and 379,880 comments on three candidates' Facebook pages during the 2017 Hong Kong Chief Executive election. In the literature on message strategies and social media, this is one of the first studies to examine both the affective responses and cognitive expressions of social media users associated with different content strategies by politicians. The results show that the dominant candidate used his strong position to construct a strong online image, with posts generating more and longer comments and thus yielding more positive feedback. While media frames do not have any impact on either affective responses or the quality of comments, the use of videos and photographs elicits more responses. The findings contribute to the literature on online engagement and campaign strategy and support the argument that social media replicate offline power relations and offline responses to political messages.

1. Introduction

Online engagement during political campaigns has become a major concern for politicians, public relations professionals, and scholars of political communication (Nielsen & Vaccari, 2013; Utz, 2009). In 2008, Barack Obama became the first African American to win a US presidential election. One major factor in the success of the Obama campaign in mobilizing voters was the strategic use of social media. By November 2008, Obama had 2.5 million people following his Facebook account (Chang, 2009, pp. 1–40). The success of Obama established online strategy as an integral part of political campaigning (Enli, 2017; Ross, Fountaine, & Comrie, 2015; Skovsgaard & Dalen, 2013). While this trend has inspired much research, many questions remain unanswered and require further exploration. The innovation and normalization hypotheses offer two competing accounts of the extension of politics to online environments. The innovation hypothesis suggests that online communication activities and information made available only online engage Internet users and increase their level of interaction with political entities, thereby contributing to changes in political structures (Larsson, 2013; Vergeer et al., 2011). The Internet allows political candidates to engage in direct dialogue with their supporters (Powell and Cowart, 2003). Because Internet users are thus encouraged to participate in the political process (Bekafigo and McBride, 2013), "A fresh wave of technological optimism has accompanied the advent of social media" (Loader & Mercea, 2011, p. 758). This theory is extended by the equalization hypothesis, which suggests that the characteristics of online engagement assist more marginal political actors (Larsson & Moe, 2014; Lilleker & Vedel, 2013).

In contrast to the innovation and equalization hypotheses, some researchers have argued that online activities simply mirror offline power struggles and the offline distribution of power (Margolis et al., 1999). Klinger and Svensson (2015) called this the normalization hypothesis, and pointed to the online replication of the inequalities present in the offline political structure, with dominant political candidates and parties enjoying similar advantages in both online and offline communication activities. The normalization hypothesis has received

* Corresponding author.
E-mail addresses: bensonlam@hsu.edu.hk (S.Y.B. Lam), s194933@hkbu.edu.hk (M.F.M. Cheung), janetlo@hkbu.edu.hk (W.H. Lo).

https://doi.org/10.1016/j.chb.2021.106800
Received 4 August 2020; Received in revised form 26 January 2021; Accepted 27 March 2021
Available online 31 March 2021

substantially more empirical support than the equalization hypothesis, at least in relation to the use of political websites (Lilleker & Vedel, 2013). However, some scholars have argued that social media, with their features of greater accessibility and lower cost, offer more likely sites for political equalization (Gibson and McAllister, 2015; Gueorguieva, 2008). Therefore, there is a hope that social media platforms, such as Facebook, might facilitate online equalization, even though this was not achieved using websites. A few scholars have found that the use of Facebook offers marginal political players opportunities to promote their political ideas (Schweitzer, 2011; Souther, 2015), while others have found that using Facebook makes no difference in equalizing opportunities (Lev-On & Haleva-Amir, 2018).

The purpose of this study is to investigate whether the effects of Facebook campaigning reflect or shift the offline political relationships and power structure in Hong Kong. To verify if the normalization or innovation hypothesis can better explain the role of social media in political campaigns, we apply several data mining methodologies to analyze over 370,000 social media messages on candidates' Facebook pages for the 2017 Hong Kong chief executive election, together with 534 posts that they made during the election period. The use of partial least squares (PLS) with a cross-validation technique can effectively identify Facebook responses (cognitive and affective) that are highly correlated with online message frames and formats (i.e., text, photo, or video). After the identification, we use regression analysis to study the effects of individual online message frames and formats on the Facebook responses. Facebook is the most popular social media platform in Hong Kong, with an 82% penetration rate in 2019 (Statista, 2019). Each of the three chief executive candidates—Carrie Lam, John Tsang, and Woo Kwok-hing—used Facebook intensively to deliver messages and engage supporters during the election period. Their messages and interactions with Facebook users influenced the news agenda of the mass media (Lo et al., 2018). Therefore, the 2017 Hong Kong chief executive election provides an ideal context to study the variations in political campaign messages and their effects on Facebook users' responses.

This study therefore makes the following valuable contributions to the literature and to data analysis methodology. First, it provides evidence for the value of Facebook as a site for political equalization. As mentioned above, few studies have used Facebook to test the competing hypotheses on the implications of online activity for political campaigning (e.g., Lev-On and Haleva-Amir, 2016), with most previous studies focusing on websites. The few studies on the topic have also yielded very different conclusions. This gap in our understanding is particularly in need of attention in the context of Hong Kong, given the high penetration rate of Facebook (82%) in the city (Statisa, 2019). This paper thus provides useful insights from a high saturation location into the development of Facebook as a political tool.

Second, previous studies of these competing hypotheses have only compared the extent of the online activities of dominant and peripheral political players, testing for differences in social media engagement. There has been a lack of discussion of promotion strategies as factors in engaging Facebook users, and whether these are more important than the offline identity of the political players (Lev-On & Haleva-Amir, 2018). In other words, previous studies cannot answer the question of whether changes in the social media strategies of established and peripheral political players can generate changes in the engagement of Internet users or merely reflect the offline power balance. This study takes a further step by comparing the effect of online promotion strategies and the identity of political players on Facebook users' engagenment (Schweitzer, 2011; Lilleker & Vedel, 2013). Addressing these research problems is important in an era in which social media are popular and common platforms for political campaigns.

Third, this study evaluates the relative importance of online message frames and formats in altering the political landscape, thus expanding the scholarship that has examined the influence exerted by message frames (e.g., Brewer et al., 2016; Fernandes, Giurcanu, Bowers, & Neely, 2010; Lee and Soo, 2012) and message formats (Ksiazek, Peer, & Lessard, 2016; Ross et al., 2015). This helps us to understand the mechanics of using Facebook for political campaigns (Lev-On & Haleva-Amir, 2018). Fourth, different from other methods that require manual identification of the associations, the combined use of PLS and cross-validation techniques to examine 534 posts and 379,880 audience comments allows this study to automatically identify content that is strongly associated with the responses of social media platform users.

2. Literature review

2.1. Evaluating the innovation and normalization hypotheses

When seeking to understand the impact of Internet-based political practices, the normalization hypothesis and the innovation hypothesis are the two most popular approaches adopted by researchers (Larsson, 2013). Scholars arguing for the innovation hypothesis have claimed that the introduction of online political activities, which allow people to actively seek information, has facilitated democratic discourse and decentralized debate (Schweitzer, 2008; Verg et al., 2011; Vergeer and Hermans, 2013). Many politicians have made heavy use of the Internet as a political marketing tool to deliver their messages and to communicate with their supporters or with voters (Bright et al., 2017; Grusell & Nord, 2016; Quinlan, Gummer, Roßmann, & Wolf, 2017). Supporters of the innovation hypothesis have argued that campaigning on social media platforms in political elections has at least three advantages. First, candidates have more control over campaign messages (Broersma & Graham, 2012) and can therefore maximize favorable publicity and minimize negative exposure (McNair, 2003, p. 7). This helps candidates to gain voters' support and commitment to their election campaigns (Lin, 2017). Second, social media platforms offer candidates a means to bypass traditional media and deliver messages directly to target audiences (Broersma & Graham, 2012). Third, small political parties with limited budgets and individual candidates without party support can make good use of these inexpensive marketing platforms for election campaigning (Larsson, 2016).

Similar to the innovation hypothesis, advocates of the equalization hypothesis have argued that online activity helps peripheral political players to establish a presence and to build and reach an audience. They claimed that the use of the Internet for political messaging has thus promoted political equality among different players (Larrson and Moe, 2014). The equalization hypothesis assumes that peripheral political players work harder to promote themselves on the Internet and that this can compensate for their offline disadvantages. However, there is little evidence to support this hypothesis. Empirical studies have mainly compared the use of digital communications by the dominant and peripheral players in a specific campaign (Corrado and Firestone, 1997); some have also analyzed particular candidates to evaluate the success of peripheral political players who make use of the Internet (Hindman, 2005; Tremlett, 2015; De Rosa, 2013). However, as stated above, most of these studies focused on the use of websites, and politicians have now shifted the focus of their online presence from websites to social media (Lev-On, Under Review).

Supporters of the normalization hypothesis have questioned whether online activities increase political participation. They have suggested that the inequalities found in society in general are replicated on the Internet, and that political parties are using traditional campaign techniques rather than adapting new online campaigning strategies (Schweitzer, 2005). It has been argued that websites are not being used to encourage dialogue, with most politicians adopting a one-way communication approach to deliver messages (Margolis et al., 1999). A recent study on the use of Facebook by German and Austrian political parties during national election campaigns found that parties and candidates tended to post messages that emphasized emotional and visual elements rather than addressing political issues (Magin, Podschuweit, Haßler, & Russmann, 2017). Research has revealed that most candidates rarely reply openly to online comments, preferring to address private

messages on social media (Magin et al., 2017), and that political mobilization shifts from participation in offline political activities to sharing messages online (Magin et al., 2017; Wallsten, 2010).

Some scholars have argued that the online world reflects the socio-economic and political structures of the offline world (Margolis et al., 1999). From this perspective, because the Internet mirrors the features of the offline world, the dominant political actors and parties will enjoy the same advantages in constructing their images online that they enjoy offline (Klinger, 2013). Therefore, any investigation of the relationship between the features of social media posts and the responses that does not consider the relative standing of the candidates may be at risk of fallacy. The innovation and normalization hypotheses hold very different assumptions about online engagement. To evaluate the two hypotheses in the context of a political campaign, we test whether the candidate's relative standing in the offline world is the predominant factor in users' responses or if online strategies do in fact contribute to engagement with and response to political communications. The normalization hypothesis would be supported if the position of candidates within the existing power structure and offline power relations are found to be the major predictor of users' responses.

