Article

The digital outcry: What incites participation behavior in an online firestorm?

Marius Johnen [ORCID]
University of Hamburg, Germany

Marc Jungblut
Ludwig-Maximilians-Universität Munich, Germany

Marc Ziegele
Johannes Gutenberg University Mainz, Germany

Abstract

Brands, celebrities, or politicians are increasingly facing enormous online outrages in response to moral misconducts. These online firestorms are characterized by high message volume, indignant tonality, and negative opinion climate. Based on the concept of moral panics, this article analyzes why people join online firestorms. We argue that participation behavior is driven by a moral compass and a desire for social recognition. Results of an experiment and a content analysis of user comments show that a higher number of participants decreases users' willingness to participate but fosters compliance with the prevalent opinion and tonality of the comments. We also observe that a higher moral arousal of the issue increases perceived similarity with previous participants, which in turn affects whether and how people participate. In total, our results indicate the importance of social context for participation behavior in an online firestorm.

Keywords

moral panic, online firestorm, social media, user comments, user participation

Corresponding author:
Marius Johnen, Institute of Marketing, University of Hamburg, Moorweidenstr. 18, D-20148 Hamburg, Germany.
Email: marius.johnen@uni-hamburg.de

In spring of 2015, the beer brand Bud Light experienced the power of indignant social media users. As part of their "#UpForWhatever" campaign, they printed the slogan "The perfect beer for removing 'no' from your vocabulary for the night" on their beer cans. After one user posted a photo of this label on social media platform Reddit, other users started blaming Bud Light for promoting date rape. The outrage spread over to other social media platforms such as Twitter, where more people joined the bandwagon and attacked Bud Light indignantly. At some point, media reported on the issue, increasing its publicity and fostering more indignation on social media. The brand was compelled to apologize, but within two or three days after the initial post, indignation subsided (Pallotta, 2015).

This incident is a characteristic example of an "online firestorm," which is a frequent form of public online communication that happens not only to brands and public institutions but also to celebrities, politicians or other individuals. Pfeffer et al. (2014) labeled this new phenomenon as a "sudden discharge of large quantities of messages containing negative word-of-mouth and complaint behavior against a person, company, or group in social media networks" (Pfeffer et al., 2014). Furthermore, such outcries are caused by a perceived moral misconduct, followed by many users' reactions within a short period of time. To a vast majority, these user articulations contain negative opinions toward the accused person and involve intense indignation (Pfeffer et al., 2014). This development can be amplified by spillover effects to online and traditional media.

Prior research has addressed this dynamic phenomenon primarily on a macro-level. For example, scholars analyzed how an online firestorm evolves on a social media platform (Jackson and Welles, 2015), which structural factors foster the development of firestorms (Pfeffer et al., 2014), how social media communication interacts with traditional media and brand reactions (Hauser et al., 2017; Hewett et al., 2016), or how news media report on this topic (Einwiller et al., 2016). On a micro-level, research has analyzed individual user motivations to participate in aggressive and offensive forms of online participation that are related to online firestorms, but are rather static, such as flaming (Alonzo and Aiken, 2004), hate speech (Erjavec and Kovačič, 2012), cyberbullying (Varjas et al., 2010), negative electronic word of mouth (Hennig-Thurau et al., 2004), and trolling (Buckels et al., 2014). Still, to our knowledge, no study has investigated the individual factors that stimulate participation in an online firestorm, which differs from other forms of online communication in terms of moral concern, consensus, hostility, disproportionality, and volatility (see explanation in next section for details).

Against this background, this study analyzes why people join in an indignant attack in a crowd-based outrage such as an online firestorm. Specifically, we investigate how specific characteristics of the firestorm (i.e. moral arousal and number of previous participants) and individually perceived social context affect users' willingness to participate as well as the content and tonality of the comments they contribute themselves. As a theoretical framework, we adapt the sociological concept of "moral panics" (Goode and Ben Yehuda, 1994) for online firestorms. We assume that individual participation behavior is driven by a "moral compass" and a desire for social recognition. The former refers to the perception that attacking the accused individual/institution is socially appropriate in situations of high moral arousal. The latter implies that the perceived chance of receiving positive feedback by others affects participation behavior.

Through a combination of an online experiment and a content analysis, we show that the number of previous participants directly affects whether and how to join an online firestorm. More specifically, a higher number of previous participants decreases willingness to participate, but fosters compliance with the prevalent opinion as well as an aggressive tonality. The actual issue, expressed by the severity of moral misconduct, indirectly affects willingness to participate and direction of expressed opinion only indirectly via perceived similarity with previous participants.

In total, this article makes three key contributions: First, we apply the concept of moral panics as a theoretical approach to explain online firestorms. Second, this article extends literature on public online communication by analyzing how the interplay of discussion characteristics and perceived social context affects different features of participation behavior. Third, as a methodological contribution, we integrate a content analysis of user comments into an experimental setting to analyze specific facets of participation behavior.

Online firestorms as moral panics

Several terms have been used to label the phenomenon that we refer to as online firestorms, for example, "collaborative brand attacks" (Rauschnabel et al., 2016) "shitstorms" (Einwiller et al., 2016), and "social media backlash" (BBC, 2015). Although scholars and journalists often emphasize the novelty of this phenomenon, we rather consider it as a well-known social phenomenon within a new context. More precisely, online firestorms show characteristics that resemble so-called moral panics (Goode and Ben Yehuda, 1994).

Cohen (1972) denoted a moral panic as a collective behavior during which "a condition, episode, person or group of persons emerges to become defined as a threat to societal values and interests" (p. 1), followed by stereotypical presentations as well as moral condemnation by societal elites. Cohen used this terminology to analyze the deviant behavior of youth subcultures. Another example of a moral panic is the medieval witchhunt in Europe, where individuals were persecuted because of being suspected to have evil magical powers (Cohen, 1972).

In their seminal work on moral panics, Goode and Ben-Yehuda (1994) identified five key characteristics of moral panics: (1) *concern* (about a seemingly threatening behavior of a person or group against moral values), (2) *hostility* (toward the accused), (3) *disproportionality* (the concern is exaggerated regarding the objective threat), (4) *consensus* (in terms of a perceived agreement about a threat by a group of people), and (5) *volatility* (a moral panic emerges and subsides quickly).

These characteristics closely resemble the process of an online firestorm. As described above, in an online firestorm, we can observe moral concern shared by a majority of participating users. This moral concern is exaggerated in terms of being a real threat to the society. In addition, opinions expressed in online firestorms contain hostility and indignation toward the accused and occur within a short period of time (Pfeffer et al., 2014).

