Exploring user responses to entertainment and political videos: An automated content analysis of YouTube

Möller, A.M.; Kühne, R.; Baumgartner, S.E.; Peter, J.

DOI
10.1177/0894439318779336

Publication date
2019

Document Version
Final published version

Published in
Social Science Computer Review

License
CC BY-NC

Article

Social Science Computer Review
2019, Vol. 37(4) 510-528

DOI: 10.1177/0894439318779336
journals.sagepub.com/home/ssc
SAGE

Exploring User Responses to Entertainment and Political Videos: An Automated Content Analysis of YouTube

A. Marthe Möller[1], Rinaldo Kühne[1], Susanne E. Baumgartner[1], and Jochen Peter[1]

Abstract
On YouTube, videos are always presented together with additional user-generated information about those videos. This *social information* is presented in the form of number of views, (dis)likes, or comments. However, we know little about the characteristics of social information about entertainment videos. To fill this gap, the present study examined the amount and valence of online entertainment videos' social information and compared this to the social information of online political videos. An automated content analysis of (dis)likes, views, and 39,602 comments presented alongside 463 videos showed that entertainment videos received more views and comments than political videos. Moreover, entertainment videos' comments were more neutral than political videos' comments. We also found that comments with a stronger positive or negative valence received more replies and likes, with the exceptions that the positive valence of political videos had no effect and that, for political videos, a stronger negative valence led to fewer likes. Finally, we found that as political videos received more comments, the positive valence of their comments became more consistent. Overall, these results show that the type of video influences the amount and valence of social information the video receives.

Keywords
social information, social influence, entertainment videos, political videos, automated content analysis, YouTube

[1] Amsterdam School of Communication Research, University of Amsterdam, Amsterdam, the Netherlands

Corresponding Author:
A. Marthe Möller, Amsterdam School of Communication Research, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, the Netherlands.
Email: a.m.moller@uva.nl

Social media have not only changed how people communicate and interact with each other but also how media contents are presented and how recipients use these contents. A platform that has substantially changed the presentation and use of audiovisual content is the video-sharing website YouTube. A special feature that distinguishes videos on YouTube from traditional offline videos is that they typically are presented together with user-generated information about those videos. This information is presented in the form of views, (dis)likes, and comments. It forms an essential part of YouTube's content because it informs viewers about the evaluations that videos receive by their audience. For example, video (dis)likes inform viewers about how positive or negative the overall audience of the video evaluated it. Therefore, we define this information as *social information*. Despite the prevalence of online social information on YouTube, we still lack a sufficient understanding of its essential characteristics, such as its amount and valence (e.g., Thelwall, Sud, & Vis, 2011; Walther, DeAndrea, Kim, & Anthony, 2010).

It is important to understand the characteristics of social information on YouTube for at least two reasons. First, the social information created in response to YouTube videos is an important outcome in its own right because it can provide insights into how video viewers use the platform. Khan (2017) distinguishes two uses of YouTube, namely, active participation and passive consumption. Active participation describes how users actively engage with the platform by creating comments and (dis)likes to express their opinion about videos. The amount of social information created in response to videos is an indication of users' active engagement with YouTube. Second, because social information is a prevalent feature on social media platforms, it constitutes a crucial context factor that influences viewers' passive consumption of YouTube, that is, how they experience the platform's content (Cameron & Geidner, 2014; Khan, 2017; Waddell, & Sundar, 2017; Walther et al., 2010). Studying the characteristics of online social information thus contributes to a thorough understanding of how responses to online videos unfold.

To learn more about viewers' engagement with YouTube and their experiences of its content, this study systematically explores the social information that is presented alongside YouTube videos. As our knowledge of the social information of videos on YouTube is limited, the present study explores multiple aspects of social information, namely, user comments, video views, and (dis)likes. This study employs an exploratory approach by focusing specifically on the social information of *popular* YouTube videos, that is, videos that are among the most viewed videos on YouTube. The 10% most popular videos on YouTube create the majority of the platform's video views (Cha, Kwak, Rodriguez, & Moon, 2007). Thus, focusing on popular videos will increase our knowledge of the social information to which YouTube users typically are exposed.

As the content of YouTube is user generated, the platform contains a wide variety of videos. This study will explore how the characteristics of user comments, views, and (dis)likes vary depending on video genre. A genre is a content category that distinguishes itself through characteristic features that match individuals' motivations for using media contents (Potter, 2009; Prior, 2005). For example, entertainment media content is primarily sought out by individuals who are looking for enjoyment (Potter, 2009; Prior, 2005; Rubin, 1983). Thus, genres are helpful for content providers and audiences because they efficiently signal which gratifications are likely to be satisfied through usage (Potter, 2009; Prior, 2005; Rubin, 1983). As viewers' motivations for watching specific video genres may differ and different gratifications are fulfilled by specific video genres, the characteristics of social information are likely to depend on the genre of the video.

In the present study, we focus on two major types of media content that have been identified in research: entertainment media content which is aimed at fulfilling users' need for enjoyment and relaxation and political content that mainly fulfills users' information-seeking need (Potter, 2009; Prior, 2005; Rubin, 1983). Although entertainment forms the main motivation for people to use YouTube (Chau, 2010), existing research on online social information has primarily focused on user responses to political media content, whereas research on the social information of entertainment media content is scarce. Therefore, this study takes the literature from political communication as a starting point to compare the social information of entertainment and political videos.

Research on the Online Social Information of Political Media Content

Researchers studying online social information have often focused on political media content. These studies show that the characteristics of political media content influence the *amount* of social information that is created in response to it (Abdul-Mageed, 2008; Ksiazek, 2018; Liu, Zhou, & Zhao, 2015; Tatar et al., 2011; Weber, 2014). Scholars found that the topics of online news items and the terminology that is used to describe those topics influence the number of comments they receive (Abdul-Mageed, 2008; Ksiazek, 2018; Tatar et al., 2011; Tsagkias, Weerkamp, & De Rijke, 2009). For example, researchers found that discussing topics surrounded by controversy such as gun control increased the number of comments that articles received (Ksiazek, 2018; Liu et al., 2015).

Not only does the amount of online social information depend on the characteristics of political media content, studies also suggest that social information's *valence* (i.e., the positivity and negativity of comments) can depend on them. Multiple content analyses investigated negative valence in the form of hostility in comments posted in response to online news articles (e.g., Coe, Kenski, & Rains, 2014; Ksiazek, 2018). In one of these studies, Coe, Kenski, and Rains (2014) found that online articles discussing news related to law and order, taxes, or the economy received more uncivil comments than articles discussing lifestyle or health. In addition, Ksiazek (2018) identified gun control, defense, and foreign policy as topics that increase the hostility in comments, in comparison to the topic of labor. The literature discussed here thus shows that the characteristics of media content can influence the amount and the valence of social information that it receives. However, when users create social information, they do not only do so in response to the media content but also in response to social information created by previous users. The characteristics of social information are therefore likely to depend not only on the media content but also on the characteristics of previously created social information.