2.2. Hong Kong chief executive election

Hong Kong's chief executive is elected by an election committee of around 1200 people. The committee members are drawn from particular sectors of society and favor the central government's decisions rather than the interests of the Hong Kong people, despite Hong Kong's supposed high degree of autonomy under the Basic Law (Lo, Lam, & Cheung, 2019). Previous studies of Hong Kong chief executive elections have mainly focused on the Chinese-style democratic process and power struggles. They have rarely discussed the effectiveness of the election strategies chosen by the candidates or the use of social media platforms (e.g., Kan, 2012; Lo, 2017). This study uses the chief executive election of 2017 to examine whether the use of Facebook promotes online political equalization.

During the 2017 Hong Kong chief executive election, the Chinese University's Centre for Communication and Public Opinion Survey measured the popularity ratings of the three candidates. The poll found that John Tsang was the most popular candidate from the beginning of the election period and continued to extend his lead over Carrie Lam, gaining support from around 60% of the general public, with Carrie Lam's support at around 30% and Woo having the least support (Cheung and Chung, 2018). Despite the popularity of John Tsang, Carrie Lam, who was widely regarded as the candidate favored by the central government, was elected as chief executive in Hong Kong by the Election Committee (Lo et al., 2019). Based on the normalization hypothesis that the online world mirrors offline political realities, it is expected that the Facebook page of John Tsang would have received responses with more favorable affect and would have elicited more comments. The identity of the candidate is thus considered as an independent variable in this study to explore whether patterns of online responses are indeed predetermined by offline political realities.

From the perspective of the innovation hypothesis, the patterns of the online responses to the three candidates are dependent on the Internet strategies that they adopt. We focus on two content features: message frame and message format. Scholars who have examined variations in social media usage have either addressed the content (e.g., Brewer et al., 2016; Fernandes et al., 2010; Lee and Soo, 2012) or the format of social media posts (Ksiazek et al., 2016; Ross et al., 2015). Most of these studies adopted a narrow scope without considering the substantial variations in the elements of different social media campaigns (Xenos, Macafee, & Pole, 2017). Relatively little research has considered both the content and format of messages to better reflect their variations in relation to user reactions in social media campaigns. The studies of Kim and Yang (2017) and Srivastava, Saks, Weed, and Atkins (2018) are exceptions, but they focused on organizational and health communication, not framing within political campaigns. In the light of this review of the literature, this study examines both frames and message formats as two independent variables, with the aim of more comprehensively analyzing the messaging strategies of political campaigns on social media and how different strategies contribute to different responses.

2.3. Impact of Content Frames and formats on audience response

Most previous studies of political content framing have focused on how traditional news media have set the frame for people acting in a political context (Craft & Wanta, 2004; Fernandes & Shumow, 2016). The limited research into framing on online media consists of studies acknowledging the framing effect that political blogs, online forums, and social media can have on the traditional news agenda (Conway, Kenski, & Wang, 2015; Meraz, 2011; Neuman, Guggenheim, Jang, & Bae, 2014; Zhou & Moy, 2007). Other scholars have noted that online media also offer candidates or other disseminators of information a means to bypass the framing of news media by delivering information directly to an audience (Anderson, 2006) and providing a public space in which citizens can voice their opinions (Goldberg, 2011; Medina, Sánchez Cobarro, & Martínez, 2017). Variations in the message frames between candidates may result in different patterns of responses from social media users.

Three types of frames are commonly found in political campaigns: strategy, policy, and personal. The strategy frame is event oriented and focuses on campaign strategies, the performance of candidates, and the win-or-lose aspect of the election. It can include polling results and emphasizes conflict and drama alongside election results (Iyengar, 1991; Bennett, 1988; Rhee, 1997). The policy frame is related to the solutions and policies proposed by the candidates to deal with societal problems (Graber, 1993; Jamieson, 1992; Patterson, 1993; D'Angelo et al., 2005). The personal frame highlights the personal lives, personal identities, lifestyles, and characters of the people involved (McGregor, Lawrence, and Cardona, 2015). With the different thematic elements, rhetorical styles, and narrative strategies adopted in the three frames, it is predicted that each frame may have a different impact on a political campaign. A further category of post used in this study is "off-topic," which refers to the irrelevant posts or announcements (such as "Happy Valentine's Day!") that do not fit into a strategy, policy, or personal frame (McGregor, Lawrence, & Cardona, 2017) (see Table 1).

In addition to the framing of posts, researchers have also identified differences in the effects of various message formats. Some scholars found that the use of multimedia components facilitates positive feedback toward a computer system because the presence of such components cause the audience to be more involved with the program (Street and Manning, 1997). Biocca (1997) argued that audiences become more fully immersed in multimedia content because it stimulates a variety of senses. Reeves and Nass (1996) argued that an audience may, at least momentarily, feel a sense of presence or "being" in a mediated world, as

Table 1
Four categories of message frames.

Strategy frame	Posts that are event-oriented, focusing on campaign strategies, endorsements, or volunteers, or a candidate's performance, character, or personality (Bennett, 1988; Enli, 2017; Iyengar, 1991; McGregor et al., 2017; Rhee, 1997; Ross et al., 2015).
Policy frame	Posts discussing social problems and solutions, including the proposals or policies suggested by candidates (Graber, 1993; Jamieson, 1992; Patterson, 1993; D'Angelo et al., 2005).
Personal frame	Posts that are candidate-centered, focusing on candidates' performance, characteristics, and personality (McGregor, Lawrence, and Cardona, 2015).
Off-topic category	Ensures the exhaustiveness of the categories by capturing irrelevant posts or announcements (such as "Happy Valentine's Day!") that do not fit into the strategy or policy frames (McGregor et al., 2017).

the formal features of the medium generate the perception that they are transported into the world being portrayed. Steuer (1992) proposed that vividness plays a key role in creating such a sense of presence. The vividness of a medium stems from its sensory breath and sensory depth: if the medium involves a greater number of the senses and a higher resolution of modality, it creates a relatively higher level of vividness. Other studies have found that media featuring a high level of vividness and creating a greater degree of presence among the audience generates a more positive attitude toward the content (Cheng & Lo, 2015; Lo and Cheng, 2015, 2020). Ardito et al. (2019) found that the presence of visual elements, such as videos and images, predicts positive engagement in a social network. Runs and Cameron (2006) found that the use of imagery could predict user engagement, while Kite, Foley, Grunseit, and Freeman (2016) reported that posts containing videos receive more "likes" and comments and are shared more widely than those only featuring photos or text.

Numerous studies on social media response have used likes, shares (or re-tweets), and comments as indicators. For example, Zhang, Peng, Zhang, Wang, and Zhu (2014) examined how content and contextual factors of posts on microblogging sites were related to the popularity of the posts; they used the counts of re-tweets and comments as a measure of popularity. Xenos et al. (2017) studied how variations in posting behavior by candidates were associated with variations in response, which was measured by the numbers of likes and comments. Similarly, Lev-On and Haleva-Amir (2018) quantified the scope of campaign engagement on social media as the numbers of likes, comments, and shares. Kim and Yang (2017) argued that these three types of responses are actually distinct, with likes representing an affective type of response (Kim & Yang, 2017) and other responses being more akin to acknowledgments (Srivastava et al., 2018). This study adds two affective reactions specific to Facebook: the "Love" emoji is counted as a favorable affective response and the "Angry" emoji is counted as an unfavorable affective response. Past studies of online political activity have provided more support for the normalization hypothesis than for the equalization theory (Jungheer, 2014; Kim & Yang, 2017; Lilleker & Vedel, 2013). Recently, Rauchflesich and Metag (2020) found that the normalization tendency could still be observed in Twitter. Therefore, based on the assumption of the normalization hypothesis that the audience's affective responses are subject to the identity of the candidates, the following hypothesis is formed:

H1: The identity of the candidate of the page on which messages are posted has a greater influence than media formats and media frames on Internet users' affective responses, including (a) the number of favorable reactions ("Like" and "Love" emoji) and (b) the number of unfavorable reactions ("Angry" emoji).

Previous research has found that the most popular posts on social media by candidates are those dealing with personal stories and campaign events rather than issues or policies (McGregor et al., 2017; Ross and Comrie, 2015). Both the strategy and policy frames are both widely used in news coverage, but studies have shown that the strategy frame is more common, especially near the time of a vote (Dekavalla, 2016; Dunaway & Lawrence, 2015). Aalberg, Stromback, and de Vreese (2012)Aalberg, Stromback, and de Vreese (2012) suggested that the prevalence of the strategy frame reflects the more strategic nature of contemporary campaigns (Aalberg et al., 2012; Dekavalla, 2016). The apparently more appealing nature of strategy framed and personal framed content (Iyengar et al., 2004; McGregor, Lawrence, and Cardona, 2015) suggests that the adoption. Su, Liu, and McLeod (2019) found that perceptions of the issue framing influenced sharing intentions. Taking this alongside the abovementioned positive effect of vividness and presence in media, we propose the following hypothesis.

H2: The candidates' message strategies on Facebook (frames and formats) contribute to variations in (a) the number of favorable responses ("Like" and "Love" emoji) and (b) the number of unfavorable reactions ("Angry" emoji).