Despite this congruence, we consider an online firestorm as a specific form of a moral panic that is different to traditional moral panics due to certain characteristics of the online communication context. First, the *formation* of an online firestorm happens more easily than in a traditional moral panic (Rost et al., 2016). For example, social network

sites (SNSs) such as Facebook provide a platform to communicate with a large number of remote and/or unfamiliar people (Ellison et al., 2011). More specifically, SNSs provide access to a large audience, and users can easily find like-minded people who share the indignation about a certain issue (Ellison et al., 2011). Thus, broad attention and support to a concern about an issue can be generated without media coverage—although journalistic reports still can increase reach (Einwiller et al., 2016). Second, the *speed* of an online firestorm is distinctly higher than in traditional moral panics. Studies on the diffusion of trends via social media show that oftentimes, only several hours elapse between the first appearance of a topic and the peak of audience attention toward it (Leskovec et al., 2009). This amplifies the short-term nature of moral panics along with a condensed development of indignation and a temporally decreased visibility of the issue. Third, the *disinhibition and indignation* in participants' behavior may appear stronger than in traditional moral panics (Rost et al., 2016). In computer-mediated communication (CMC), users often are either not aware of or not concerned about the potential negative consequences of their behavior (Suler, 2004). These perceptions can result in disinhibited behavior, with users being more inclined to accuse deviants and engage in hostile behavior. Besides, factors such as the absence of non-verbal cues, invisibility, and asynchronicity can foster rude language, hatred, and threats (Walther, 2011). Previously published indignant comments can further intensify this tendency by suggesting that aggression and hostility are dominant group norms. These group norms can then put pressure on subsequent individuals to adjust their own behavior to fit with these norms (Hsueh et al., 2015; Spears et al., 1990). Thus, uncivil behavior of participants is supposed to be more prevalent than in traditional moral panics.

Furthermore, the key features of moral panics proposed by Goode and Ben-Yehuda (1994) can serve as criteria to distinguish online firestorms from related opinion-based forms of public online communication, such as electronic word-of-mouth (eWOM), user comments on online news (in short form: news comments), and flaming (Hennig-Thurau et al., 2004; O'Sullivan and Flanagin, 2003; Ziegele and Quiring, 2013). Compared to these related forms, online firestorms have a more specific topic of communication (i.e. a moral concern), comprise a greater and "condensed" level of hostility and aggressiveness, include a more disproportionate presentation of the discussed issue, show less diversity in opinions and polarization (i.e. higher perception of consensus), and involve a sudden accumulation and rapid declines of user comments (i.e. higher volatility). Based on these specific circumstances, we consider an investigation of the *individual* motivations to engage in online firestorms as necessary.

Theoretical framework for participation behavior in online firestorms

Goode and Ben-Yehuda (1994) identified two core motivations of participating in moral panics. On the one hand, they assumed that the expressed concern by participants is based on *morality*, "that is, deeply and genuinely felt attitudes and sentiments" (Goode and Ben-Yehuda, 1994: 159). Thus, participation behavior is driven by the need to defend one's moral sense against deviants. On the other hand, *self-interest* is considered to be a key motivation of participation behavior. More specifically, the pursuit of gaining

something of value such as power, respectability, or recognition should have an effect on participation (Goode and Ben-Yehuda, 1994).

Based on these two components, we argue that in an online firestorm as a specific form of a moral panic, participation behavior is driven by (1) a "moral compass" and (2) a desire for social recognition. These two drivers are strongly connected to the social context due to our expounded consideration of online firestorms as a specific form of moral panic. A *moral compass* refers to an individual's judgments about the issue at hand as well as about the corresponding user reactions. If the moral misconduct is perceived to violate one's sense of morality, this should stimulate the need to join the attack (Eisenberg, 2000; Lindenmeier et al., 2012). This violation can refer to both intrinsic moral values and perceived social norms (Rost et al., 2016). In addition, the reactions and judgments of other participants are highly visible in the online communication context. These reactions/comments could likely suggest a specific social context which impacts the behavior of subsequent users. Perceived morality of a group is relevant when evaluating and joining that group (Leach et al., 2007). Thus, besides evaluating the moral misconduct itself, an individual should also judge the societal and moral appropriateness of other users attacking the denounced actor. If this judgment turns out to be positive, an individual is more likely to identify with the acting group and adapt corresponding group behavior (Spears et al., 1990; Tajfel and Turner, 1986). Consequently, we assume that the stronger the perception of a moral violation and/or the more the users' attack against the accused is perceived to be appropriate, the more likely individuals join the attack in an online firestorm.

A *desire for social recognition* is related to self-interest motives (Goode and Ben Yehuda, 1994). Prior research on public online communication has shown that self-enhancing motives, such as the desire to be liked and to win esteem, are relevant drivers for participation (Baumeister, 1982; Berger, 2014). Gaining positive social recognition by others is hence considered as a goal and benefit (Baumeister, 1982). We assume that in a rather anonymous online context, in which oneself and other users may not be identifiable, social recognition is not necessarily limited to imagined reactions but still can be expressed by observable rewards. More specifically, within an online firestorm, social recognition can be obtained through positive feedback of other, predominantly unknown users in terms of "likes," "retweets;" approving comments, and so on. In the past few years, this behavior of attempting to receive social recognition in social media by expressing a moral concern has also been discussed controversially under the term "virtue signaling" in the media (Bartholomew, 2015). Thus, we argue that individuals assess online firestorms with regard to the chance of receiving positive social recognition after participation, based on past user comments (e.g. reach, climate of opinion, and discussion culture) as well as the expected future development of the online firestorm. Furthermore, this evaluation should influence the perception of an appropriate participation behavior, that is, which opinion to express and what tonality to use to increase the chance of getting positive feedback. Therefore, we assume that the higher the expected social recognition for participation, the more likely individuals will join the attack in an online firestorm.