Several scholars studied how the amount of comments written in response to online news articles depends on the characteristics of previously written comments. In an experiment, Ziegele, Weber, Quiring, and Breiner (2017) investigated how previously written user comments influence users' willingness to post new comments. They found that comments including questions or incivility increased users' tendency to reply to these comments. Yun, Park, Holody, Yoon, and Xie (2013) found that in comment threads discussing articles about abortion, comments expressing an opinion incongruent to the opinion expressed in the article received more replies than comments expressing an opinion congruent to the article. Moreover, Song, Dai, and Wang (2016) conducted a content analysis of comments discussing a political topic on a Chinese social network site. They found that messages that are very positive or very negative are shared by others more often than neutral messages, with messages expressing anger being shared most often (Song, Dai, & Wang, 2016). In sum, these studies show that the amount of user comments written is influenced by previously written comments.

Previously written user comments also affect the *valence* of political media content's social information. Ziegele et al. (2017) found that users are less likely to write uncivil comments if previous comments include questions or information supplementary to the media content, while users are more likely to post uncivil comments if previous comments included incivility (Ziegele, Weber, Quiring, & Breiner, 2017). Chmiel et al. (2011) found that replies on political websites often contain the same valence as the comments in response to which they were written. That is, while negative comments received mostly negative replies, positive comments received mostly positive replies (Chmiel et al., 2011). Finally, a study by Siersdorfer, Chelaru, Nejdl, and San Pedro (2010) researching a large sample of YouTube videos discussing various topics implies that this pattern also occurs across other components of social information. Siersdorfer et al. (2010) found that as comments are more positive, users are more inclined to assign a like to them.

In sum, the literature on online political media content showed that the amount and valence of political media content's social information depends on content characteristics and previously created social information. However, with the exception of Siersdorfer et al. (2010), the studies discussed here focused exclusively on political media content, and it is unknown whether these findings also apply to entertainment media content. Because entertainment media content is typically geared toward creating enjoyment and political media content typically aims at informing audiences (Potter, 2009; Prior, 2005; Rubin, 1983), there may be differences between entertainment and political videos' social information. However, the literature described here did not systematically compare the social information of entertainment and political media content. It is therefore necessary to investigate which specific differences in the social information of entertainment and political videos may occur.

Video Type and Online Social Information

One of the differences between the social information of entertainment videos and that of political videos that is likely to occur is a difference in the number of video views that both types of videos receive. Video type is likely to affect the number of views for a rather simple reason: One of the most important motivations for people to watch YouTube videos is entertainment (Chau, 2010; Khan, 2017). Audiences are likely to seek out media content that adheres to their needs (Rubin, 2009), and as entertainment videos directly cater to a prevalent need of viewers, their number of views is likely to be high. It is therefore plausible that, overall, entertainment videos receive more video views than political videos. Based on this, we hypothesized:

Hypothesis 1: Online entertainment videos receive more views than online political videos.

Entertainment and political videos may also differ with regard to the number of comments they receive. Previous research indicates that political videos may elicit more intense discussions and more user comments because political content is often controversial (Chen & Berger, 2013; Jeong, 2003; Liu et al., 2015). Political videos often address issues that are highly relevant and have consequences for a considerable part of society, such as immigration, gun control, or tax policy. Moreover, these issues are often presented using frames that emphasize the conflict between political camps and viewpoints (Schuck et al., 2013; Semetko & Valkenburg, 2000). Because of this, the audiences of political videos are likely to be more inclined to express their opinion than the audiences of entertainment videos who are not exposed to political controversy. Hence, once they are viewed, political videos should elicit more comments than entertainment videos. In line with this reasoning, Choi et al. (2015) found that conversation threads on the online platform Reddit that discuss news include more comments than threads discussing other topics. Hence, we hypothesized:

Hypothesis 2: Online political videos receive more comments than online entertainment videos.

Next to the amount of views and comments, the valence of videos' social information may also differ. On YouTube, the valence of social information can be presented in different ways. While video likes express a positive valence, video dislikes express a negative valence. Similarly, comments can include positive expressions (e.g., "This is a fantastic video!") or negative expressions (e.g., "I hate this video!"). It should be noted that social information can simultaneously be positive and negative, for example, when a video receives both likes and dislikes. Similarly, a comment may include positive expressions as well as negative expressions (e.g., “Great story in this video, music in it was horrible though.”). Moreover, comments can vary in the intensity of their valence. For example, a comment with a high proportion of positive words (e.g., “This video is super amazing awesome and so much fun!”) is more positive than a comment with a low proportion of positive words (e.g., “This video is fun.”).

In the case of YouTube videos, video genre is likely to influence the valence of social information. Political videos are likely to elicit more negative responses than entertainment videos. As discussed before, political videos often discuss topics about which people disagree and present them in a manner that emphasizes these conflicts (Schuck et al., 2013; Semetko & Valkenburg, 2000). Expressions of disagreement are often associated with a negative valence (Hancock, Landrigan, & Silver, 2007). The negative valence of online political videos is likely to be adopted in the (dis)likes and the comments that they receive: Scholars found that people adopt the valence that they detect in messages and express it in their own messages (Hancock, Gee, Ciaccio, & Lin, 2008; Kramer, Guillory, & Hancock, 2014). Moreover, comments written by individuals who engage in an online discussion of controversial issues are likely to include disagreement, which can result in a negative valence (Hancock et al., 2007).

In contrast, entertainment videos are likely to elicit relatively positive responses. While the content of political videos is likely to address topics about which people disagree, entertainment videos primarily aim to create positive experiences by entertaining viewers. The content of entertainment videos is therefore likely to be more positive than that of political videos. As viewers are likely to adopt the valence expressed in entertainment videos in their own comments (Hancock et al., 2008; Kramer et al., 2014), the social information of entertainment videos is likely to be more positive than that of political videos. We thus hypothesized:

Hypothesis 3a: Online entertainment videos receive more likes than online political videos.

Hypothesis 3b: Online political videos receive more dislikes than online entertainment videos.

Hypothesis 3c: Comments written in response to online entertainment videos are more positive than comments written in response to online political videos.

Hypothesis 3d: Comments written in response to online political videos are more negative than comments written in response to online entertainment videos.

The Role of Previously Created Social Information

In addition to the type of video, previously created social information may influence the amount and valence of new social information (Chmiel et al., 2011; Siersdorfer, Chelaru, Nejdl, & San Pedro, 2010; Ziegele et al., 2017). With regard to the *amount* of social information, previous studies found that comments that are very positive or very negative trigger more responses than neutral messages (Song et al., 2016; Ziegele et al., 2017). This can be explained by the notion that comments that are very positive or very negative elicit more arousal than neutral comments, making recipients more inclined to respond (Berger & Milkman, 2012). Thus, the amount of replies that comments receive is likely to depend on their valence. Specifically, comments with a strong positive or negative valence are likely to receive more replies than neutral comments. With regard to the *valence* of social information, scholars found that people’s responses to comments often have the same valence as that of the comment to which they respond (Chmiel et al., 2011; Siersdorfer et al., 2010; Ziegele et al., 2017). This is in line with the general finding that people tend to adopt the valence of messages that they read and express it as their own messages (Hancock et al., 2008; Kramer et al., 2014).