As mentioned above, most social media studies measure user responses in terms of number of likes, shares, and comments (Xenos et al., 2017). However, this can only reflect the quantities of user reactions and not the quality of political expression, which may be a more suitable indicator for understanding the public sphere offered by social media. Some scholars have argued that liking a post is a kind of affective response, while leaving a comment is a form of cognitive feedback (Kim & Yang, 2017; Srivastava et al., 2018). Comments are recognized as engagement at a cognitive level, ranging from simple to complex thoughts, and are thus believed to have a higher value than likes (John, Mochon, Emrich, & Schwartz, 2017; Kim & Yang, 2017; Srivastava et al., 2018). Scholars have argued that comments are a kind of interpersonal communication (Ballantine, Lin, & Veer, 2015; Smock, Ellison, Lampe, & Wohn, 2011) or mass personal communication (O'Sullivan and Carr, 2017; Walther, Carr, & Choi, 2010). These cognitive expressions may affect an audience's perception of a post (Waddell & Sundar, 2017). Accordingly, comments are important as indicators of the quality of the political expressions received by politicians in their online election campaigns and as reflections of the nature of the public sphere on social media. Past studies found that Facebook was having a positive influence in increasing the civility of political discussion (Halpern, Gibbs, and Rowe, 2015).

The length of a comment can be considered one indicator of the quality of cognitive engagement. Arguing that longer messages could accommodate more complex sentences and ideas, Halpern and Gibbs (2013) used message length as a measure of the level of deliberation in user comments on Facebook and YouTube. To further reflect the quality of cognitive engagement with a post, this study expands the scope of cognitive expressions by counting the number of comments, the number of words in each comment, and the number of long comments. Based on the greater support for the normalization hypothesis, the effects of the identity of the candidates on cognitive responses is again compared with those of media formats and frames. To evaluate the normalization hypothesis in relation to audiences' cognitive responses, the following hypothesis is formed.

H3: The identity of the candidate represented on the page on which messages are posted has a greater influence than media formats and media frame on the quality of cognitive responses, as measured by (a) the number of comments, (b) the number of words in the comments, and (c) the number of long comments.

Studies supporting the innovation hypothesis have suggested that the adoption of different frames in media coverage or messages can bring about very different cognitive outcomes. Strategy frames and personal frames may direct people to focus on process rather than substance (De Vreese, 2012) and reduce people's levels of political trust, political participation, and engagement (Pedersen, 2012; Shehata, 2014). Policy or issue frames guide people to focus on the substance of policies and may contribute to political participation and substantive deliberation (Lawrence, 2000; Shehata, 2014). However, as the personality- and strategy-framed messages posted on political candidates' social media, which focus on the positive side of the candidate, are largely different from the frames used in news reports, and strategy and personal frames are more appealing to the audience than issue frames, it is questionable whether issue frames are better at promoting substantive deliberation.

While the studies mentioned above have found that the presence of multimedia components make an audience more involved and engaged with a computer system, no prior research has examined whether the use of multimedia components produces better quality comments. Accordingly, the following hypothesis is formed.

H4: The candidate's message strategies on Facebook (frames and formats) have an influence on (a) the number of comments and (b) the number of long comments.

2.4. Social media data mining

Table 2 provides a brief review of social media data mining methods that have been used in three studies exploring different topics related to

Table 2
Overview of data mining methodologies for social media research.

Source	Vicario, Gaito, Quattrociocchi, Zignani, and Zollo (2017)	Marozzo and Bessi (2018)	Morgan, Shafiq, and Lampe (2013)	This study
Data	Twitter and Facebook	Twitter	Twitter	Facebook
Area	Community Detection in Social Media	Polarization of Social Media Users	Ideology of News Outlets in Social Media	Political Expression in Social Media
Data Analysis Strategies				
Data Mining Methodology	Community Detection Algorithm	Random Forest	KS Test, Logistic Regression	Partial Least Square, Cross Validation, Regression Analysis
Automatic variable selection	✗	✗	✗	✓
Statistical inference	✗	✗	✓	✓
Model evaluation	✗	✓	✓	✓

Twitter and Facebook. The differences in the method adopted in this study and those of the other three studies are as follows: (i) automatic variable selection, (ii) statistical inference, and (iii) model evaluation.

(i) Automatic variable selection

One of the advantages of the method used in this study over other related methods is that it can automatically select dependent and independent variables that are highly correlated to each other. The study uses PLS with a cross-validation technique. PLS is used to identify the associations between every pair of the dependent and independent variables and the cross-validation technique identifies the highly correlated connections. This approach is different from other methods that require manual selection of the dependent and independent variables. Morgan et al. (2013) studied the effects of the perceived ideology of news outlet on consumption and sharing of news on Twitter. They manually selected 12 news outlets and 3 news categorizations as two different dependent variables. They then applied the Kolmogorov–Smirnov (KS) test and logistic regression to study the relationship between seven different traits per tweet and these two dependent variables. Marozzo and Bessi (2018) analyzed the polarization of social media users and news sites for the Italian constitutional referendum. They employed the random forest machine learning method to predict the polarization of users, which required the manual selection of useful historical information. Vicario et al. (2017) employed community detection algorithms to cluster Facebook pages and Twitter accounts with common users. They studied different characteristics of the detected communities but did not include any dependent variables.

(ii) Statistical inference

Statistical inference is the use of statistics to draw conclusions about some unknown aspect of a population based on a random sample from that population. This study adopted ordinary linear regression to infer the views of Hong Kong Facebook users towards the HK chief executive election based on a sample collected from Facebook pages. We found that author and message format had significant effects on both affective and cognitive expressions. These results were drawn by hypothesis testing. Morgan et al. (2013) used the KS test to analyze the relationship among the 12 selected news outlets and 3 news categorizations. They also adopted logistic regression and applied hypothesis testing to infer that the selected seven different traits per tweet had significant effects on the two selected dependent variables. However, Marozzo and Bessi (2018) and Vicario et al. (2017) only used data visualization to show the characteristics of the data.

(iii) Model evaluation

Goodness of fit is a measure that evaluates how well a model fits a set of observations. We used condition numbers, Q–Q plots, residual plots, and R^2 to evaluate the goodness of fit of the ordinary regression models. The condition numbers, QQ plots, and residual plots were used to show that the fitted regression models fulfilled the key assumptions of the regression models. The R^2 represents the proportion of variance in the dependent variable that the dependent variables explain well. Other studies have used different methods for evaluation. Marozzo and Bessi (2018) divided the data into different periods and used the previous time periods to predict the polarization of users in the next period. They then compared the predicted polarization of users with the true polarization of users. Morgan et al. (2013) adopted pseudo R^2 to evaluate the logistic regression models. However, Vicario et al. (2017) focused on data visualization to show the characteristics of the data.

3. Methodology

This study analyzed the posts published on the Facebook campaign pages for candidates in the 2017 Chief Executive election in Hong Kong and the user comments to these posts. Data were collected for the period from January 1, 2017 (when the last candidate began running for the election) to March 31, 2017 (the fourth day after the election). Using Facepager to extract data from Facebook, 509 candidate posts (excluding outliers) and 379,880 user comments (excluding image icons) were taken from the official pages of the three candidates—Carrie Lam, John Tsang, and Woo Kwok-hing—during the 3-month election period. The study combined traditional and computational content analysis methods and drew on a huge amount of social media data to examine the candidates' social media posts and the comments they generated. Content analysis is widely adopted to study frames (e.g., Dekavalla, 2016; Shehata, 2014) and the analysis of social media comments on political campaign pages (e.g., Hagar, 2014; Woolley, Limperos, & Oliver, 2010). Automated processing can also help extract useful information from articles in a more objective manner (e.g., Hong & Nadler, 2012; Song, Dai, & Wang, 2016). While human coders were responsible for coding message frames and formats, a computational method was used to count user responses and to code the comments. The researchers created a codebook containing guidelines and reference materials for the coding process. Message frames were identified by the presence of certain textual features, including the choices of words and phrases (Rhee, 1997). After training, one coder was responsible for coding all of the posts and a second coder then coded a random sample comprising 30% of the posts. Cohen's Kappa tests yielded an overall intercoder reliability of 0.88. A computational method was used to measure over 370,000 user comments. We also validated the key assumptions of the ordinary least squares (OLS) regression model (see Table 3).

3.1. independent variables

The independent variables were candidate, format of candidate post, and message frame. For each of the 509 posts, the candidate was Carrie Lam (Lam), John Tsang (Tsang), or Woo Kwok-hing (Woo). The message formats were text, photo, and video.

Table 3
Model validation and checking.

Regression of the Dependent Variable	"Like"	"Love"	"Angry"	No. of Comments	Total Word Count	Long Comments
Condition Number	14.31	10.34	8.06	9.94	20.29	7.95
R^2	0.67	0.98	0.97	0.98	0.99	0.95