Hypotheses development

Based on this theoretical framework, we derive our hypotheses to explain participation behavior in an online firestorm. Figure 1 summarizes the assumed relationships between
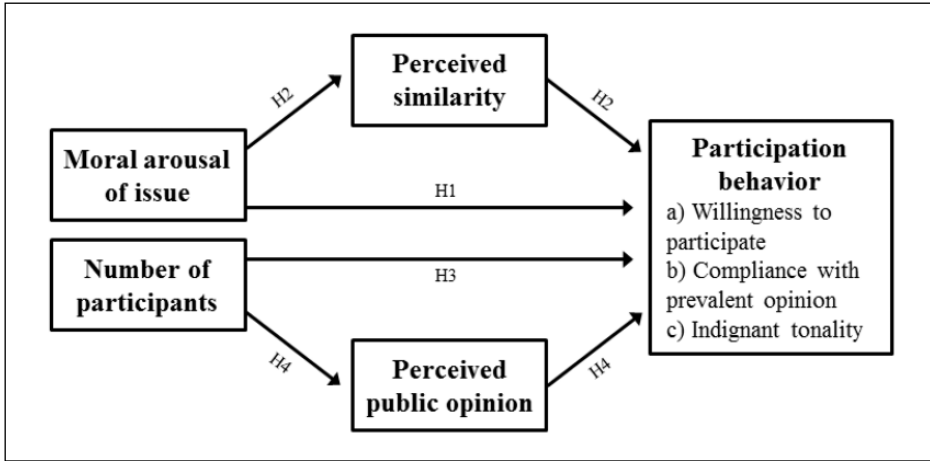
Figure 1. Conceptual framework for participation behavior in an online firestorm.

online firestorm characteristics (i.e. moral arousal of the issue and number of previous participants), perceived social context (i.e. perceived similarity of participants and perceived public opinion about the issue), and participation behavior on the theoretical grounding of an individual's moral compass and desire for social recognition. As the aim of our research is to explain why people join in an indignant attack in an online firestorm, we have three dependent variables: willingness to participate ("why join in?"), compliance with the prevalent opinion ("why join in an attack?"), and indignant tonality in a comment ("why join indignantly?").

As outlined before, moral concern as a discussion topic is a key characteristic of online firestorms. Prior research has shown the relevance of topic features for online user participation (e.g. Berger, 2014). For example, when an issue incites cognitive involvement or dissonance, users are more willing to post a comment (Ziegele and Quiring, 2013). Furthermore, Berger and Milkman (2012) revealed that content with a high emotional arousal is more likely to elicit active user engagement. In the context of online firestorms, the arousal of moral concern may also vary, depending on the individual moral compass, that is, the dissent from one's moral beliefs. Accordingly, the more severe an individual perceives a denounced moral misconduct, the stronger should this perception evoke negative emotions (Lindenmeier et al., 2012; Rost et al., 2016). Hence, a higher moral arousal of the discussed issue in an online firestorm should lead to a higher willingness to participate and join the attack. Furthermore, with an increase of negative moral emotions, observers should be more likely to behave in an indignant way themselves, for example, by writing a more indignant comment (Eisenberg, 2000). Thus, we postulate the following:

> H1. High moral arousal of an issue in an online firestorm will increase an individual's willingness to participate (H1a), the compliance with the prevalent opinion (H1b), and the level of indignant tonality of the comment (H1c).

Users' perceptions of online discussions can impact identity formation and identification processes, which then influence participation behavior (Hsueh et al., 2015; Spears and Lea, 1994). As elaborated above, we assume that a user judges the appropriateness of an online firestorm based on her or his perceived societal and moral appropriateness of attacking the denounced actor. The more this evaluation turns out to be positive, the more likely the user is to regard the group of actively participating users as relevant and, thus, identifies with them in a desire to feel as a part of this group. As a result, a user might adopt the group's behavioral patterns and join the firestorm in a similar way. Research from social psychology has demonstrated that in various CMC situations, individual discussion behavior is affected by the perception of the behavior of others (Lee, 2006). The SIDE model (Spears et al., 1990) suggests that the lack of individual information that is prevalent in CMC de-emphasizes intra-group differences and thereby enhances the identification with an existing group. A high attachment to the group, then, makes an individual more susceptible to social influence. Consequently, a higher identification with the other participants of the firestorm should increase a user's willingness to participate as well as the compliance with the prevalent group opinion. Finally, research on group identification processes has demonstrated that group identification holds an impact on opinion polarization (Lee, 2006) and might foster self-enhancement and attacks toward outgroups (e.g. Hsueh et al., 2015). Consequently, we assume:

> H2. Perceived similarity of participants mediates the effect of moral arousal of the issue on willingness to participate (H2a), on compliance with the prevalent opinion (H2b), and on indignant tonality in a comment (H2c).

Within social media environments, user postings often display aggregated statistics of the number of participants who have already engaged with the postings (Knobloch-Westerwick et al., 2005). Past research has found that such "popularity indications" can increase recipients' engagement with online content (Knobloch-Westerwick et al., 2005). Regarding the impact of the number of previous participants on the willingness to join an online firestorm, it can first be assumed that users infer the size of the potential audience of their own comment from these statistics. A higher number of popularity indicators should then represent a larger audience. If users' participation behavior is driven by a moral compass and a desire for social recognition, a larger audience should then increase users' perceptions that (1) the moral concern is shared by a significant number of others and is more societally appropriate (i.e. perceived consensus), and (2) their own comment will more likely be recognized and appreciated by these other users. These perceptions should stimulate the individual willingness to join the online firestorm.

However, we assume that the relationship between the number of previous participants and recipients' willingness to participate is not necessarily a linear one. A very large number of participants could make recipients perceive that every possible argument has already been voiced in the discussion and that the value or the impact of an additional contribution will be very small. Regarding social recognition, this perception could also decrease the subjective chance to write a comment that stands out from the mass and receives positive feedback. Hence, compared with a moderate amount of

popularity indications, a very large amount could decrease recipients' motivation to participate in the online firestorm. We thus hypothesize the following:

H3a: The number of previous participants of an online firestorm and recipients' willingness to participate have an inverted U-relationship.

Regarding the impact of the number of previous participants on the content and tonality of individual postings, theories such as the spiral of silence suggest that positive social recognition in public environments is more likely to occur when users comply with the opinion of the (perceived) majority (Noelle-Neumann, 1974). Hence, the larger the number of previous participants with consonant opinions, the more can recipients be expected to comply with this opinion, that is a perceived consensus. At the same time, based on the recipients' desire for social recognition, it can be assumed that they will not simply repeat the arguments of the previous users, but try to add something "unique" to their own comment to stand out from the mass. In the context of online firestorms this uniqueness can, for example, be achieved by voicing an even more radical, provocative, or indignant opinion than other users. Therefore, we hypothesize as follows:

H3b/H3c: The larger the number of previous participants, the higher the compliance with the prevalent opinion (H3b) and the higher the level of indignant tonality in a comment (H3c).