Based on this, we propose that the more positive a comment is, the more likes it receives, while the more negative a comment is, the fewer likes it receives. Hence, we hypothesized:

Hypothesis 4: The more positive the comment, (a) the more replies and (b) the more likes it receives.

Hypothesis 5: The more negative the comment, (a) the more replies but (b) the fewer likes it receives.

However, previously posted comments might not only influence replies and likes to these comments, but the valence of the entirety of previous comments might also influence the valence of subsequent comments. To explain this, the referent informational influence framework (Turner, 1982) is useful. The framework describes social influences within groups. A group is formed by two or more people if they are aware of a common social category membership (Turner, 1982). On social media, a group can be formed when individuals collectively participate in creating social information by expressing their opinion about media content (Postmes, Spears, & Lea, 2000). The referent informational influence theory states that when individuals consider themselves members of a social group, they conform to behavioral norms that characterize that group (Turner, 1982). In line with this reasoning, Postmes, Spears, and Lea (2000) found that, in online conversations, group members increasingly conformed to communication patterns of the group: As time passed, the communication style of group members became increasingly similar because members increasingly used humor or incivility in their messages.

On YouTube, user comments provide information about how people communicate with each other. As the number of comments increases, characteristics of a group's communication style (e.g., the valence of messages) become more evident, and authors of new comments are more likely to conform to these standards. Thus, when new comments are posted to a video, either the positively or the negatively toned expressions are likely to become more salient because the authors of the new comments conform to the group norm. Through this self-reinforcing process, the homogeneity of the comments' valence should increase over time. Or put differently, the variance in the positive and negative valence of user comments should decrease when the number of comments increases. Based on this reasoning, we hypothesized that:

Hypothesis 6: As the number of comments written in response to a video increases, the comments' valence becomes more consistently positive or negative.

Hypotheses 4–6 describe how previously created social information affects the amount and valence of subsequently created social information in response to entertainment and political videos. Presumably, the experiences that viewers have when watching entertainment videos may differ from their experiences when watching political videos. As discussed before, the audiences of political videos are likely to be more actively engaged in the content than entertainment videos viewers because they are exposed to political issues presented in a manner that emphasizes conflict and controversy. Because of this, political video viewers may also be more actively involved in examining and creating social information than entertainment video viewers. As a consequence, the influence of previously created social information on new social information may differ depending on the video type. However, based on our current knowledge, it is unclear how these differences emerge and how they are reflected in the social information of entertainment and political videos. To learn more about how the effects of previously created social information on the number of likes and replies that comments receive (Hypotheses 4 and 5) and on the consistency of the comments' valence (Hypothesis 6) may vary between entertainment and political videos, we proposed three research questions:

Research Question 1: How do the relationships between the positive valence of user comments and (a) the number of replies that they receive and (b) the number of likes that they receive differ between online entertainment videos and online political videos?

Research Question 2: How do the relationships between the negative valence of user comments and (a) the number of replies that they receive and (b) the number of likes that they receive differ between online entertainment videos and online political videos?

Research Question 3: How does the relationship between the number of comments written in response to videos and the consistency of the comments' valence differ between online entertainment videos and online political videos?

Method

Sample Selection

To draw a conclusion about the proposed hypotheses and research questions, the social information of frequently viewed entertainment videos and political videos on YouTube was collected. Institutional Review Board approval for this study was granted by the ethical committee of the authors' university (ERB number 2018-YME-9226). To determine the initial sample of videos, we used the website Social Blade (https://www.socialblade.com), which presents statistics about several social media platforms, including YouTube. The statistics about YouTube provided by Social Blade were used for two reasons. First, the statistics presented a starting point to identify videos that were likely to be entertainment or political videos. Social Blade categorizes YouTube video channels according to their topic. This categorization is based on the topic of the channels' most recent videos as determined by the uploaders of the videos. We selected potential entertainment videos from channels with videos categorized in 11 topics, namely, entertainment, comedy, film, gaming, shows, how to and style, music, people and blogs, pets and animals, sports, or travel. In addition, potential political videos were selected from channels with videos on news and politics.

Second, the statistics presented by Social Blade ensured that the sample included only popular videos. For entertainment videos, for each of the 11 entertainment topics, we selected the 50 YouTube channels with the highest total number of viewers. From each of these channels, the most frequently watched video was selected. When, in five cases, data were missing on how often each of a channel's videos had been watched, we selected the most recent video. In the two cases where information about the channel's most recent videos was also missing, we selected the highest rated video. This resulted in 550 potential entertainment videos.

As political videos were only selected from channels on news and politics, selecting only the 50 most popular channels would result in few political videos. Thus, to increase the number of political videos in the sample while at the same time ensuring that only the most popular political videos would be included in the sample, the 150 channels with the most viewers discussing news and politics were selected. From each of the selected channels, the most frequently watched video was chosen. In one case, data were missing on the number of times that videos were watched. We selected the most recent video instead. In total, 150 potential political videos were selected.

Data Collection

After identifying potential entertainment and political videos, the social information of the videos was collected between January and April 2017. This was done through a Python 3.5 script that employs the Google Application Programming Interface (API) and is openly accessible (Van de Velde, 2017). Through the Google API, the data of 537 entertainment videos were found as well as the data of 147 political videos. Manual inspection of the 16 videos with missing social information showed that three videos were not available on YouTube and two videos were not found because the video links retrieved from Social Blade were invalid. The remaining 11 videos were available on YouTube, but for unknown reasons they were not found by the Google API. These videos were searched for again and this time, the social information of eight more videos was found through the API. Hence, the social information of 544 entertainment videos and 148 political videos was collected.

The information that was found through the Google API was downloaded in two steps. First, information and summary statistics about each video were collected. This includes the videos' title, the video description, the number of likes and dislikes that the video received, the number of times that it was viewed, and the number of comments that it received. Second, the comments of each video were collected. This was done for the first five pages of comments posted in response to each video, which are the approximately 147 newest comments of each video. For each comment, we extracted data including information indicating to which video the comment was posted, the text of the comment, the number of likes that the comment received, and the number of replies that it received. In total, the information of 101,889 comments written in response to 692 videos was collected.

Sample Validation

At the outset of the study, we planned to work with the categorization of entertainment and political videos as provided by Social Blade. However, after manual inspection of some of the videos and doubts about the validity of the Social Blade categorization, we decided to hand code whether the videos were entertainment or political videos. At the time of hand coding, however, 43 videos were no longer available, resulting in 649 videos that could be hand coded.

Entertainment videos were defined as videos discussing topics that can offer entertainment either while watching the videos or as a direct consequence of watching them. Videos can do so by (a) offering amusement, for example, through comedy or games (Gray, 2009), (b) offering escapism, for example, through a (fictional) narrative (Vorderer, Klimmt, & Ritterfeld, 2004), or (c) connecting to viewers' personal interests, for example by presenting sports or music (Vorderer, 2001). Political videos were defined as videos that discuss topics which are relevant for, and have a direct or indirect influence on, a considerable number of individuals within a society and therefore create a need for discussion or action (Potter, 2009; Prior, 2005; Rubin, 1983; Young & Leonardi, 2012).