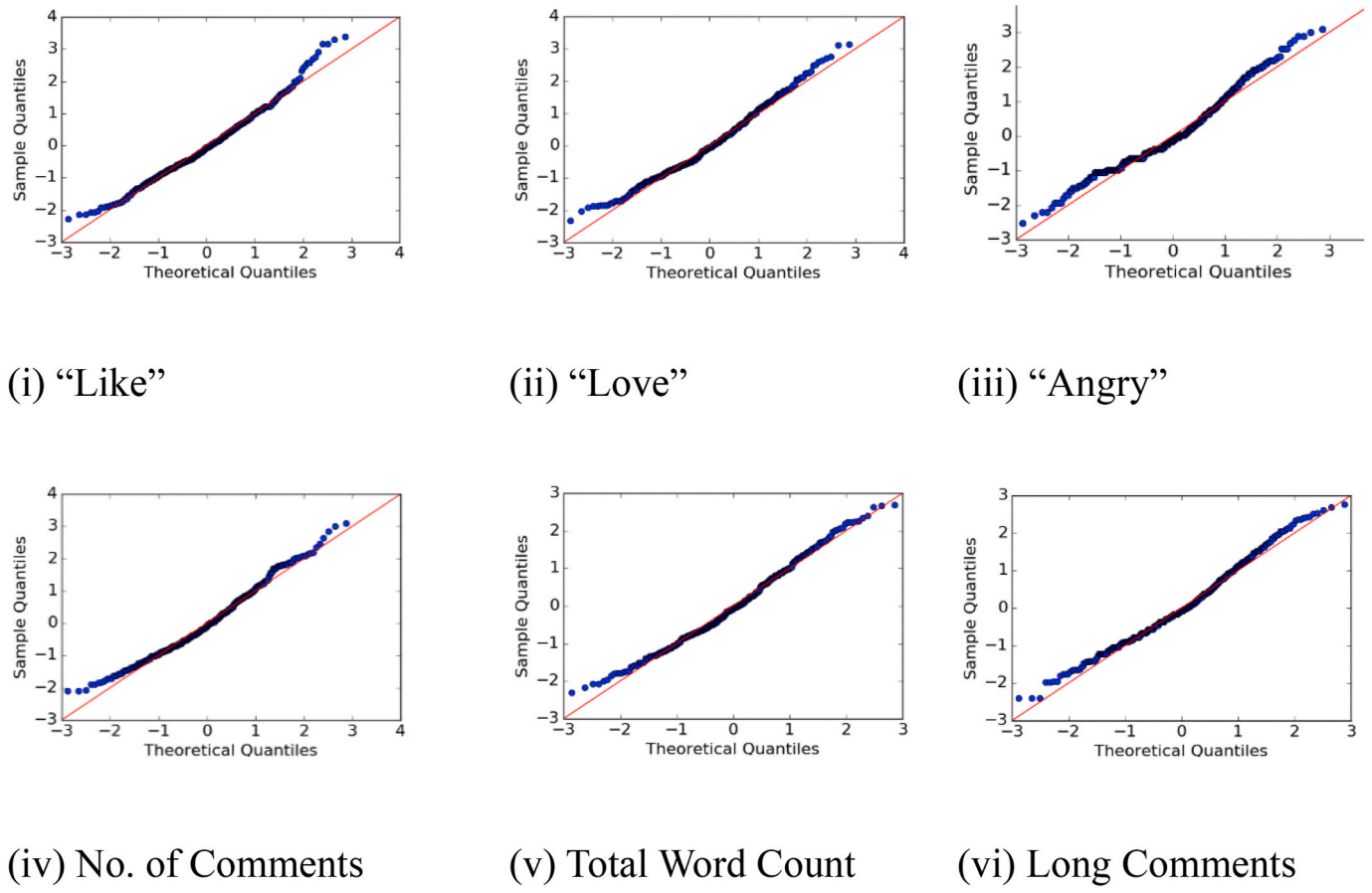


Fig. 1. QQ plots for the six regression models.

Content Frames. Coders were trained to classify candidates' posts into one of four *message frames*: strategy frame, issue frame, personal frame, or off-topic (see Table 1). Coders were asked to determine which category was the best fit for each post. Cohen's Kappa tests yielded an intercoder reliability of 0.84.

Content Formats. Coders counted the presence (yes or no) of different formats (text, photo, and video) and the average Kappa value was 0.88. There were few text-only posts ($n = 16$), and more posts featured both text and photos ($n = 303$) than text and video ($n = 153$).

3.2. Dependent Variables

Affective Response. To measure user responses, we used computer-assisted content analysis to count the number of *favorable Facebook reactions* ("Like" and "Love" emoji) and *unfavorable Facebook reactions* ("Angry" emoji).

Quality of Comments. With computer assistance, we also recorded the *number of user comments*, *number of words in the comments*, and *number of long comments*. The number of comments refers to the total number of user comments on a post. The number of long comments refers to the number of user comments of more than 30 words.

Partial Least Squares (PLS) and Ordinary Least Squares (OLS) regression.

Hypotheses 1–4 were tested with PLS and OLS regression by first automatically selecting the independent variables and dependent variables with a high statistical correlation. The analysis provides statistical evidence to verify which of the innovation or normalization theories best fits the data from the Hong Kong Chief Executive election. A high correlation between the candidate and the dependent variables would indicate that the normalization hypothesis fits the situation well, and a high correlation between the message frames and formats would indicate that the innovation hypothesis fits the situation well. After selecting the independent and dependent variables, we studied the relationship between the selected independent variables and each of the selected dependent variables using OLS regression.

(i) Selecting Independent and Dependent Variables Using PLS Regression

PLS regression assigns different weightings to the independent and dependent variables such that the more highly weighted dependent variables and independent variables have the largest statistical correlation. Larger weights thus indicate larger connections between the independent and dependent variables. To identify the number of connections and select the most relevant dependent and independent variables, we incorporated the machine learning technique of cross

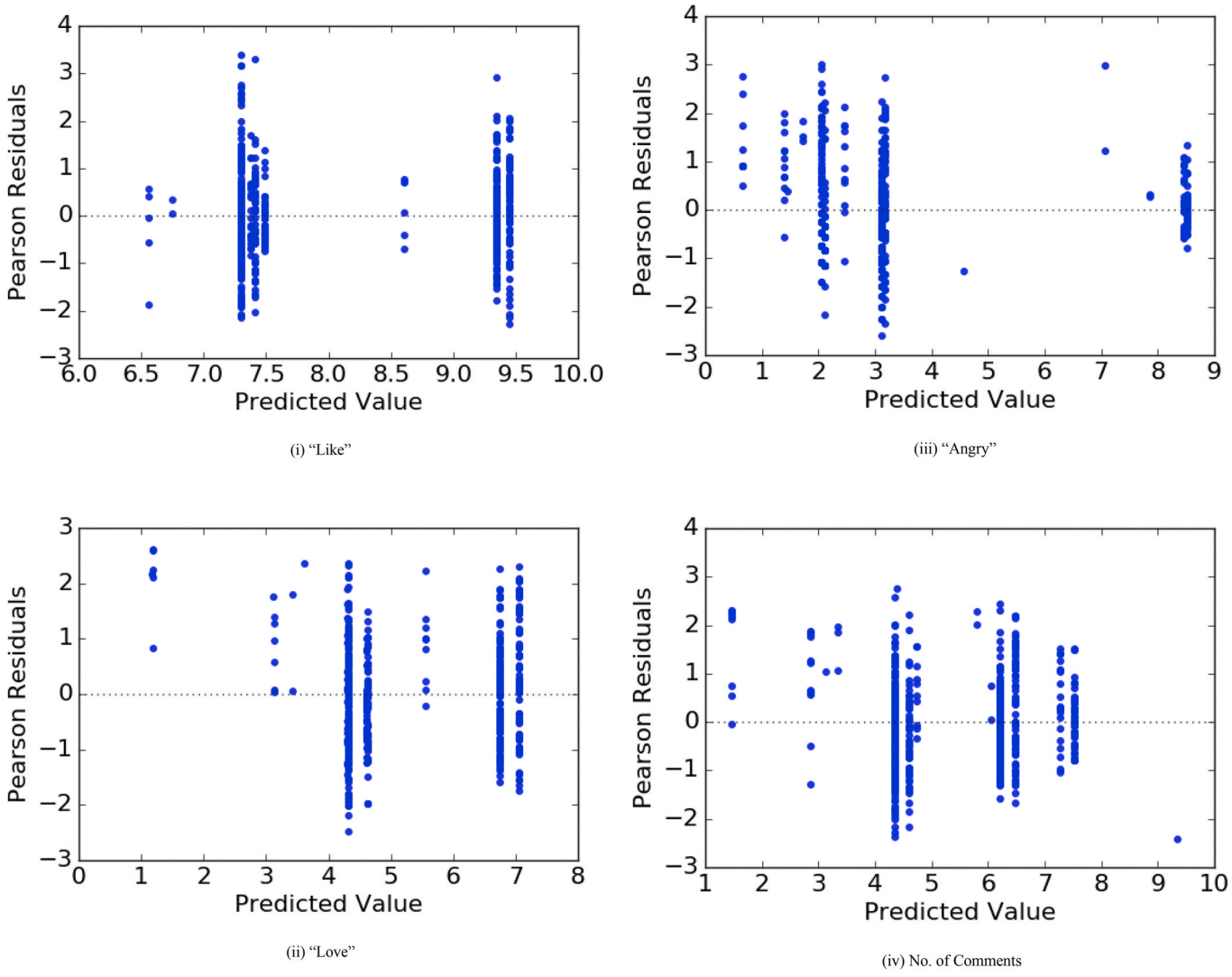


Fig. 2. Residual Plots (Predicted Values vs. Pearson Residuals) for the Six Regression Models.

validation into the PLS regression. First, the whole dataset, including dependent and independent variables, were randomly split into two halves. One half was used to train the PLS regression while the other half was used to test the accuracy of the trained model. Subsequently, we tried different numbers of connections to identify the number that gave the highest accuracy on the test dataset. Finally, the variables were selected for large weightings obtained by the PLS regression. The whole procedure was then repeated with a different method of random splitting. We then examined if the detected number of connections and selected variables were the same. In our experiments, all random splits gave the same result. We detected two connections, shown together with the selected variables in Table 4. The two connections showed that candidate (Lam, Tsang, and Woo) and format (text, photo, and video) were most closely related to the selected dependent variables. Thus, we treated the candidates and message formats as the independent variables and the six selected variables shown for the two connections in Table 4 as the dependent variables. The six dependent variables required the construction of six OLS regression models, the details of which are given in the next subsection.