Finally, studies have shown that news comments and eWOM messages can shift observers' perceptions of the public opinion toward issues in the direction of the postings (Lee and Jang, 2010). Users therefore seem to use the posts of others to infer how the society (or a part thereof) thinks about issues. This phenomenon can be explained with exemplification theory (e.g. Zillmann, 2002). Exemplars allow individuals to apply cognitive heuristics to generalize from the characteristics of one exemplar to those of a whole group or category of people. This application of cognitive heuristics can also explain why online users often infer public opinion from single user-generated postings (Peter et al., 2014). Although exemplification research has demonstrated that individuals make broad generalizations based on a small number of opinions, it has been proposed that the effect multiplies when the number of consonant opinions increases (Zerback and Fawzi, 2016; Zillmann, 2002).

Regarding online firestorms, these arguments suggest that a hostile and indignant opinion climate can quickly emerge from a few negative user postings (Lee and Jang, 2010). The perception that these comments represent public opinion could be even stronger when a high number of users have already participated with consistent opinions. For recipients who hold a similar opinion, spiral of silence theory suggests that the perception of one's own opinion being shared by the majority of the society can increase the individual willingness to join in and voice a similar opinion (Noelle-Neumann, 1974). Based on these arguments, we propose the following:

H4a/b/c: Perceived public opinion mediates the effect of the number of previous participants on recipients' willingness to participate (H4a), on compliance with the prevalent opinion (H4b), and on indignant tonality in a comment (H4c).

Method

Procedure and sample

We conducted an online experiment with a fictitious online firestorm as a stimulus. The online context of the experiment was supposed to provide a more realistic environment for an online firestorm than a traditional laboratory experiment. We manipulated characteristics of the online firestorm (moral arousal of the issue and number of previous participants) and measured participants' perceived social context, willingness to participate, as well as several control variables in a subsequent survey. Furthermore, compliance with the prevalent opinion and level of indignant tonality were assessed through a content analysis of comments written by respondents as a reaction to the stimulus.

We recruited $N=663$ active social media users through a commercial online panel in a large Western European country, being nationally representative regarding age and gender ($M_{age} = 42.01$, $SD_{age} = 12.91$, 54% women; 26% high school diploma). Participants received a cash incentive. We randomly assigned each respondent to one condition in a 2 (moral arousal of issue: low vs. high) × 3 (number of previous participants: low vs. medium vs. high number of user reactions) between-subjects design. This design procedure as well as our verbal instructions should minimize the likelihood that respondents discern our hypotheses (Shimp et al., 1991). Besides, to ensure that subjects did not tailor their behavior based on observation, they were informed on the introduction page that their behavior was anonymized.

Stimulus material

To simulate a typical setting of an online firestorm, we used a fictitious user posting on a brand's Facebook page accompanied by complying and indignant comments by other users. We chose an existing airline to maintain adequate realism perceptions and involvement by respondents. In a pre-study ($N=444$), we tested different scenarios regarding credibility, realism, and discriminability of different stimuli. Thereby, the latter refers to a sufficient difference in perceived moral arousal of the firestorm as well as in the perceived number of previous participants.

Based on the results of our pre-study, we selected the stimuli that received the best ratings regarding our criteria so that respondents in our main study were confronted with the following stimuli: On the first page, in a Facebook layout, respondents were shown an initial user posting that contained a complaint about a derogatory behavior of two stewardesses against another passenger during a flight. The user described the furtive dialogue of the two stewardesses and considered it to be "unacceptable, unprofessional, and disrespectful." She also accused the brand for having a ruinous human resources management. In the low moral arousal condition, the stewardesses derided the eating habits of a passenger, compared him with pigs and laughed about him. In the

high moral arousal condition, they defamed a passenger due to his religion by calling him "vermin" and making jokes about his death. Below this posting, each respondent received the usual Facebook statistics about user reactions to this posting, that is, number of likes, comments, and shares. Depending on the experimental condition, these numbers were low (51 likes, 12 comments, 5 shares), medium (510, 121, 53) or high (5010, 1212, 513).

On the next page, the same posting was presented to respondents, now including 10 comments from fictional users. All comment authors expressed a consensus in negative opinion and indignation toward the brand or the stewardesses (e.g. "such a behavior is just embarrassing," "how can you hire those people?" "I hope these stewardesses will be fired!"). Specifically, indignation was shown by a negative emotional arousal, threats, insults, or sarcasm. The content was the same across all conditions, but the information about number of likes to each comment was proportionally different according to the respective condition (i.e. 3 to 7 likes [low] vs. 28 to 72 likes [medium] vs. 280 to 728 likes [high]).

Measures

Following the experimental manipulation, we measured respondents' willingness to participate on 7-point scales using three items from Nekmat and Gonzenbach (2013) (e.g. *If you encountered such a posting and comments on Facebook, how likely would you add a comment?* 1=*very unlikely*, 7=*very likely*, α=.94). Subsequently, we asked respondents who showed at least a minimum level of willingness to participate ($M_{WTP} > 1$ on the aggregated 7-point scale; $n=475$) to write a comment in an open textbox that was embedded below the stimulus. We used this comment to measure respondents' compliance with the prevalent opinion and indignant tonality (see next section). To increase external validity of our findings, we excluded those respondents who showed a negative tendency in their willingness to participate (i.e. $M_{WTP} < 4$), resulting in a reduced subsample of $n=244$. Nevertheless, as a robustness check, we repeated the following analyses (i.e. content analyses and regression analyses) including all respondents who wrote a comment ($M_{WTP} > 1$, $N=475$), but our findings did not change substantially.

Afterwards, respondents indicated their perceived similarity with the participants of the firestorm on a three-item 7-point Likert scale adapted from Feick and Higie (1992) (e.g. *The people who took part … are quite a bit like me*, α=.96). Subsequently, we measured respondents' perceived public opinion based on Zerback and Fawzi (2016) by asking the following:

> Suppose that there will be a national poll with 1,000 respondents tomorrow, with all participants being informed prior about this issue. What do you estimate: What percentage of respondents would express a negative opinion toward the brand regarding the issue?

Using a slider bar, respondents could indicate the estimated percentage on a 0 to 100% scale (recoded to fit a range from 0 to 10).

Furthermore, we measured several control variables that might affect our dependent variables. We asked respondents about their social media productivity (four items concerning activity on Facebook, Twitter, comment sections on news media, and YouTube,

adapted from Hocevar et al., 2014; α=.78), about their general involvement with the issue (three items adapted from Mittal, 1985; α=.94), and brand perception (three items adapted from Zaichkowsky, 1985; α=.82).

Finally, to check the success of our manipulation, we measured respondents' perception of the stimuli at the end of the survey by using two items: "*There was a high number of 'likes' in response to the initial posting*" (7-point scale; for number of previous participants) and "*I consider the events in the airplane to be ... morally acceptable vs. morally question-able*" (7-point semantic differential; for moral arousal of issue). These items were selected based on qualitative assessments by 12 communication scholars and were also tested in our pre-study. Results of both procedures ensured that these items obtain satisfactory convergence and are a valid measure of our manipulation (Perdue and Summers, 1986).