As the videos were coded based on their topic, it was important that the content of the video was clearly understandable. Therefore, coders were instructed to only code videos in English or with English subtitles. In the sample, 185 videos were neither spoken in English nor did they have English subtitles. Thus, in total, 464 videos were hand coded to categorize them as either entertainment or political. A preliminary look at a selection of the videos showed that the first 5 min of a video provides sufficient information to understand its content. Therefore, coders were instructed to watch half of each video plus one additional minute before they coded it, except when a video was shorter than 5 min, in which case coders were instructed to watch the entire video (n = 284, 61.2%). One coder coded all the videos and a second coder coded a randomly selected 22% of the videos (n = 102) to determine intercoder reliability. The coders agreed in 92% of the cases, indicating a good intercoder reliability. Subsequent discussion about the remaining 8% led to the agreement to include the coding results of the first coder in the analyses. Based on the results of the coding procedure, 416 videos of which the social information was collected were categorized as entertainment videos and 47 videos were categorized as political videos. One video was coded as neither entertainment nor political. This video was excluded from the analyses.

Data Processing

After the final sample for the study was determined, the collected data were processed using a second Python 3.5 script. With this script, the language of each comment was determined using Google's language detection library (Google, 2017). Because the algorithm used in the subsequent part of the script can only deal with texts written in English, comments that were not written in English were excluded from the analyses. The remaining 39,602 English comments were further analyzed to test the hypotheses of this study.

The valence of each comment was assessed through the SentiStrength algorithm (Thelwall, Buckley, & Paltoglou, 2012). This algorithm is designed to detect the tone of voice of short web texts and has been tested and validated for online forum posts, Tweets, and comments written on YouTube (Thelwall et al., 2012). After applying the SentiStrength algorithm, each comment received a score indicating the strength of the comment's positive valence, ranging from 1 (*neutral*) to 5 (*extremely positive*). In addition, comments received a score indicating the strength of their negative valence ranging from -1 (*neutral*) to -5 (*extremely negative*).

For the data on the video level, the average positive valence of all comments written in response to a video was calculated by taking the mean of the positive valence of all comments of this video. In a similar manner, a variable indicating the mean negative valence of all comments written in response to a video was created. To assess the variability of positive valence, the variance in the positive valence of all the comments written to each video was calculated. Finally, a similar variable was created for the variance of the negative valence of all the comments written to each video (see Table 1). As the variance of comments' negative and positive valence could only be calculated for videos with at least two or more collected comments, 48 videos with less than two comments were not included in the analyses of the variability in the comments' valence.

Data Analysis

The processed data were analyzed using the Statistical Package for the Social Science (SPSS) 24. Before the hypotheses were tested, it was checked whether the data were normally distributed. Scores on variables that had a distribution with a skew and kurtosis between -3 and 3 were regarded as normally distributed. A descriptive analysis showed that for variables on the video level, the scores on video view count, video comment count, video like count, and video dislike count were skewed (see Table 1). Another descriptive analysis showed that for variables on the comment level, the scores on the comments' like count and reply count were skewed (see Table 1). Therefore, a log transformation was performed on these variables using the following formula: *logged variable* = LN(*original variable* + 1). Inspection of the new, logged variables showed that the scores on the logged variables of video view count, video comment count, video like count, and video dislike count were now normally distributed. Analyses including these variables were therefore run with the new, transformed variables. Scores on the comments' like count and reply count, however, remained skewed (see Table 1).

To test Hypotheses 1 through 3d, scores of entertainment videos and political videos on several variables were compared. The comparisons were conducted by running a series of analyses of variance (ANOVAs) with the following dependent variables: video view count (Hypothesis 1), video comment count (Hypothesis 2), video like count (Hypothesis 3a), video dislike count (Hypothesis 3b), video comments' mean positive valence (Hypothesis 3c), and video comments' mean negative valence (Hypothesis 3d). When testing Hypotheses 2 through 3d, we included the video view count as a control variable in the model to preclude that differences in the dependent variables were caused by different numbers of video views for entertainment and political videos. To test Hypotheses 4a through 5b and to answer Research Questions 1 and 2, analyses were performed on

Table 1. Descriptive Statistics of Dependent Variables.

Variable	Min	Max	*M*	*SD*	Skewness	Kurtosis
Video level						
Mean comment positive valence	1.00	2.47	1.61	0.27	0.20	0.54
Mean comment negative valence	−2.70	−0.92	−1.39	0.32	−1.27	1.86
Variance comment positive valence	0.00	1.67	0.66	0.28	−0.01	0.56
Variance comment negative valence	0.00	2.15	0.62	0.44	0.63	−0.10
Video view count (untransformed)	127,133	3,000,000,000	151,185,432	355,980,233	3.93	17.45
Video view count (logged)	11.75	21.73	17.32	1.76	0.01	0.37
Video comment count (untransformed)	0.00	4,780,584	63,136.13	253,820.77	14.91	269.87
Video comment count (logged)	0.00	15.38	9.32	1.99	−0.41	1.07
Video like count (untransformed)	140	13,000,632	530,953.49	1391,092.87	5.22	34.30
Video like count (logged)	4.95	16.38	11.46	1.96	−0.15	0.23
Video dislike count (untransformed)	5.00	1,680,907	46,565.36	133,741.49	7.31	70.56
Video dislike count (logged)	1.79	14.33	9.06	2.00	−0.45	0.87
Comment level						
Comment like count (untransformed)	0.00	2,022.00	1.40	21.84	54.91	3,832.25
Comment like count (logged)	0.00	7.61	0.25	0.61	3.69	18.80
Comment reply count (untransformed)	0.00	193.00	0.29	2.46	35.86	2,023.13
Comment reply count (logged)	0.00	5.27	0.10	0.37	4.94	30.32

the comment level. The analyses tested the relationship between comments' positive valence and the number of replies (Hypothesis 4a) and likes (Hypothesis 4b) they receive as well as the relationship between comments' negative valence and the number of replies (Hypothesis 5a) and likes (Hypothesis 5b) they receive. As indicated earlier, the scores on the comment reply count and the comment like count were skewed. Thus, we employed regression models for count data. More specifically, we ran negative binomial regression models on the original, untransformed variables because the assumption of equidispersion was violated (Coxe, West, & Aiken, 2009). Regression coefficients indicate how a one-unit increase in comment valence multiplicatively affects the respective count outcome. For the measure of positive valence, coefficients above 1.0 indicate that a more positive valence increases the expected count. For the measure of negative valence, coefficients below 1.0 indicate that a more negative valence increases the expected count. When Hypotheses 4a and 5a were tested, only data of comments that were not replies were used because on YouTube, comments that are replies cannot receive replies. The analyses were repeated for comments written in response to entertainment videos and political videos separately to answer Research Questions 1 and 2. Finally, Hypothesis 6 was tested by running linear regression analyses on the video comment count and the variance in comments' positive valence and negative valence. These analyses were repeated for entertainment videos and political videos separately to answer Research Question 3.