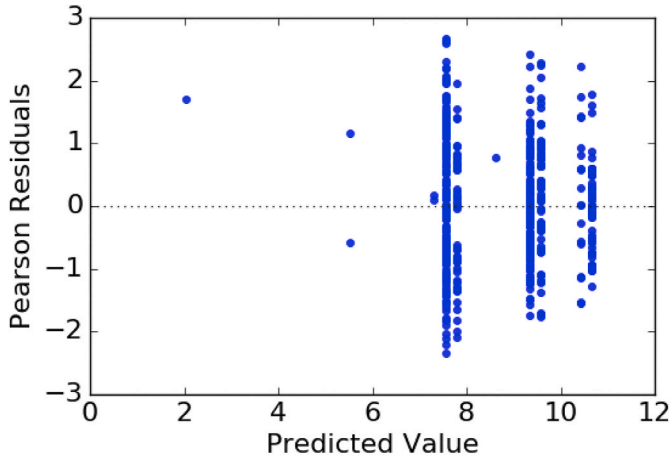
Candidate and message format were the independent variables for the study. However, the fact that each Facebook post was published on the page of one of the three candidates violated one of the key assumptions of OLS regression, which is the independence among the independent variables. To overcome this violation, we introduced the following regression model:

$$y = \delta_{Lam}x_{Lam} + \delta_{Tsang}x_{Tsang} + \beta_{Woo-Text}x_{Text} + \beta_{Woo-Photo}x_{Photo} + \beta_{Woo-Video}x_{Video} \quad (1)$$

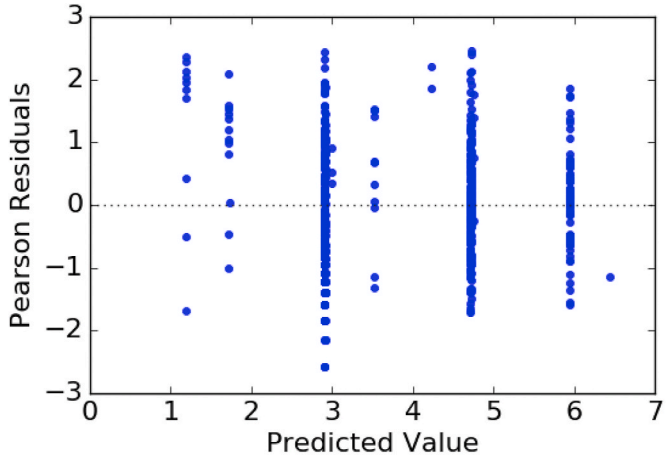
Here, y is the dependent variable, and this is taken as the log transformation of the selected dependent variables shown in Table 4. $x_{Lam}, x_{Tsang}, x_{Text}, x_{Photo}, x_{Video}$ are the indicator variables, with a value of either 0 or 1. The values of the variables ($x_{Lam}, x_{Tsang}, x_{Text}, x_{Photo}, x_{Video}$) were set to 1 if the post was published by Lam, published by Tsang, published in text format, published in photo format, or published in video format, respectively; otherwise, they were set to 0. The variables δ_{Lam}, δ_{Tsang}, $\beta_{Woo-Text}$, $\beta_{Woo-Photo}$, and $\beta_{Woo-Video}$ are the regression coefficients. This regression shows that the baseline of the Facebook post is published by Woo. If a post is published by Woo, we must have $x_{Lam} = x_{Tsang} = 0$. The regression model then becomes

$$y = \beta_{Woo-Text}x_{Text} + \beta_{Woo-Photo}x_{Photo} + \beta_{Woo-Video}x_{Video}. \quad (2)$$

In other words, the coefficients $\beta_{Woo-Text}, \beta_{Woo-Photo}$, and $\beta_{Woo-Video}$ are the regression coefficients for text, photo, and video messages that are published by Woo, respectively. For the posts that are published by Lam, we have $x_{Lam} = 1$ and $x_{Tsang} = 0$. The regression model then becomes



(v) Total Word Count



(vi) Long Comments

Fig. 2. (*continued*).

Table 4
Selected independent and dependent variables using PLS regression.
(ii) Analyzing the Connections Using OLS Regression

Selected Independent Variable	Selected Dependent Variable
First Connection Found by PLS Regression	
(1) Lam; (2) Woo; (3) Photo; (4) Video	(1) "Angry"; (2) Number of Comments; (3) Number of words in the comments; (4) Number of long comments (messages with over 30 words)
Second Connection Found by PLS Regression	
(1) Lam; (2) Tsang; (3) Woo	(1) "Like"; (2) "Love"; (3) "Angry"

$$y = \delta_{Lam} + \beta_{Woo-Text}x_{Text} + \beta_{Woo-Photo}x_{Photo} + \beta_{Woo-Video}x_{Video}. \quad (3)$$

This means that the effects of text, photo, and video messages of Lam on the dependent variables are $\delta_{Lam} + \beta_{Woo-Text}$, $\delta_{Lam} + \beta_{Woo-Photo}$, and $\delta_{Lam} + \beta_{Woo-Video}$, respectively. The regression coefficient δ_{Lam} represents the difference in the impacts of Lam's and Woo's posts. If δ_{Lam} was significantly greater than zero, Lam's posts had, in general, a more positive impact on the dependent variable than Woo's posts. If δ_{Lam} was significantly less than zero, Lam's posts had, in general, a more negative impact on the dependent variable than Woo's posts. Similar to Lam, the regression coefficient δ_{Tsang} represents the difference in the impact of Tsang's posts and Woo's posts.

4. Results

To verify the four hypotheses (H1–4), we considered the analysis results including the regression coefficients, their *p*-values and their confidence intervals, as shown in Table 5.

Verification of H1.

We first analyzed the results of the affective responses to address H1, testing whether appearing on a particular candidate's page had a greater influence than media formats and media frames on Internet users' affective responses. The corresponding dependent variables for affective response were favorable responses ("Like" and "Love") and unfavorable responses ("Angry").

The impact of Woo's posts on these three dependent variables was first examined. *Woo-Photo*, *Woo-Text*, and *Woo-Video* were all types of messages that Woo could have posted on his Facebook page and were the independent variables. The *p*-values of these three independent variables were all less than 0.001, implying that Woo's posts had a significant effect on the three dependent variables ("Like", "Love," and "Angry").

The impact of the posts published by Woo was used as a reference to analyze the impact of Tsang's posts on the dependent variables. The regression coefficients of the variable *Tsang* were 2.04, 2.43, and 1.06 for "Like", "Love," and "Angry," respectively. These three quantities represent the difference in impact of the posts published by Woo and

Table 5
Regression Coefficients for the Six Regression Models. *p*-values are shown in round brackets. The 95% confidence intervals are shown in square brackets.

Dependent Variable	Lam	Tsang	Woo-Photo	Woo-Text	Woo-Video
"Like"	0.08 (0.46) [-0.13 0.29]	2.04 (0) [1.91 2.18]	6.56 (0) [6.17 6.95]	0.74 (0) [0.35 1.13]	6.67 (0) [6.26 7.09]
"Love"	−0.01 (0.90) [-0.26 0.23]	2.43 (0) [2.27 2.59]	3.13 (0) [2.79 3.46]	1.18 (0) [0.85 1.51]	3.44 (0) [3.07 3.81]
"Angry"	6.41 (0) [6.21 6.60]	1.06 (0) [0.93 1.19]	1.39 (0) [1.18 1.61]	0.66 (0) [0.45 0.88]	1.45 (0) [1.21 1.70]
Number of Comments	3.00 (0) [2.74 3.25]	1.89 (0) [1.72 2.05]	2.95 (0) [2.61 3.28]	1.19 (0) [0.85 1.52]	3.35 (0) [2.98 3.72]
Total Word Count of Comments	2.87 (0) [2.61 3.13]	1.79 (0) [1.62 1.96]	5.52 (0) [4.84 6.20]	2.03 (0) [1.36 2.71]	5.76 (0) [5.06 6.45]
Long Comments (>30 words)	3.04 (0) [2.77 3.30]	1.81 (0) [1.63 1.98]	1.72 (0) [1.43 2.01]	1.19 (0) [0.90 1.47]	1.73 (0) [1.40 2.06]

Tsang. Their *p*-values were all zero. This indicates that Tsang's posts had an effect on the dependent variable and this effect was higher than for Woo's posts.

Similar to the variable *Tsang*, the variable *Lam* represented the difference in impact of the posts published by Woo and Lam. The coefficients were close to zero for favorable responses (0.08 for "Like" and −0.01 for "Love"). The *p*-values for favorable responses were greater than 0.05, while the *p*-value for the unfavorable response was less than 0.05. This indicates that the effects of posts by Lam was similar to those by Woo for favorable responses and greater than Woo for unfavorable responses.

In response to H1a, Tsang's posts were more attractive and provoked more favorable responses than those of Lam and Woo, as measured by "Like" and "Love" emoji. For H1b, we observed that the regression coefficients for the variables Lam and Tsang were positive (6.41 for Lam and 1.06 for Tsang) with zero *p*-values. The confidence interval for the variable Lam was [6.21, 6.60], which was beyond the confidence interval for Tsang [0.93, 1.19]. In other words, Lam's posts were the most impactful for unfavorable responses, followed by Tsang and lastly Woo. Therefore, H1 is supported.

4.1. Verification of H3

We then analyzed the results for cognitive responses to address H3, testing whether appearing on a particular candidate's page influenced cognitive responses to a post. The number of comments, total word count of comments, and number of long comments were the three dependent variables for H3a, H3b, and H3c, respectively.

Similar to the testing of H1, the impact of Woo's posts was first examined and the *p*-values of the independent variables, *Woo-Photo*, *Woo-Text*, and *Woo-Video*, for the three dependent variables were all zero. The regression coefficients of the three variables were positive, indicating that Woo's posts had a positive effect on the three dependent variables.

The regression coefficients for the variables *Lam* and *Tsang* were all positive, with *p*-values of zero. This implies that the posts of Lam and Tsang had a higher impact on the three dependent variables than Woo's posts. The confidence intervals of the variable *Lam* were all beyond the confidence intervals of those of the variable *Tsang*. For example, the confidence interval of *Lam* for the dependent variable number of comments was [2.74, 3.25] while the confidence interval of *Tsang* for the same variable was [1.72, 2.05]. The smallest value in the confidence interval for *Lam* (2.74) was larger than the largest value in the confidence interval for *Tsang* (2.05). This indicates that the impact of Lam's posts was higher than that of Tsang's posts. Thus, H3 is supported.

4.2. Verification of H2 and H4

We first analyzed the impact of the message format of Woo's posts on affective responses ("Like", "Love," and "Angry") and cognitive responses (number of comments, total word count of comments, and number of long comments). We observed that the *p*-values of *Woo-Photo*, *Woo-Text*, and *Woo-Video* for all six dependent variables were all zero, indicating that the impact of each message format was significant on all dependent variables. The regression coefficients of *Woo-Video* were the largest, followed by those of *Woo-Photo* and *Woo-Text*. The confidence intervals of *Woo-Video* and *Woo-Photo* were beyond those of *Woo-Text*, indicating that *Woo-Video* and *Woo-Photo* had more impact. There was overlapping between the confidence intervals of *Woo-Video* and *Woo-Text*. In other words, the impacts of these two formats were similar.