Content analysis

We content analyzed all comments from participants whose willingness to participate was above the center scale (i.e. $M_{WTP}>4$) in order to assess compliance with the prevalent opinion and tonality. Overall, this resulted in a corpus of 244 texts. The content analysis was conducted by two of the authors and one trained research assistant.

Users' compliance with the prevalent opinion was calculated on the basis of an author's general tendency toward the post (i.e. having a positive, negative, or ambivalent opinion), her/his opinion about the attacked brand, her/his evaluation of the stewardesses, and the evaluation of the author of the original post, as these indicators expressed the prevalent opinion in the discussion. The four variables were each measured on a 5-point scale. The resulting variable "compliance with the prevalent opinion" is the arithmetic mean of all non-missing values from these four items, resulting again in a 5-point measure with 5 indicating a high compliance with the prevalent negative opinion ($M=4.23$, $SD=1.31$, α=.87).

We also used the content analysis to identify indignant tonality in the comments.

We examined whether a participant's comment contained insults, sarcasm, threats, or negative emotions with a high degree of arousal such as anger, fury, disgust, or outrage. A comment was regarded as using indignant tonality when the expressed opinion within the comment was negative (i.e. $M_{compliance}>3$) and it either contained an insult, a threat, sarcasm, or negative emotions with a high degree of arousal. Consequently, indignant tonality was a binary variable in the research design, with 37% of the comments being coded as indignant.

Initially, 50 comments were coded by each coder to ensure reliability. We then measured Krippendorff's alpha coefficients for all single variables. Overall, the reliability measures were satisfying (see Table 1).

Results

Manipulation checks

Our intended manipulations regarding the moral arousal of the issue (two conditions) and the number or previous participants (three conditions) were successful. Respondents considered the events in the high-arousal condition to be more morally questionable than in the low-arousal condition ($M_{high}=6.23$, $SD=1.21$; $M_{low}=5.22$, $SD=1.68$, $t(661)=8.89$,

Table 1. Descriptive statistics and reliability measures of content analysis.

Variable	Incidents	*M* (*SD*)	Krippendorff's alpha
General tendency of the post	244	4.18 (1.40)	.841
Expressed opinion toward brand	67	3.65 (1.86)	1.00
Expressed evaluation of the stewardesses	216	4.68 (1.02)	.748
Expressed evaluation of the author of the post	34	4.00 (1.57)	.855
Usage of insults in the post	15	–	1.00
Usage of threat(s)	17	–	.745
Usage of sarcasm	9	–	.681
Usage of arousing negative emotions	107	–	.681

Number of coded incidents, mean values, and standard deviations refer to the reduced sample of $N = 244$. The first four variables are coded on a 1–5 scale (for the first three variables, 5 indicates a negative opinion, for the fourth variable, 5 indicates a positive opinion). The calculation of Krippendorff's alpha is based on 50 comments coded by all three coders.

$p < .001$). Also, the perceived number of participants differed across the experimental conditions ($M_{low} = 4.20$, $SD = 1.51$; $M_{med} = 4.57$, $SD = 1.55$; $M_{high} = 4.83$, $SD = 1.60$; $t_{low \text{ vs. } med} (442) = 2.61$, $p = .005$; $t_{med \text{ vs. } high} (443) = 1.68$, $p = .047$).

Results for willingness to participate

We used Hayes' (2013) PROCESS macro to analyze our data and to test our hypotheses. Thus, we conducted simple mediation analyses (model 4) using ordinary least squares path analysis. Willingness to participate served as dependent variable, perceived similarity and perceived public opinion were used as mediating variables and moral arousal, number of participants as well as our several control variables were independent variables. Table 2 (second column) shows the combined results of the model with perceived similarity as a mediator of moral arousal and the model of perceived public opinion as a mediator of the number of previous participants. As we hypothesized an inverted U-relationship between the number of previous participants and respondents' willingness to participate, we used the medium number of participants as a reference category in the dummy coding for this variable.

Results show that moral arousal of the issue had a positive total effect on willingness to participate ($\beta_c = .394$, $p = .003$), supporting H1$_a$. In other words, willingness to participate increased by .394 scale points (on a 7-point scale) in the high moral arousal condition compared to the low moral arousal condition. This effect appeared as indirect-only mediation (Zhao et al., 2010) as the moral arousal of the issue did not have a direct significant effect on willingness to participate ($\beta_{c'} = .158$, $p = .200$). However, respondents in the high moral arousal condition significantly differed from those in the low moral arousal condition in terms of their perceived similarity with participants ($\beta_a = .649$, $p < .001$). A higher perceived similarity then increased respondents' willingness to participate ($\beta_b = .364$, $p < .001$). Accordingly, a bias-corrected 95% bootstrap confidence interval (CI) for this indirect effect ($\beta_{ab} = .236$) based on 10,000 bootstrap samples did not contain zero (0.132 to 0.366), supporting H2$_a$.

Moreover, the number of previous participants significantly affected willingness to participate. More precisely, a high number of previous participants decreased willingness to participate by .352 scale points compared to a medium number ($\beta_{c'high}=-.352$, $p=.018$). However, we did not find a significant difference between the conditions of low and medium number of previous participants on willingness to participate ($\beta_{c'low}=-.038$, $p=.799$), thus H3$_a$ is only partially supported. Furthermore, perceived public opinion was not affected by the number of previous participants ($\beta_{a,low}=.197$, $p=.319$; $\beta_{a,high}=.105$, $p=.597$) and did not have an effect on willingness to participate ($\beta_b=.046$, $p=.120$). Consequently, we could not observe any relative indirect effect of number of previous participants on willingness to participate via perceived public opinion; both bias-corrected 95% bootstrap CI contained zero ($\beta_{ab,low}=.009$, lower CI$=-.005$, upper CI$=.048$; $\beta_{ab,high}=.005$, lower CI$=-.010$, upper CI$=.041$). Therefore, H4$_a$ is not supported.

Results for compliance with the prevalent opinion

To analyze compliance with the prevalent opinion in an online firestorm, we refer to the variable that we computed based on the content analysis of respondents' reaction to the stimulus. We used the same procedure as for the dependent variable "willingness to participate," except with the reduced sample of $N=244$ and with a different dummy coding for number of previous participants due to our hypothesis (i.e. low number as a reference category).