Results

Hypothesis 1 stated that entertainment videos receive more views than political videos. In support of this hypothesis, we found that there was a difference between entertainment videos and political videos with

Table 2. Mean scores of Entertainment and Political Videos on Video Views, Video Comments, Likes, Dislikes, and Comment Valence (Hypotheses 1–3).

	EntertainmentVideos		PoliticalVideos				
	M	*SD*	*M*	*SD*	Part. η^2	*p*	Hypotheses
Video views count (logged)	17.49	1.73	15.84	1.31	.08	<.001	H1 supported
Video comment count (logged)	9.38	1.99	8.77	1.91	.02	.001	H2 not supported
Video like count (logged)	11.64	1.91	9.91	1.75	.00	*n.s.*	H3a not supported
Video dislike count (logged)	9.20	2.00	7.83	1.65	.02	.009	H3b not supported
Mean comment positive valence	1.59	0.27	1.74	0.23	.02	.002	H3c not supported
Mean comment negative valence	−1.33	0.24	−1.98	0.40	.34	<.001	H3d supported

Note. n.s. = not significant.

regard to their number of views, $F(1, 460) = 39.96, p < .001$, part. $\eta^2 = .08$. Online entertainment videos received more views ($M_{\text{logged}} = 17.49$, $SD_{\text{logged}} = 1.73$) than online political videos ($M_{\text{logged}} = 15.84$, $SD_{\text{logged}} = 1.31$). According to the second hypothesis, political videos should receive more comments than entertainment videos. Contrary to our hypothesis, entertainment videos received more comments ($M_{\text{logged}} = 9.38$, $SD_{\text{logged}} = 1.99$) than political videos ($M_{\text{logged}} = 8.77$, $SD_{\text{logged}} = 1.91$), $F(1, 444) = 10.61, p = .001$, part. $\eta^2 = .02$, when controlling for the number of received views. Based on these results (see Table 2), Hypothesis 1 was supported while Hypothesis 2 was rejected.

Hypothesis 3a stated that online entertainment videos receive more likes than online political videos. We found (see Table 2) no difference between the number of likes that entertainment videos received and the number of likes that political videos, $F(1, 452) = 1.94$, $p = .164$, received. Thus, Hypothesis 3a was rejected. Hypothesis 3b stated that online political videos receive more dislikes than online entertainment videos. Contrary to our hypothesis, we found (see Table 2) that entertainment videos received more dislikes ($M_{\text{logged}} = 9.20$, $SD_{\text{logged}} = 2.00$) than political videos ($M_{\text{logged}} = 7.83$, $SD_{\text{logged}} = 1.65$), $F(1, 452) = 6.94$, $p = .009$, part. $\eta^2 = .02$, while controlling for the number of received views. Therefore, Hypothesis 3b was rejected. According to Hypothesis 3c, comments written in response to online entertainment videos have a stronger positive valence than comments written in response to online political videos. In contrast to this hypothesis, we found (see Table 2) that comments written in response to political videos had a stronger positive valence ($M = 1.74$, $SD = 0.23$) than comments written in response to entertainment videos ($M = 1.59$, $SD = 0.27$), $F(1, 417) = 9.84, p = .002$, part. $\eta^2 = .02$. Hypothesis 3d stated that comments written in response to online political videos have a stronger negative valence than comments written in response to online entertainment videos. Our results supported this hypothesis, as we found (see Table 2) that comments written in response to political videos had a stronger negative valence ($M = -1.98$, $SD = 0.40$) than comments written in response to entertainment videos ($M = -1.33$, $SD = 0.24$), $F(1, 417) = 209.88$, $p < .001$, part. $\eta^2 = .34$.

According to Hypothesis 4a, the positive valence of comments has a positive effect on the number of replies that they receive. In support of this hypothesis, we found a significant, positive effect of the comments' positive valence on the number of received replies (see Table 3). An increase in a comment's positive valence led to more replies ($e^{0.033} = 1.03, p = .013$). Hypothesis 4b stated that as comments have a stronger positive valence, they receive more likes. There was a significant, positive effect of comments' positive valence on the number of received likes which supported Hypothesis 4b (see Table 3). An increase in a comment's positive valence led to more likes ($e^{0.221} = 1.25$, $p < .001$).

According to Hypothesis 5a, comments receive more replies as they express a stronger negative valence. Supporting Hypothesis 5a, the results showed a significant, positive effect of the comments'

Table 3. Effect of the Mean Comment Positive and Negative Valence on the Number of Received Replies and Comments (Hypotheses 4 and 5 and Research Questions 1 and 2).

	All Videos						Entertainment Videos						Political Videos					
	Mean Comment Positive Valence			Mean Comment Negative Valence			Mean Comment Positive Valence			Mean Comment Negative Valence			Mean Comment Positive Valence			Mean Comment Negative Valence		
Variables	PR	CI 95%	*p*	PR	CI 95%	*p*	PR	CI 95%	*p*	PR	CI 95%	*p*	PR	CI 95%	*p*	PR	CI 95%	*p*
Comment reply count (untransformed)	1.03	[.01, .06]	.013	0.74	[−.32, −.27]	<.001	1.03	[.00, .06]	.031	0.75	[−.31, −.26]	<.001	1.00	[−.06, −.07]	*n.s.*	0.86	[−.20, −.11]	<.001
Comment like count (untransformed)	1.25	[.21, .24]	<.001	0.94	[−.07, −.05]	<.001	1.28	[.23, .27]	<.001	0.87	[−.15, −.12]	<.001	0.99	[−.06, .03]	*n.s.*	1.14	[.10, .16]	<.001

Note. CI = confidence interval; *n.s.* = not significant; PR = poisson regression (negative binomial).

Table 4. Effect of the Number of Comments Posted in Response to a Video on the Variance in the Comments' Valence (Hypothesis 6 and Research Question 3).

	All Videos			Entertainment Videos			Political Videos		
	Video Comment Count (Logged)			Video Comment Count (Logged)			Video Comment Count (Logged)		
Variables	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*
Variance Comment Positive Valence	−.00	.01	*n.s.*	.01	.01	*n.s.*	−.06	.02	< .001
Variance Comment Negative Valence	.00	.01	*n.s.*	.01	.01	*n.s.*	.01	.03	*n.s.*

Note. n.s. = not significant; *SE* = standard error.

negative valence on the number of received replies (see Table 3). As a comment's negative valence increased, it received more replies ($e^{-0.295} = 0.74, p < .001$). Hypothesis 5b stated that as comments' negative valence becomes stronger, they receive less likes. We found a significant, positive effect of the comments' negative valence on the number of received likes (see Table 3). As a comment's negative valence increased, it received more likes ($e^{-0.060} = 0.94, p < .001$). As this is not in line with Hypothesis 5b, the hypothesis was rejected.