Redoing the regression analysis for *Lam* and *Tsang*, the results showed that the degrees of impact for Lam's formats and Tsang's formats were the same as for Woo's formats. Therefore, the impact of photo and video on any dependent variable was similar, whereas the impact of text was much lower than that of photo and video.

For the message frames, only the three candidates and formats of the posts were selected in the PLS regression. This may suggest that the connections between the message frames and the dependent variables were not strong, with the message frames having little or no impact on either the favorable or unfavorable responses, or on the number of comments overall and the number of long comments.

To further verify these relationships, we adopted the following statistical hypotheses to test if messages frames were significant independent variables for H2 and H4.

H_0 : With the message frames in the regression as independent variables.

H_1 : Without the message frames in the regression as independent variables.

The null hypothesis states that the four message frames (strategy frame, issue frame, personal frame, and off-topic) were incorporated into the regression model as the independent variables (i.e., H_0). Thus, there were a total of nine independent variables: the three candidates, three formats, and three message frames. The alternative hypothesis is the regression as it was performed above (i.e., H_1). The rejection of the null hypothesis would support the relatively unimportant role played by the message frames compared with the candidates and formats of the posts.

There were six dependent variables, so six pairs of statistical hypotheses were tested, with the results shown in Table 6. All of the statistical null hypotheses were rejected. In other words, in addressing H2 and H4, it was found that the message frames adopted by a candidate had no impact on favorable responses, unfavorable responses, number of comments, number of words in the comments, or number of long comments, whereas the message formats did have an impact. Thus, H2 is partially supported and H4 is rejected.

5. Discussion

With the heavy use of social media in political campaigning all over the world, there is a need to examine whether message strategies affect user responses on social media platforms or if online communication activities are merely mirroring offline power structures. Adopting the normalization hypothesis and innovation hypothesis as the theoretical framework, this study examined whether Facebook is a site for political equalization or an online platform that merely replicates offline power inequalities. The use of PLS together with a cross-validation technique also contributes to the literature by examining the effect of both online (message frames and message formats) and offline (candidate identity) factors on user responses. The study has practical insights for politicians

Table 6
p-values of the Statistical Hypothesis Tests.

Dependent Variable	"Like"	"Love"	"Angry"	No. of Comments	Total Word Count	Long Comments
P-value	0	0	0	0	0	0

about the mechanics of Facebook popularization and campaigning.

5.1. Automatic feature selection of the PLS method

Different from prior studies, which have identified independent variable manually, the method adopted in this study automates the selection of variables with associations. The PLS with cross-validation techniques used to examine 534 posts and 379,880 audience comments automatically selected dependent and independent variables that were highly correlated with each other.

We initially found that the dependent variables (*Like*, *Love*, *Angry*, *No. of comments*, *Total Word Count*, and *Long Comments*) were highly associated with the independent variables (*Lam, Tsang, Woo, Woo-Photo, Woo-Text*, and *Woo-Video*). We then conducted further analysis that included additional dependent and independent variables and examined if more associations can be found. These variables are shown in Table 8. Table 4 shows the six dependent variables that were highly associated with the independent variables. This means the six identified associations were the key relationships among the 17 dependent and 31 independent variables.

The automated variable selection process and the data analysis approach adopted in this study reduced the possibility of overlooking possible associations between variables and improved the reliability of the results. The method is thus a promising data analysis approach for investigating the mechanism of social media campaigns with many independent and dependent variables.

5.2. An online replication of power inequalities

Scholars have argued that online communication activities, such as using Facebook pages for publicity, can empower marginal political players, allowing them to promote their political ideas and engage Internet users, thus contributing to changes in the political and power landscape (Gibson and McAllister, 2015; Larsson, 2013; Souther, 2015; Vergeer et al., 2011). However, the results of this study did not suggest that social media, where flexible online campaigning strategies can be designed, are ideal spaces for shifting offline power structures. The findings of H1 and H3 supported the normalization hypothesis of an online replication of the inequalities present in the offline political structure. Although message format was found to relate to positive responses to posts by Facebook users, the identity of candidates played a crucial role in triggering affective and cognitive responses.

Lam's posts provoked the greatest number of "angry" emoji, comments, words in comments, and long comments, and this holds regardless of what message frames and message formats were used. This can plausibly be explained by the fact that Tsang and Lam were more famous candidates than Woo, with Tsang being a former financial secretary and Lam a former chief secretary. As supporters of the normalization hypothesis have suggested, the dominant political actors have advantages when constructing their images online through their posts. As a result of their offline positions, posts on the pages of Lam and Tsang attracted more attention and provoked greater controversy, especially in the case of Lam, who was considered to be the candidate favored by the central government (Lo et al., 2019). This could explain why Lam received more "angry" emoji, comments, words in the comments, and long comments than Tsang. However, Lam's posts did not receive a greater number of favorable reactions than Woo's, whereas Tsang's posts did. This reflects the offline reality of the time: the polls showed that Tsang was far more popular than Lam and the gap increased over the election period, thus Tsang's Facebook messages unsurprisingly received more positive reactions.

Table 7
Condition numbers and R^2 for the six regression models.

Normality assumption: The residuals follow a normal distribution	This assumption is fulfilled if the QQ plot of the theoretical normal quantile and the residual roughly form a diagonal line. The QQ plots of the six regression models are shown in Fig. 1. The blue dots plot the theoretical normal quantile and the residuals. The red lines are the diagonal lines. We can see that the blue dots are clustered around the red lines, and thus the normality assumptions were fulfilled. Moreover, outliers would be identifiable as blue dots positioned away from the red lines. Because there are no such dots, we can also conclude that the residuals are outlier free.
Linearity relationship: The mean value is a linear function of the independent variables	This assumption can be examined by checking the plots of the residuals versus the predicted values. If there is no obvious trend (upward, downward, or forming a curvilinear structure) shown in the plots, this assumption is fulfilled. The residual plots of the six regression models are shown in Fig. 2. We can see that there is no obvious trend, and thus, the assumption was fulfilled. Most residual plots in Fig. 2 have three segments because we treated the three candidates as the independent variables and in our regression model we used the variables separately. This caused the predicted values to have three segments.
Independence assumption: The independent variables are independent	This assumption can be examined by checking the plots of the residuals versus the predicted values. If there is no obvious trend (that is going upward, downward, or forming a curvilinear structure) shown in the plots, this assumption is fulfilled. The residual plots of the six regression models are shown in Fig. 2. We can see that there is no obvious trend and thus the assumption was fulfilled. Most residual plots in Fig. 2 have three segments because we treated the three candidates as the independent variables, and in our regression model, we used the variables separately. This caused the predicted values to have three segments. Moreover, condition numbers can also be used to check if the independent variables are independent. If the condition number is less than or equal to 25, the variables are independent. This is also shown in Table 7.
Goodness of fit of the model: The model fits the data well	We use the R^2 to provide an indication of the goodness of fit. A value of over 0.7 is considered a good fit. The R^2 of the six regressions are shown in Table 7. It can be seen that nearly all R^2 are close to 1, which means the models were an excellent fit to the data.

Our findings have provided evidence for the online replication of the inequalities present in the offline political structure in Hong Kong, where Lam and Tsang, as dominant political candidates, enjoyed advantages in Facebook campaigning, not only in relation to the affective aspect but also in cognitive responses. This supports the idea that the online world is indeed a mirror of the socioeconomic and political structures existing in the offline world (Margolis et al., 1999; Klinger, 2013).

Table 8
Independent and dependent variables used in the extended analysis.

Independent Variable[1]	Dependent Variable[2]
(1) Lam; (2) Tsang; (3) Woo; (4) Content Frame-personal; (5) Content Frame-campaign; (6) Content Frame – policy/issue; (7) Content Frame – Off-topic; (8) Endorsement; (9) Infographics; (10) Mention of the opponent – no mention; (11) Mention of the opponent – praise; (12) Mention of the opponent – neutral; (13) Mention of the opponent – attack; (14) Number of hashtags; (15) Photo; (16) Text; (17) Video	(1) Likes; (2) Loves; (3) Hahas; (4) Wows; (5) Sads; (6) Angries; (7) Views; (8) Shares; (9) Comments; (10) Thankful; (11) No. of comments & replies; (12) Total word count; (13) Message over 10 words; (14) Message over 20 words; (15) Messages over 30 words; (16) No. of responses Lam being positive mentioned; (17) No. of responses Lam being negative mentioned; (18) No. of responses Tsang being positive mentioned; (19) No. of responses Tsang being negative mentioned; (20) No. of responses Woo being positive mentioned; (21) No. of responses Woo being negative mentioned; (22) No. of responses about the campaign themes; (23) No. of responses Lam positively named by a rival; (24) No. of responses Lam formally named by a rival; (25) No. of responses Lam negatively named by a rival; (26) No. of responses Tsang positively named by a rival; (27) No. of responses Tsang formally named by a rival; (28) No. of responses Tsang negatively named by a rival; (29) No. of responses Woo positively named by a rival; (30) No. of responses Woo formally named by a rival; (31) No. of responses Woo negatively named by a rival.

5.3. Minimal impact of online message strategies

This study filled a research gap in the literature on whether the identity of political players or their choice of campaigning strategies is more important in engaging Facebook users (Lev-On & Haleva-Amir, 2018). Through the testing of H2 and H4, it also contributed to the literature by examining the effects of both online (message frames and message formats) and offline (candidate identity) factors on Facebook users' affective and cognitive responses. The findings indicated that message frames had no impact on audience response, while message formats had an impact only on positive reactions but not on negative reactions or cognitive responses. This suggests that format is more important than content for garnering a positive response.