Again, we did not find a direct effect of moral arousal of the issue in an online firestorm on the direction of the expressed opinion ($\beta_{c'}=.120$, $p=.390$; see Table 2, third column). Also, the total effect was not significant, in contrast to H1$_b$ ($\beta_c=.217$, $p=.169$). However, moral arousal of the issue indirectly increased compliance with the prevalent opinion through its effect on perceived similarity. Specifically, respondents in the high moral arousal condition showed a marginally stronger perceived similarity of participants in the stimuli than those in the low moral arousal condition ($\beta_a=.437$, $p=.076$), and respondents with a higher perceived similarity expressed a higher agreement with the prevalent opinion ($\beta_b=.285$, $p<.001$). The bias-corrected CI for the indirect effect ($\beta_{ab}=.128$) was slightly above zero (.001 to .285), which supports H2$_b$.

The number of previous participants in the online firestorm only directly affected the direction of the expressed opinion. Confronted with a higher number of previous participants, respondents complied with the prevalent opinion more strongly, which supports H3$_b$ ($\beta_{c', medium}=.338$, $p=.041$; $\beta_{c', high}=.580$, $p<.001$). Perceived public opinion, however, was not affected by the number of previous participants when only analyzing the respondents with a positive willingness to participate ($\beta_{a,medium}=-.095$, $p=.756$; $\beta_{a,high}=.365$, $p=.254$). However, respondents increasingly complied with the prevalent opinion when they perceived that the comments of the other participants represented public opinion ($\beta_b=.124$, $p=<.001$). Nevertheless, we did not find a relative indirect effect of the number of previous participants on compliance with the prevalent opinion via perceived public opinion as both bias-corrected 95% bootstrap CI contained zero ($\beta_{ab,medium}=-.012$, lower CI$=-.111$, upper CI$=.062$; $\beta_{ab,high}=.045$, lower CI$=-.020$, upper CI$=.166$). Thus, H4$_b$ is rejected.

Table 2. Mediation models for participation behavior in online firestorms.

Predictors	Willingness to participate			Compliance with opinion			Indignant tonality in comment			
	β	t	p	β	t	p	β	Odds	z	p
Moral arousal of issue (1: high)	.158	1.283	.200	.120	0.861	.390	.808	2.242	2.666	.008
Number of participants (1: low)	-.038	-0.254	.799	.338	2.051	.041	.284	1.329	0.785	.433
Number of participants (1: med.)			Ref. category			Ref. category	Ref. category			
Number of participants (1: high)	-.352	-2.372	.018	.580	3.365	<.001	.772	2.164	2.055	.040
Perceived similarity	.364	10.829	<.001	.285	7.761	<.001	.332	1.394	3.931	<.001
Perceived public opinion	.046	1.556	.120	.124	3.531	<.001	.182	1.200	2.137	.033
Control variables										
Social media productivity	.956	9.846	<.001	.013	0.142	.887	-.049	0.952	-0.249	.804
Issue involvement	.226	5.441	<.001	.012	0.233	.816	-.098	0.907	-0.850	.395
Brand perception	.091	1.852	.065	-.097	-1.782	.076	-.282	0.754	-2.357	.018
Age	.015	3.196	.002	.006	1.066	.288	-.002	0.998	-0.164	.870
Gender	-.177	-1.426	.154	.151	1.087	.278	.382	1.466	1.260	.208
Constant	-1.799	-3.801	<.001	1.534	2.735	.007	-2.502	0.819	-2.065	.039
Direct effects on mediators	β	t	p	β	t	p				
Moral arousal of issue →Perceived similarity	.649	4.597	<.001	.437	1.783	.076	[Same as in model for compliance with prevalent opinion]			
Number of participants (low/med.) →Perceived public opinion[a]	.197	0.997	.319	-.095	-0.312	.756				
Number of participants (high) →Perceived public opinion[a]	.105	0.529	.597	.365	1.143	.254				

(Continued)

Table 2. (Continued)

Indirect effects	Willingness to participate			Compliance with opinion			Indignant tonality in comment		
	β	Boot SE	95% CI	β	Boot SE	95% CI	β	Boot SE	95% CI
Moral arousal of issue →Perceived similarity	.236	.059	[0.132, 0.366]	.128	.071	[0.001, 0.285]	.145	.098	[−0.009, 0.366]
Number of participants (low/med.) →Perceived public opinion[a]	.009	.012	[−0.005, 0.048]	−.012	.042	[−0.111, 0.062]	−.017	.069	[−0.193, 0.102]
Number of participants (high) →Perceived public opinion[a]	.005	.011	[−0.010, 0.041]	.045	.044	[−0.020, 0.166]	.066	.076	[−0.044; 0.260]
N (adjusted R^2)	663 (.406)			244 (.360)			244 (Nagelkerke R^2=.291)		
F/LR statistics	$F(10, 652)=46.17, p<.001$			$F(10, 233)=14.64, p<.001$			LR $\chi^2 (10)=59.69, p<.001$		

Note: β coefficients are unstandardized; number of bootstrap samples for bias-corrected bootstrap confidence interval (CI): 10,000; all predictors also included in models for mediators. [a]Direct and indirect effects of multicategorical variable 'number of participants' are relative effects to reference category 'low number of participants' for 'compliance with opinion' and 'indignant tonality' and to reference category 'medium number of participants' for 'willingness to participate'.Results for indignant tonality

In a last step, we analyzed the inciting factors of indignant tonality in an online firestorm. We used the same procedure and sample as in the previous analysis for compliance with the prevalent opinion. As our dependent variable is binary, we used a logistic regression for the estimation of paths on indignant tonality.

The data in Table 2 (fourth column) show that a higher moral arousal of the issue in an online firestorm directly increased the likelihood of writing an indignant comment, which supports H1$_c$ ($\beta_{c'}$=.808, p=.008; β_c=.953, p=.003). In other words, participants wrote an indignant comment twice as likely in the high moral arousal condition compared to the low moral arousal condition (odd ratio=2.242). Perceived similarity mediated the effect of moral arousal on the dependent variable only by trend. This variable significantly increased the likelihood of writing an indignant comment (β_b=.332, p<.001). Still, the bias-corrected bootstrap 95% CI for the indirect effect contained zero (β_{ab}=.145, lower 95% CI=−.009, upper 95% CI=.366; lower 90% CI=.017; upper 90% CI=.333).