Hypothesis 6 posed that as a video receives more comments, the comments' valence becomes more consistently positive or negative. The results showed (see Table 4) that the number of comments that videos receive did not affect the variance in the positive valence expressed through comments (p = .853) nor did it affect the negative valence expressed through comments p = .813). Based on these results, Hypothesis 6 was rejected.

Research Question 1a asked how the relationship between the positive valence of user comments and the number of replies that they receive differs for entertainment and political videos. For entertainment videos, we found that an increase in a comment's positive valence led to more replies ($e^{0.032} = 1.03$, $p = .031$). For comments written in response to political videos, we found that an increase in a comment's positive valence did not lead to more replies (p = .973). Thus, only for comments written in response to entertainment videos did we find that they received more replies as their positive valence became stronger (see Table 3). Research question 1b asked how the relationship between the positive valence of comments and the number of likes that they receive differs between entertainment and political videos. For comments written in response to entertainment videos, the results showed that an increase in a comment's positive valence led to more likes ($e^{0.249} = 1.28, p < .001$). For comments written in response to political videos, we found that an increase in a comment's positive valence had no significant effect on the likes that the comment received (p = .533). Thus, while the positive valence did positively affect the number of likes that comments of entertainment videos receive, it did not affect the number of likes that comments of political videos receive (see Table 3).

Research Question 2a asked how the relationship between the negative valence of user comments and the number of replies that they receive differs for entertainment and political videos. For entertainment videos, we found than an increase in a comment's negative valence led to more replies ($e^{-0.284} = 0.75$, $p < .001$). For comments written in response to political videos, we found that an increase in a comment's negative valence led to more replies ($e^{-0.154} = 0.86, p < .001$). Thus, for both comments written in response to entertainment videos and to political videos, we found that they received more replies as their negative valence became stronger (see Table 3). Research Question 2b asked how the relationship between the negative valence of comments and the number of likes that they receive differs between entertainment and political videos. For comments written in response to entertainment videos, we found that an increase in a comment's negative valence led to more likes ($e^{-0.136} = 0.87$, $p < .001$). However, for comments written in response to political videos, we found that an increase in a comment's negative valence led to fewer likes ($e^{0.132} = 1.14$, $p < .001$). Thus, while for entertainment videos, comments' negative valence increased the number of likes that comments received, for political videos, comment's negative valence decreased the number of likes that comments received (see Table 3).

Research Question 3 asked how the relationship between the number of comments that videos receive and the variance in the comments' valence differs for entertainment and political videos. The result of the regression analysis on the entertainment videos showed that the number of comments that an entertainment video received had no significant effect on the variance of its comments' positive valence ($p = .365$). The results of the regression analysis on the political videos showed that as the number of received comments increased, the variance in their positive valence decreased ($\beta = -0.58$, $B = -.06$, $SE = .02$, $p < .001$). For entertainment videos, we found that the number of received comments did not affect the variance in its comments' negative valence ($p = .175$). For political videos, the results also showed that the number of comments that an entertainment video received had no significant effect on the variance in its comments' negative valence ($p = .758$). Thus, only for political videos did the number of received comments decrease the variance in the comments' positive valence (see Table 4).

Discussion

The present study analyzed the characteristics of the social information of YouTube videos. The results contribute to our knowledge in multiple ways. First, our results indicate that the video type is related to viewing frequency and to the amount of social information that is created in response to the video. Entertainment videos were watched more often than political videos, as differences in the numbers of views between entertainment and political videos indicate. We also found that when controlling for the number of video views, entertainment videos received more comments and more dislikes than political videos. These findings imply that YouTube users are more engaged with the content of entertainment videos, leading them to express their opinions about entertainment videos more often than their opinions about political videos. This is in line with previous studies which found that YouTube is predominantly used to satisfy entertainment needs (Chau, 2010; Khan, 2017).

In addition, our results show that the comments that entertainment videos receive are more neutral than the comments of political videos, which, accordingly, contain information with a more extreme positive as well as negative valence. The more extreme valence of user comments to political videos may have different causes: the controversial nature of political media content (Chen & Berger, 2013; Jeong, 2003; Liu et al., 2015), polarization that often occurs when political content is discussed (Sunstein, 2002), or the higher inclination of viewers of political videos to express extreme views. Thus, our first conclusion is that the audience of entertainment videos is more active in creating user comments, but the audience of political videos writes more extreme comments.

A second main contribution of this study is the finding that existing social information itself is a driver of new social information. We hypothesized that more positive comments receive more replies and more likes. However, we only found this to be true for entertainment videos but not for political videos. This indicates that viewers of political videos respond differently to previously created social information than entertainment video viewers. But more research is needed to fully understand the mechanisms behind these findings. In line with our assumptions, we found that more negative comments receive more replies. We also hypothesized that more negative comments receive fewer likes because likes express a positive valence instead of a negative valence. However, for entertainment videos we found that more negative comments receive more likes. A reason for this may be that when video viewers assign a like to a negative comment, they do so to express agreement with the negative comment. In that case, the like confirms the negative valence of the comment.

A third contribution of this study is that the results expand our knowledge of how Turner's (1982) referent informational influence framework applies to online social information. Based on Turner's framework, we hypothesized that as more comments are written in response to a video, a more consistent usage of positive and negative valence would emerge. However, we only found this relationship for the positive valence of comments written in response to political videos. An explanation for this finding is that positive comments may stand out against the negativity prevalent in political videos: because political videos are often negative, positive comments are easily noticed and, thus, the group norm of expressing a positive valence in comments may be more visible. The fact that we did not find this pattern for entertainment videos may be because group norms are less visible in the comments of entertainment videos than in the comments of political videos. Group norms in the comments section of political videos may be more visible because of the more extreme valence of the comments: as the valence of comments about political videos is more extreme, it can be more easily recognized, have a stronger impact on the valence of newly created comments, and, thus, increase the homogeneity of the comments' valence. In contrast, comments about entertainment videos are more neutral and thus less likely to indicate group norms and to increase the homogeneity of user comments' valence. This implies that in online communication through social information, the visibility of group norms might be an important factor: if norms are visible enough, they can be adopted by users as this study found to be the case for comments about political videos. Future research could thus benefit from investigating how attributes of social information affect its visibility and impact on media users' behaviors.

In addition to making an empirical and theoretical contribution, our findings have practical implications for online platforms. Many platforms want to prevent users from posting negative comments and using uncivil language. Although strategies exist to deal with this problem, such as moderating comments or actively engaging with commenters to influence group norms (Ksiazek, 2015; Stroud, Scacco, Muddiman, & Curry, 2015), they require a large effort for platform managers. Our findings offer an alternative to these labor-intensive solutions. Platforms can design the presentation of social information in such a way that the visibility of comments which present undesired behaviors is decreased. For example, platform managers can detect the valence of comments and decrease the visibility of negative comments by placing them at the end of the list of comments. This would make it less likely that the behavioral norms presented in these comments are adopted in new comments. This way, they can decrease the presence of negative comments in a more efficient manner.