This study also provided supporting evidence on social media use for elections and political communication. The finding somewhat confirmed the effectiveness of videos and photos in attracting user responses, in this case favorable reactions ("Like" and "Love"), although video was not found to be better than photos in this respect. This is consistent with previous studies that found more vivid messages to be more effective. The use of videos and photos appears to be a more promising strategy for yielding positive affective responses and feedback from social media users. Video was not found to be a better medium for generating positive response than photos, although other factors such as the quality of the message may have altered these results. Text-only posts appear to be comparatively weak for attracting favorable affective reactions.

Although previous research has found that Facebook offered opportunities for politicians to promote their political ideas (Schweitzer, 2011; Southern, 2015), much social media campaigning places more emphasis on visual elements than on political issues (Magin et al., 2017). The findings of this study have justified that choice. Theoretically, it suggests that message formats are more significant factors than message frames. In practice, media professionals and politicians may consider using more video and visual images to engage social media users for more affective responses in political campaigns, although such considerations will not outweigh the effects of offline power structures.

The study further contributed to the literature on social media and the public sphere by examining the use of message frames and the variations in Facebook user responses. As messages published on social media can be made globally accessible and open for discussion, some scholars believe that these online platforms can serve as an online public sphere that can enhance freedom of speech and foster alternative discourses (Fuchs, 2012; Halpern et al., 2013; Loader et al., 2011; Van Dijick, 2012). Lee, So, and Leung (2015) argued that in the presence of political issues or elections, social media turns into an insurgent public sphere that allows people to connect, discuss issues, construct identities, articulate common goals, and engage in collective action. However, this study found that message formats and message frames do not have an effect on the quality of comments. This implies that no particular online strategy adopted by a politician will evoke more discussion; rather, the identity of the candidate determines how Internet users respond in comments. Politicians appear to have little capacity to change their relationship with Internet users and to evoke more discussion among them, even though politicians have the powered to be "primary definers" in relation to messages (Hall, Critcher, Jefferson, Clarke, & Roberts, 2013) in their social media campaigns. This may further suggest that social media are not ideal platforms for public discussion, given the difficulties in encouraging discussion on a topic regardless of a politician's choice of message frames and message formats. Last but not least, this study examined the case for Hong Kong and their people's particular context with high technology and their lack of political freedom. This conclusion could be different in other contexts like Europe and Latin America. It opens a new perspective on this kind of research.

5.4. Limitations of this study

This study has several limitations. First, it focuses on candidate posts and user responses related to the 2017 Hong Kong Chief Executive election, for which there were only three candidates. Hence, the results may not be generalizable to other elections with a large number of candidates. It may not be applicable to elections between parties. Future research should explore other types of political elections and determine whether the results of this study are also identified in online political campaigning around the world. Second, this study focuses on Facebook, but other social media platforms may have different limitations and features in relation to posts, reactions, and comments. Third, we cannot rule out the possibility that user responses may have been influenced by offline events during the election and reporting on other media platforms, such as traditional news media. Facebook may serve as a platform for audiences to express their emotional or cognitive responses to offline events. Future research could explore the relationships between offline events, online campaigns, and user responses.

Credit author statement

Dr. Lam Shu Yan Benson, Conceptualization, Methodology, Software, Data curation, Writing – original draft, Writing- Reviewing and Editing, Supervision Project administration. Dr. Cheung Mei Fung Meily, Conceptualization, Writing – original draft, Writing- Reviewing and Editing, Supervision and Project administration. Dr. Lo Wai Han, Conceptualization, Writing – original draft, Writing- Reviewing and Editing, and Project administration.

[1] Independent Variables:

Mention of the opponent coders were trained to classify candidates' posts into four categories: no mention; praise; neutral or attack.

[2] Dependent Variables:

No. of response about Lam/Tsang/Woo Positive/negative mentions are obtained using sentiment analysis. A comment is classified either positive sentiment or negative sentiment based on Naive Bayes classifier.

Acknowledgment

The work described in this paper was partially supported by the grant from Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. UGC/FDS14/P04/17).

References

Aalberg, T., Strömbäck, J., & de Vreese, C. (2012). The framing of politics as strategy and game: A review of concepts, operationalizations and key findings. *Journalism, 13*(2), 162–178.

Ballantine, P. W., Lin, Y., & Veer, E. (2015). The influence of user comments on perceptions of Facebook relationship status updates. *Computers in Human Behavior, 49*, 50–55.

Bennett, W. L. (1988). *News: The politics of illusion* (2nd ed.). New York: Longman.

Biocca, F. (1997). The cyborg's dilemma: Progressive embodiment in virtual embodiment in virtual environments. *Journal of Computer-Mediated Communication, 3*(2). https://doi.org/10.1111/j.1083-6101.1997.tb00070.x. Retrieved from http://www.ascusc.org/jcmc.

Brewer, P. R., Habegger, M., Harrington, R., Hoffman, L. H., Jones, P. E., & Lambe, J. L. (2016). Interactivity between candidates and citizens on a social networking site: Effects on perceptions and vote intentions. *Journal of Experimental Political Science, 3*(1), 84–96.

Bright, J., Hale, S. A., Ganesh, B., Bulovsky, A., Margetts, H., & Howard, P. (2017). Does campaigning on social media make a difference? Evidence from candidate use of twitter during the 2015 and 2017 UK elections. *Communication Research*, 1–22.

Broersma, M., & Graham, T. (2012). Social media as beat. Tweets a news source during the 2010 British and Dutch elections. *Journalism Practice, 6*(3), 403–419.

Chang, V. (2009). *Obama and the power of social media and technology*. Stanford graduate school of business. https://www.gsb.stanford.edu/faculty-research/case-studies/obama-power-social-media-technology.

Cheng, B. K. L., & Lo, W. H. (2015). The effects of melodramatic animation in crime-related news. *Journalism & Mass Communication Quarterly, 92*(3), 559–579.

Conway, B. A., Kenski, K., & Wang, D. (2015). The rise of twitter in the political campaign: Searching for intermedia agenda-setting effects in the presidential primary. *Journal of Computer-Mediated Communication, 20*(4), 363–380.

Craft, S., & Wanta, W. (2004). US public concerns in the aftermath of 9/11: A test of second-level agenda setting. *International Journal of Public Opinion Research, 16*(4), 456–463.

D'Angelo, P., Calderone, M., & Territola, A. (2005). Strategy and issue framing: An exploratory analysis of topics and frames in campaign 2004 print news. *Atlantic Journal of Communication, 13*(4), 199–219.

Dekavalla, M. (2016). Issue and game frames in the news: Frame-building factors in television coverage of the 2014 Scottish independence referendum. *Journalism, 19*(11), 1588–1607.

Dunaway, J., & Lawrence, R. G. (2015). What predicts the game frame? Media ownership, electoral context, and campaign news. *Political Communication, 32*(1), 43–60.

Enli, G. (2017). Twitter as arena for the authentic outsider: Exploring the social media campaigns of trump and clinton in the 2016 US presidential election. *European Journal of Communication, 32*(1), 50–61.

Fernandes, J., Giurcanu, M., Bowers, K. W., & Neely, J. C. (2010). The writing on the wall: A content analysis of college students' facebook groups for the 2008 presidential election. *Mass Communication & Society, 13*(5), 653–675.

Fernandes, J., & Shumow, M. (2016). Framing politics in transnational communities: Spanish-language immigrant media and election coverage in south Florida. *Journalism, 17*(8), 961–978.

Fuchs, C. (2012). The political economy of privacy on Facebook. *Television & New Media, 13*(2), 139–159.

Graber, D. A. (1993). *Mass media and American politics* (4th ed.). Washington, DC: CQ Press.

Grusell, M., & Nord, L. (2016). Setting the trend or changing the game? Professionalization and digitalization of election campaigns in Sweden. *Journal of Political Marketing*. https://doi.org/10.1080/15377857.2016.1228555 [online].

Hagar, D. (2014). Campaigning online: Social media in the 2010 niagara municipal elections. *Canadian Journal of Urban Research, 23*(1), 74–98.

Hall, S., Critcher, C., Jefferson, T., Clarke, J., & Roberts, B. (2013). *Policing the crisis: Mugging, the state and law and order*. Macmillan International Higher Education.

Halpern, D., & Gibbs, J. (2013). Social media as a catalyst for online deliberation? Exploring the affordances of facebook and YouTube for political expression. *Computers in Human Behavior, 29*(3), 1159–1168.

Hong, S., & Nadler, D. (2012). Which candidates do the public discuss online in an election campaign?: The use of social media by 2012 presidential candidates and its impact on candidate salience. *Government Information Quarterly, 29*(4), 455–461.

Iyengar, S. (1991). *Is anyone responsible? How television frames political issues*. Chicago: University of Chicago Press.

Jamieson, K. H. (1992). *Dirty politics*. New York: Oxford University Press.

John, L., Mochon, D., Emrich, O., & Schwartz, J. (2017). *What's the value of a like? Harvard business review*. March-April 2017 Issue. Retrieved from https://hbr.org/2017/03/whats-the-value-of-a-like.

Junger, A. (2004). *Twitter in politics: A comprehensive literature review'*. Social Science Research Network. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2402443. (Accessed 28 April 2020).

Kan, K. (2012). Beijing's invisible hand. Power struggles and media meddles in the Hong Kong Chief Executive Election. *China Perspectives, 2*, 81, 41.

Kim, C., & Yang, S. U. (2017). Like, comment, and share on Facebook: How each behavior differs from the other. *Public Relations Review, 43*(2), 441–449.

Kite, J., Foley, B. C., Grunseit, A. C., & Freeman, B. (2016). Please like me: Facebook and public health communication. *PloS One, 11*(9), Article e0162765. https://doi.org/10.1371/journal.pone.0162765

Klinger, U. (2013). Mastering the art of social media: Swiss parties, the 2011 national election and digital challenges. *Information, Communication & Society, 16*(5), 717–736. https://doi.org/10.1080/1369118X.2013.782329

Ksiazek, T. B., Peer, L., & Lessard, K. (2016). User engagement with online news: Conceptualizing interactivity and exploring the relationship between online news videos and user comments. *New Media & Society, 18*(3), 502–520.