Furthermore, we found that in the condition of a high number of previous participants, the likelihood of writing an indignant comment was significantly higher, supporting H3$_c$ ($\beta_{c',medium}$=.284, p=.433; $\beta_{c', high}$=.772, p=.040). Besides, the more respondents perceived public opinion to be congruent with the opinions of the online firestorm, the more likely they wrote an indignant comment (β_b=.182, p=.033). Still, the indirect effect of number of previous participants on indignation in a comment via perceived public opinion was not significant ($\beta_{ab,medium}$=−.017, lower CI=−.193, upper CI=.102; $\beta_{ab,high}$=.066, lower CI=−.044, upper CI=.260). Consequently, H4$_c$ is rejected.

Discussion

Online firestorms have become a frequent phenomenon in social media over the past few years. Celebrities, politicians, public institutions, as well as corporate brands can be subject to sudden crowd-based attacks as a result of a seemingly moral misconduct. Although online firestorms usually last only for a short period of time, they may cause severe negative consequences beyond that time frame, such as resignations, dismissals, image damages, or financial losses (Hewett et al., 2016). Thus, understanding the mechanisms of online firestorms appears necessary to reduce their negative effects.

To explain behavioral patterns in online firestorms, we transferred the concept of moral panics to the computer-mediated communication context. This allowed us to derive two key mechanisms that should drive individual participation behavior in online firestorms, namely morality and self-interest. In this context, a "moral compass" suggests that individuals rely on a judgment about the societal appropriateness of the attack. This judgment about social norms affects the evaluation of and identification with the accusatory crowd, and in turn has an effect on participation behavior.

In our empirical study, the moral arousal of the events only affected the expressed tonality directly, but not the willingness to participate or compliance with the prevalent opinion. Instead, we observed indirect-only effects via perceived similarity for those two participation features. Thus, the individual perceptions of other participants are a crucial factor to explain the effect of moral arousal of the issue in an online firestorm. Similarly, our findings also suggest that the moral compass is mostly affected by the perceived social appropriateness of attacking the denounced actor rather than by intrinsic moral values.

Moreover, our findings suggest that the desire for social recognition is a strong driver of participation in online firestorms. In our study, participants used information about the number of previous participants to adjust their participation behavior. More specifically, a higher number of previous participants decreased the willingness to participate but increased compliance with the prevalent opinion and the likelihood of writing an indignant comment. These results suggest that the information about the number of previous participants affect individuals' expectation of social recognition. When too many people already participated in the firestorm, the difficulty of standing out from the mass makes participation rather unattractive as the chance for social recognition is low. This decreased attractiveness of participation could also be a reason for the short-term nature of online firestorms (Pfeffer et al., 2014). Furthermore, a higher tendency to conform to previous comments with an increase in number of previous participants may be explained by participants' expectation that this behavior elicits more positive feedback than a deviant comment. Still, in a later stage of an online firestorm, the necessity to stand out from the mass fosters stronger indignation in a comment. These two findings of an increasingly homogeneous opinion climate and an amplified aggressive atmosphere due to the desire of social recognition contribute to understanding the negative dynamics in online firestorms (Pfeffer et al., 2014) and should be investigated more thoroughly in future research.

Taken as a whole, the perceived social context is a key factor when analyzing participation behavior in online firestorms. When people expect social support, either due to the similarity of previous participants or due to a consistent public opinion, they rather join in an indignant attack against a brand, celebrity, or politician. Similar to research findings on flaming (O'Sullivan and Flanagin, 2003), users in online firestorms thus seem to consider perceived situational social norms in their decision whether and how to participate. However, the peculiarities of online firestorms in terms of moral concern, hostility, disproportionality, consensus, and volatility suggest a stronger relevance of social context than other forms of online communication.

While the number of previous participants had some effects on users' participation behavior, our participants apparently did not use this number as a cue for gauging public opinion toward the issue. While this finding is unexpected in light of the predictions of exemplification theory, recent studies in another context have reported similar results (Zerback and Fawzi, 2016). The sheer number of participants might not be a sufficiently strong cue for how the whole society thinks about an issue. Rather, it has been argued that the direction of individual opinions themselves will influence this perception (Zerback and Fawzi, 2016). Hence, recipients' estimates of public opinion might have differed if we had contrasted solely indignant discussions against discussions, which also included more moderate opinions. Owing to our theoretical considerations about the nature of online firestorms, we did however not vary the distribution of different opinions in the comment sections. Still, future research might consider manipulating the distribution of opinions as an option to measure whether this affects recipients' perception of the opinion climate.

Our study has some limitations that should be considered when interpreting the results and that may be addressed in future studies. First, we used an isolated case of an online firestorm on a single social media platform. Typically, online firestorms spread over several platforms and also to news media (Pfeffer et al., 2014). These multiple confrontations might

lead to amplified perceptions of an opinion climate. Besides, although being rare in a typical online firestorm, our stimuli contained no contradictory opinions or reactions, neither of supporting users or the accused subject. Furthermore, respondents in our study only received a limited number of social cues (e.g. no information about origin or interests). In particular, our experimental design induced a high level of anonymity, which may be less distinct on certain platforms. These additional factors can also affect potential benefits regarding the desire for social recognition, thus being potentially important for participation behavior (Hutchens et al., 2015). For instance, a lower level of anonymity might either foster participation due to a higher value of social recognition by identifiable others, or hinder participation due to concerns about contempt by disagreeing (known) participants.

Additionally, we limited our study design to a specific example of an online firestorm and to two relevant discussion characteristics. However, other features that differentiate online firestorms may have an influence on participation behavior. An online firestorm against a celebrity or politician might therefore induce stronger tendencies to participate. Similarly, other characteristics that describe the nature of the issue in an online firestorm, such as moral controversy, proximity, or topicality, may have an effect on participation behavior. In this regard, future research should also consider field data when analyzing participation behavior in online firestorms.

Despite these limitations, our research provides important theoretical and practical implications toward understanding individual motivations to participate in online firestorms and toward handling these public outcries. Both moral arousal of the issue and the number of previous participants affected individual participation behavior, while the actual issue mostly showed indirect effects via perceived similarity of participants. This finding underlines the importance of other users' behavior in online communication environments for individual decision making. From a business perspective, public relation (PR) and social media managers should be aware that any moral complaint seems to have a certain potential to escalate if the social fabric of the firestorm is homogeneous. Thus, observing group dynamics and opinion diversity is important, as a homogeneous negative opinion climate may foster further indignation. Consequently, appropriate and fast responses and/or mobilizing supporters might mitigate the negative spiral in online firestorms (Rauschnabel et al., 2016). Besides, addressing users' desire for social recognition by rewarding constructive criticism with a response and hence, with additional attention, could mitigate the negative dynamics of a digital outcry.