At least three limitations of the current study must be noted. First, because YouTube contains a plethora of videos and their social information, a representative sample of videos can hardly be collected. For feasibility reasons, we relied on information provided by Social Blade to initially sample potential English-language entertainment and political videos with a high number of videos views. The initial reliance on information from Social Blade and the focus on English-language videos biases our sample. Hence, the findings of this study should be interpreted as a first exploration of online videos' social information, and more research on this topic is needed to complete our knowledge of the characteristics of online social information. Second, the strength of the effect sizes found in this study varied substantially. While some of the effect sizes were relatively strong (e.g., the effect of video type on the negative valence of comments), other effect sizes were relatively weak (e.g., the effect of video type on the number of received comments and dislikes). These differences in effect sizes should be considered when evaluating the results of this study. Third, this study found that the social information of YouTube videos depends on the type of the video and on previously created social information. However, our research design does not allow us to preclude that there exist additional determinants of social information. Notably, the characteristics of the audiences of entertainment and political videos may influence social information too. For example, Khan (2017) found that men are more likely to read comments on YouTube than women. In addition, men are more likely to watch political videos because they have a relatively high political engagement compared to women (Jerit & Barabas, 2017; Verba, Burns, & Schlozman, 1997). Future research could shed more light on this matter by investigating what differences exist between audiences of different video types and how this relates to systematic differences in the amount and valence of videos' social information. By gathering information about YouTube audiences using survey data or by analyzing commenters' avatars and user names, scholars could investigate possible relationships between audience characteristics and the social information that they create.

The current study explored a topic about which little research exists and which lacks an established theoretical framework. Hence, this study was a first attempt to systematically investigate the social information of entertainment media content and the factors that determine its amount and valence. As there is still much left to explore with regard to the current topic, future research could contribute to our knowledge by developing theories that deepen our understanding of the factors that determine the characteristics of online social information. This could be done by expanding the present study through the investigation of social information on other online platforms. Although social information constitutes a major part of YouTube's content, it also plays an important role on other social media, such as Facebook. A notable difference between YouTube and Facebook is that while on the first platform users are exposed to social information that is mainly created by strangers, on Facebook this social information is created mostly by acquaintances. Research indicates that this factor may alter how users respond to social information created by others (Postmes, Spears, Sakhel, & De Groot, 2001; Walther et al., 2010). Hence, investigating social information on other platforms such as Facebook can broaden our insights into the factors on which online social information's characteristics depend.

The present study contributed to our knowledge by exploring to what social information YouTube users are exposed and how this depends on the type of video that they watch. This way, it contributed to a broader understanding of how social information is created and what consequences this may have for the video views, (dis)likes, and comments seen by online video viewers. For researchers studying social media, it is important to know that much can be gained by looking beyond the content that is provided by social media platforms—knowledge about the total content to which social media users are exposed can be gained by examining platforms' social information.

Authors' Note

The authors would like to thank Damian Trilling and Bob van de Velde for their help and advice in using Python to write the scripts for this study. Moreover, they would like to thank Amber van der Wal for her helpful contributions to the development of the codebook that was used for coding the YouTube videos. The data collected and analyzed for the current study can be made available through a data set that is in line with the requirements of the General Data Protection Regulation (effective May 25, 2018) by request to A. Marthe Möller (a.m.moller@uva.nl).

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

Abdul-Mageed, M. M. (2008). Online news sites and journalism 2.0: Reader comments on Al Jazeera Arabic. *tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, *6*, 59–76. Retrieved from: https://www.triple-c.at/index.php/tripleC/article/view/78.

Berger, J., & Milkman, K. L. (2012). What makes online content viral? *Journal of Marketing Research*, *49*, 192–205. doi:10.1509/jmr.10.0353

Cameron, J., & Geidner, N. (2014). Something old, something new, something borrowed from something blue: Experiments on dual viewing TV and Twitter. *Journal of Broadcasting & Electronic Media*, *58*, 400–419. doi:10.1080/08838151.2014.935852

Cha, M., Kwak, H., Rodriguez, P., & Moon, S. (2007). I tube, you tube, everybody tubes: Analyzing the world's largest user generated content video system. *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement, San Diego*, 1–14. doi:10.1002/yd.376

Chau, C. (2010). YouTube as a participatory culture. *New Directions for Youth Development*, *128*, 65–74. doi:10.1002/yd.376

Chen, Z., & Berger, J. (2013). When, why, and how controversy causes conversation. *Journal of Consumer Research*, *40*, 580–593. doi:10.1086/671465

Chmiel, A., Sienkiewicz, J., Thelwall, M., Paltoglou, G., Buckley, K., Kappas, A., & Hołyst, J. A. (2011). Collective emotions online and their influence on community life. *PLoS One*, *6*, 1–8. doi:10.1371/journal.pone.0022207

Choi, D., Han, J., Chung, T., Ahn, Y. Y., Chun, B. G., & Kwon, T. T. (2015). Characterizing conversation patterns in Reddit: From the perspectives of content properties and user participation behaviors. *Proceedings of the 2015 ACM on Conference on Online Social Networks, Palo Alta*, 233–243.

Coe, K., Kenski, K., & Rains, S. A. (2014). Online and uncivil? Patterns and determinants of incivility in newspaper website comments. *Journal of Communication*, *64*, 658–679. doi:10.1111/jcom.12104

Coxe, S., West, S. G., & Aiken, L. S. (2009). The analysis of count data: A gentle introduction to Poisson regression and its alternatives. *Journal of personality assessment*, *91*, 121–136. doi:10.1080/00223890802634175

Google. (2017). *Google language-detection*. Retrieved from: https://pypi.python.org/pypi/langdetect.

Gray, J. (2009). *Television entertainment*. New York, NY: Routledge.

Hancock, J. T., Gee, K., Ciaccio, K., & Lin, J. M. (2008). I'm sad you're sad: Emotional contagion in CMC. *Proceedings of CSCW '08: ACM Conference on Computer Supported Cooperative Work, San Diego*, 295–298. doi:10.1145/1460563.1460611

Hancock, J. T., Landrigan, C., & Silver, C. (2007). Expressing emotion in text-based communication. *Proceedings of the SIGCHI conference on Human factors in computing systems, San Jose*, 929–932. doi:10.1145/1240624.1240764

Jeong, A. C. (2003). The sequential analysis of group interaction and critical thinking in online threaded discussions. *The American Journal of Distance Education*, *17*, 25–43. doi:10.1207/S15389286AJDE1701

Jerit, J., & Barabas, J. (2017). Revisiting the Gender Gap in Political Knowledge. *Political Behavior*, *39*, 817–838. doi:10.1007/s11109-016-9380-6

Khan, M. L. (2017). Social media engagement: What motivates user participation and consumption on YouTube? *Computers in Human Behavior*, *66*, 236–247. doi:10.1016/j.chb.2016.09.024

Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massivescale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, *111*, 10779–10779. doi:10.1073/pnas.1412469111

Ksiazek, T. B. (2015). Civil interactivity: How news organizations' commenting policies explain civility and hostility in user comments. *Journal of Broadcasting and Electronic Media*, *59*, 556–573. doi:10.1080/08838151.2015.1093487

Ksiazek, T. B. (2018). Commenting on the news: Explaining the degree and quality of user comments on news websites. *Journalism Studies*, *19*, 650–673. doi:10.1080/1461670X.2016.1209977

Liu, Q., Zhou, M., & Zhao, X. (2015). Understanding News 2.0: A framework for explaining the number of comments from readers on online news. *Information & Management*, *52*, 764–776. doi:10.1016/j.im.2015.01.002

Postmes, T., Spears, R., & Lea, M. (2000). The formation of group norms in computer-mediated communication. *Human Communication Research*, *26*, 341–371. doi:10.1111/j.1468-2958.2000.tb00761.x

Postmes, T., Spears, R., Sakhel, K., & De Groot, D. (2001). Social influence in computer-mediated communication: The effects of anonymity on group behavior. *Personality and Social Psychology Bulletin*, *27*, 1243–1254. doi:10.1177/01461672012710001

Potter, W. J. (2009). *Arguing for a general framework for mass media scholarship*. Thousand Oaks, CA: Sage.