Larsson, A. O. (2016). Online, all the time? A quantitative assessment of the permanent campaign on facebook. *New Media & Society, 18*(2), 274–292. https://doi.org/10.1177/1461444814538798

Larsson, A. O., & Moe, H. (2014). Triumph of the underdogs? Comparing twitter use by political actors during two Norwegian election campaigns. *SAGE Open, 4*(4), 2158–2440.

Lawrence, R. G. (2000). *The politics of force: Media and the construction of police brutality*. London: University of California Press.

Lee, E. J., & Oh, S. Y. (2012). To personalize or depersonalize? When and how politicians' personalized tweets affect the public's reactions. *Journal of Communication, 62*(6), 932–949.

Lee, P. S., So, C. Y., & Leung, L. (2015). Social media and Umbrella Movement: Insurgent public sphere in formation. *Chinese Journal of Communication, 8*(4), 356–375.

Lev-On, A., & Haleva-Amir, S. (2018). Normalizing or equalizing? Characterizing facebook campaigning. *New Media & Society, 20*(2), 720–739.

Lilleker, D. G., & Vedel, T. (2013). The Internet in campaigns and elections. In W. H. Dutton (Ed.), *Oxford handbook of Internet studies* (pp. 401–420). Oxford: Oxford University Press.

Lin, H. C. (2017). How political candidates' use of Facebook relates to the election outcomes. *International Journal of Market Research, 59*(1), 77–96.

Lo, S.-H. (2017). Factionalism and Chinese-style democracy: The 2017 Hong Kong chief executive election. *Asia Pacific Journal of Public Administration, 39*(2), 100–119. https://doi.org/10.1080/23276665.2017.1325620

Loader, B. D., & Mercea, D. (2011). Networking democracy? Social media innovations and participatory politics. *Information, Communication & Society, 14*(6), 757–769.

Lo, W. H., & Cheng, B. K. L. (2015). *The use of melodramatic animation in news, presence and news credibility: A path model*. Journalism Studies. http://www.tandfonline.com/doi/abs/10.1080/1461670X.2015.1087814?journalCode=rj os20.

Lo, W. H., & Cheng, K. L. B.((2020). Does virtual reality attract visitors? The mediating effect of presence on consumer response in virtual reality tourism advertising. *Information Technology & Tourism, 22*, 537–562. https://doi.org/10.1007/s40558-020-00190-2

Lo, W. H., Lam, S. Y., & Cheung, M. F. (2019). *The dynamics of political elections: A big data analysis of intermedia framing between social media and news media. Who has a say in political elections? Framing in the era of big data. Social science computer review*. Online First. https://doi.org/10.1177/0894439319876593

Magin, M., Podschuweit, N., Haßler, J., & Russmann, U. (2017). Campaigning in the fourth age of political communication. A multi-method study on the use of Facebook by German and Austrian parties in the 2013 national election campaigns. *Information, Communication & Society, 20*(11), 1698–1719.

Marozzo, F., & Bessi, A. (2018). Analyzing polarization of social media users and news sites during political campaigns. *Social Network Analysis and Mining, 8*(1). https://doi.org/10.1007/s13278-017-0479-5

McGregor, S. C., Lawrence, R. G., & Cardona, A. (2017). Personalization, gender, and social media: Gubernatorial candidates' social media strategies. *Information, Communication & Society, 20*(2), 264–283.

McNair, B. (2003). *An introduction to political communication*. London: Routledge.

Medina, R. Z., Sánchez Cobarro, P. D. H., & Martínez, H. M. (2017). The importance of the" strategic game" to frame the political discourse in Twitter during 2015 Spanish Regional Elections. *Communications Society, 30*(3).

Meraz, S. (2011). Using time series analysis to measure intermedia agenda-setting influence in traditional media and political blog networks. *Journalism & Mass Communication Quarterly, 88*(1), 176–194.

Morgan, J. S., Shafiq, M. Z., & Lampe, C. (2013). Is news sharing on Twitter ideologically biased?. In *Proceedings of the ACM conference on computer supported cooperative work, CSCW* (pp. 887–896). https://doi.org/10.1145/2441776.2441877

Neuman, W. R., Guggenheim, L., Jang, S. M., & Bae, S. Y. (2014). The dynamics of public attention: Agenda-setting Theory meets Big data. *Journal of Communication, 64*, 193–214.

Nielsen, R. K., & Vaccari, C. (2013). Do people "like" politicians on facebook? Not really. Largescale direct candidate-to-voter online communication as an outlier phenomenon. *International Journal of Communication, 7*, 2333–2356.

O'Sullivan, P. B., & Carr, C. T. (2018). Masspersonal communication: A model bridging the mass-interpersonal divide. *New Media & Society, 20*(3), 1161–1180.

Patterson, T. E. (1993). *Out of order*. New York: Knopf.

Pedersen, R. (2012). The game frame and political efficacy: Beyond the spiral of cynicism. *European Journal of Communication, 27*(3), 225–240. https://doi.org/10.1177/0267323112454089

Quinlan, S., Gummer, T., Roßmann, J., & Wolf, C. (2017). *'Show me the money and the party!' – variation in Facebook and Twitter adoption by politicians*. Information, Communication & Society. https://doi.org/10.1080/1369118X.2017.1301521

Rauchfleisch, A., & Metag, J. (2020). Beyond normalization and equalization on Twitter: Politicians' Twitter use during non-election times and influences of media attention. *Journal of Applied Journalism & Media Studies, 9*(2), 169–189.

Reeves, B., & Nass, C. (1996). *The media equation: How people teat computer, television and new media like real people and places*. Standford, CA: CSLI Publication and Cambridge University Press.

Rhee, J. W. (1997). Strategy and issue frames in election campaign coverage: A social cognitive account of framing effects. *Journal of Communication, 47*(3), 26–48.

Ross, K., Fountaine, S., & Comrie, M. (2015). Facing up to facebook: Politicians, publics and the social media (ted) turn in New Zealand. Media. *Culture & Society, 37*(2), 251–269. https://doi.org/10.1177/0163443714557983

Rus, H. M., & Cameron, L. D. (2016). Health communication in social media: Message features predicting user engagement on diabetes-related facebook pages. *Annals of Behavioral Medicine, 50*(5), 678–689.

Shehata, A. (2014). Game frames, issue frames, and mobilization: Disentangling the effects of frame exposure and motivated news attention on political cynicism and engagement. *International Journal of Public Opinion Research, 26*(2), 157–177.

Skovsgaard, M., & Dalen, A. V. (2013). Dodging the gatekeepers? Social media in the campaign mix during the 2011 Danish elections. *Information, Communication & Society, 16*(5), 737–756.

Smock, A. D., Ellison, N. B., Lampe, C., & Wohn, D. Y. (2011). Facebook as a toolkit: A uses and gratification approach to unbundling feature use. *Computers in Human Behavior, 27*(6), 2322–2329.

Song, Y., Dai, X. Y., & Wang, J. (2016). Not all emotions are created equal: Expressive behavior of the networked public on China's social media site. *Computers in Human Behavior, 60*, 525–533.

Srivastava, J., Saks, J., Weed, A. J., & Atkins, A. (2018). Engaging audiences on social media: Identifying relationships between message factors and user engagement on the American Cancer Society's Facebook page. *Telematics and Informatics, 35*, 1832–1844.

Statista. (2019). Number of daily active Facebook users worldwide as of 3rd quarter 2018. Statista – the Statistics Portal. https://www.statista.com/statistics/346167/facebook-global-dau/.

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication, 42*, 73–93.

Street, R. L., Jr, & Manning, T. (1997). Information environments for breast cancer education. In R. L. StreetJr, W. R. Gold, & T. Manning (Eds.), Health promotion and interactive technology: Theoretical applications and future directions (pp. 121-139). Mahwah, NJ: Lawrence Erlbaum Associates.

Su, M. H., Liu, J. W., & McLeod, D. M. (2019). Pathways to news sharing: Issue frame perceptions and the likelihood of sharing. *Computers in Human Behavior, 91*, 201–210.

Van Dijck, J. (2012). Facebook as a tool for producing sociality and connectivity. *Television & New Media, 13*(2), 160–176.

Vicario, M. D., Gaito, S., Quattrociocchi, W., Zignani, M., & Zollo, F. (2017). *News consumption during the Italian referendum: A cross-platform analysis on facebook and twitter*. IEEE International Conference on Data Science and Advanced Analytics (DSAA).

Waddell, T. F., & Sundar, S. S. (2017). # thisshowsucks! the overpowering influence of negative social media comments on television viewers. *Journal of Broadcasting & Electronic Media, 61*(2), 393–409.

Wallsten, K. (2010). 'Yes we can': How online viewership, blog discussion, campaign statements, and mainstream media coverage produced a viral video phenomenon. *Journal of Information Technology & Politics, 7*(2–3), 163–181. https://doi.org/10.1080/19331681003749030

Walther, J. B., Carr, C. T., & Choi, S. S. W. (2010). Interaction of interpersonal, peer, and media influence sources online: A research agenda for technology convergence. In *A networked self* (pp. 25–46). Routledge.

Woolley, J. K., Limperos, A. M., & Oliver, M. B. (2010). The 2008 presidential election, 2.0: A content analysis of user-generated political facebook groups. *Mass Communication & Society, 13*(5), 631–652.

Xenos, M. A., Macafee, T., & Pole, A. (2017). Understanding variations in user response to social media campaigns: A study of facebook posts in the 2010 US elections. *New Media & Society, 19*(6), 826–842.

Zhang, L., Peng, T. Q., Zhang, Y. P., Wang, X. H., & Zhu, J. J. (2014). Content or context: Which matters more in information processing on microblogging sites. *Computers in Human Behavior, 31*, 242–249.

Zhou, Y., & Moy, P. (2007). Parsing framing processes: The interplay between online public opinion and media coverage. *Journal of Communication, 57*, 79–98.