Prior, M. (2005). News vs. entertainment: How increasing media choice widens gaps in political knowledge and turnout. *American Journal of Political Science*, *49*, 577–592. doi:10.1111/j.1540-5907.2005.00143.x

Rubin, A. M. (1983). Television uses and gratifications: The interactions of viewing patterns and motivations. *Journal of Broadcasting & Electronic Media*, *27*, 37–51. doi:10.1080/08838158309386471

Rubin, A. M. (2009). Uses-and-Gratification perspective on media effects. In J. Bryant & M. B. Oliver (Eds.), *Media effects: Advances in theory and research* (pp. 532–550). New York, NY: Routledge.

Schuck, R. T., Vliegenthart, R., Boomgaarden, H. G., Elenbaas, M., Azrout, R., Van Spanje, J., & de Vreese, C. H. (2013). Explaining campaign news coverage: How medium, time, and context explain variation in the media framing of the 2009 European parliamentary elections. *Journal of Political Marketing*, *12*, 8–28. doi:10.1080/15377857.2013.752192

Semetko, H. A., & Valkenburg, P. M. (2000). Framing European politics: A content analysis of press and television news. *Journal of Communication*, *50*, 93–109. doi:10.1111/j.1460-2466.2000.tb02843.x

Siersdorfer, S., Chelaru, S., Nejdl, W., & San Pedro, J. (2010). How useful are your comments? Analyzing and predicting Youtube comments and comment ratings. *Proceedings of the 19th International Conference on World Wide Web, Raleigh*, *15*, 891–900. doi:10.1145/1772690.1772781

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. doi:10.1016/j.chb.2016.02.086

Stroud, N. J., Scacco, J. M., Muddiman, A., & Curry, A. L. (2015). Changing deliberative norms on news organizations' Facebook sites. *Journal of Computer-Mediated Communication*, *20*, 188–203. doi:10.1111/jcc4.12104

Sunstein, C. R. (2002). The law of group polarization. *Journal of Political Philosophy*, *10*, 175–195. doi:10.1111/1467-9760.00148

Tatar, A., Antoniadis, P., Dias de Amorim, M., Leguay, J., Limbourg, A., & Fdida, S. (2011). Predicting the popularity of online articles based on user comments. *Proceedings of the International Conference on Web Intelligence, Mining and Semantics*. Sogndal, Norway. doi:10.1145/1988688.1988766.

Thelwall, M., Buckley, K., & Paltoglou, G. (2012). Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology*, *63*, 163–173. doi:10.1002/asi.21662

Thelwall, M., Sud, P., & Vis, F. (2011). Commenting on YouTube videos: From Guatemalan Rock to El Big Bang. *Journal of the American Society for Information Science and Technology*, *1463*, 616–629. doi:10.1002/asi.21679

Tsagkias, M., Weerkamp, W., & De Rijke, M. (2009). Predicting the volume of comments on online news stories. *Proceedings of the 18th ACM conference on Information and knowledge management*, Hong Kong, 1765–1768. doi:10.1145/1645953.1646225

Turner, J. C. (1982). Towards a cognitive redefinition of the social group. In H. Tajfel (Ed.), *Social identity and intergroup relations* (pp. 15–40). New York, NY: Cambridge University Press.

Van de Velde, R. N. (2017). *YouTube Scraper*. Retrieved from: https://github.com/bobvdvelde/YouTube_scraper/tree/master.

Verba, S., Burns, N., & Schlozman, K. L. (1997). Knowing and caring about politics: Gender and political engagement. *The Journal of Politics*, *59*, 1051–1072. doi:10.2307/2998592

Vorderer, P. (2001). It's all entertainment—sure. But what exactly is entertainment? Communication research, media psychology, and the explanation of entertainment experiences. *Poetics*, *29*, 247–261. doi:10.1016/S0304-422X(01)00037-7

Vorderer, P., Klimmt, C., & Ritterfeld, U. (2004). Enjoyment: At the heart of media entertainment. *Communication Theory*, *14*, 388–408. doi:10.1111/j.1468-2885.2004.tb00321.x

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*, 393–409. doi:10.1080/08838151.2017.1309414

Walther, J. B., DeAndrea, D., Kim, J., & Anthony, J. C. (2010). The influence of online comments on perceptions of antimarijuana public service announcements on YouTube. *Human Communication Research*, *36*, 469–492. doi:10.1111/j.1468-2958.2010.01384.x

Weber, P. (2014). Discussions in the comments section: Factors influencing participation and interactivity in online newspapers' reader comments. *New Media & Society*, *16*, 941–957. doi:10.1177/1461444813495165

Young, L. E., & Leonardi, P. M. (2012). Social issue emergence on the Web: A dual structurational model. *Journal of Computer-Mediated Communication*, *17*, 231–246. doi:10.1111/j.1083-6101.2011.01566.x

Yun, G. W., Park, S.-Y., Holody, K., Yoon, K. S., & Xie, S. (2013). Selective moderation, selective responding, and balkanization of the blogosphere: A field experiment. *Media Psychology*, *16*, 295–317. doi:10.1080/15213269.2012.759462

Ziegele, M., Weber, M., Quiring, O., & Breiner, T. (2017). The dynamics of online news discussions: effects of news articles and reader comments on users' involvement, willingness to participate, and the civility of their contributions. *Information, Communication & Society*, 1–17. doi:10.1080/1369118X.2017.1324505

Author Biographies

A. Marthe Möller (MSc) is a PhD candidate at the Amsterdam School of Communication Research (University of Amsterdam). E-mail: a.m.moller@uva.nl

Rinaldo Kühne is an assistant professor at the Amsterdam School of Communication Research (University of Amsterdam). E-mail: r.j.kuhne@uva.nl

Susanne E. Baumgartner is an assistant professor at the Amsterdam School of Communication Research (University of Amsterdam). E-mail: s.e.baumgartner@uva.nl

Jochen Peter is a professor of Media Entertainment at the University of Amsterdam and the current scientific director of the Amsterdam School of Communication Research. E-mail: j.peter@uva.nl