Acknowledgements

Funding

ORCID iD

Marius Johnen https://orcid.org/ 0000-0001-6635-8038

References

Alonzo M and Aiken M (2004) Flaming in electronic communication. *Decision Support Systems* 36(3): 205–213.

Bartholomew J (2015) The awful rise of "virtue signaling." *The Spectator*, 18 April. Available at: https://www.spectator.co.uk/2015/04/hating-the-daily-mail-is-a-substitute-for-doing-good/

Baumeister RF (1982) A self-presentational view of social phenomena. *Psychological Bulletin* 91(1): 3–26.

BBC (2015) Twitter ad sparks social media backlash. Available at: http://www.bbc.com/news/technology-34665439

Berger J (2014) Word of mouth and interpersonal communication: a review and directions for future research. *Journal of Consumer Psychology* 24(4): 586–607.

Berger J and Milkman KL (2012) *Journal of Marketing Research* 49(2): 192–205.

Buckels EE, Trapnell PD and Paulhus DL (2014) Trolls just want to have fun. *Personality and Individual Differences* 67: 97–102.

Cohen S (1972) *Folk Devils and Moral Panics: The Creation of the Mods and Rockers*. Oxford: Martin Robertson.

Einwiller S, Viererbl B and Himmelreich S (2016) Journalists' coverage of online firestorms in German-language news media. *Journalism Practice* 11(9): 1178–1197.

Eisenberg N (2000) Emotion, regulation, and moral development. *Annual Review of Psychology* 51(1): 665–697.

Ellison NB, Steinfield C and Lampe C (2011) Connection strategies: social capital implications of Facebook-enabled communication practices. *New Media & Society* 13(6): 873–892.

Erjavec K and Kovačič MP (2012) "You don't understand, this is a new war!": analysis of hate speech in news web sites' comments. *Mass Communication & Society* 15(6): 899–920.

Feick L and Higie RA (1992) The effects of preference heterogeneity and source characteristics on ad processing and judgements about endorsers. *Journal of Advertising* 21(2): 9–24.

Goode E and Ben-Yehuda N (1994) Moral panics: culture, politics, and social construction. *Annual Review of Sociology* 20(1): 149–171.

Hauser F, Hautz J, Huttler K, et al. (2017) Firestorms: modeling conflict diffusion and management strategies in online communities. *The Journal of Strategic Information Systems* 26(4): 285–321.

Hayes AF (2013) *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression Based Approach*. New York: Guilford Press.

Hennig-Thurau T, Gwinner KP, Walsh G, et al. (2004) Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet? *Journal of Interactive Marketing* 18(1): 38–52.

Hewett K, Rand W, Rust TR, et al. (2016) Brand buzz in the echoverse. *Journal of Marketing* 80(3): 1–24.

Hocevar KP, Flanagin AJ and Metzger MJ (2014) Social media self-efficacy and information evaluation online. *Computers in Human Behavior* 39: 254–262.

Hsueh M, Yogeeswaran K and Malinen S (2015) "Leave your comment below": can biased online comments influence our own prejudicial attitudes and behaviors? *Human Communication Research* 41(4): 557–576.

Hutchens MJ, Cicchirillo VJ and Hmielowski JD (2015) How could you think that?!?!: understanding intentions to engage in political flaming. *New Media & Society* 17(8): 1201–1219.

Jackson SJ and Welles BF (2015) Hijacking #myNYPD: social media dissent and networked counterpublics. *Journal of Communication* 65(6): 932–952.

Knobloch-Westerwick S, Sharma N, Hansen DL, et al. (2005) Impact of popularity indications on readers' selective exposure to online news. *Journal of Broadcasting & Electronic Media* 49(3): 296–313.

Leach CW, Ellemers N and Barreto M (2007) Group virtue: the importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology* 93(2): 234–249.

Lee EJ (2006) When and how does depersonalization increase conformity to group norms in computer-mediated communication? *Communication Research* 33(6): 423–447.

Lee EJ and Jang YJ (2010) What do others' reactions to news on internet portal sites tell us? Effects of presentation format and readers' need for cognition on reality perception. *Communication Research* 37(6): 825–846.

Leskovec J, Backstrom L and Kleinberg J (2009) Meme-tracking and the dynamics of the news cycle. In: *Proceedings of the 15Th ACM SIGKDD international conference on knowledge discovery and data mining,* Paris, 28 June–1 July, pp. 497–505. New York: ACM.

Lindenmeier J, Schleer C and Pricl D (2012) Consumer outrage: emotional reactions to unethical corporate behavior. *Journal of Business Research* 65(9): 1364–1373.

Mittal B (1985) A comparative analysis of four scales of consumer involvement. *Psychology & Marketing* 12(7): 663–682.

Nekmat E and Gonzenbach WJ (2013) Multiple opinion climates in online forums role of website source reference and within-forum opinion congruency. *Journalism & Mass Communication Quarterly* 90(4): 736–756.

Noelle-Neumann E (1974) The spiral of silence: a theory of public opinion. *Journal of Communication* 24(2): 43–51.

O'Sullivan PB and Flanagin AJ (2003) Reconceptualizing 'flaming' and other problematic messages. *New Media & Society* 5(1): 69–94.

Pallotta F (2015) Bud light's #UpForWhatever slogan causes social media uproar. *CNN Money Online*, 28 April. Available at: http://money.cnn.com/2015/04/28/media/bud-light-bottle-slogan-uproar/

Perdue BC and Summers JO (1986) Checking the success of manipulations in marketing experiments. *Journal of Marketing Research* 23(4): 317–326.

Peter C, Rossmann C and Keyling T (2014) Exemplification 2.0. *Journal of Media Psychology* 26(1): 19–28.

Pfeffer J, Zorbach T and Carley KM (2014) Understanding online firestorms: negative word-of-mouth dynamics in social media networks. *Journal of Marketing Communications* 20(1–2): 117–128.

Rauschnabel PA, Kammerlander N and Ivens BS (2016) Collaborative brand attacks in social media: exploring the antecedents, characteristics, and consequences of a new form of brand crises. *Journal of Marketing Theory and Practice* 24(4): 381–410.

Rost K, Stahel L and Frey BS (2016) Digital social norm enforcement: online firestorms in social media. *PLoS ONE* 11(6): e0155923.

Shimp TA, Hyatt EM and Snyder DJ (1991) A critical appraisal of demand artifacts in consumer research. *Journal of Consumer Research* 18(3): 273–283.

Spears R and Lea M (1994) Panacea or panopticon? The hidden power in computer-mediated communication. *Communication Research* 21(4): 427–459.

Spears R, Lea M and Lee S (1990) De-individuation and group polarization in computer-mediated communication. *British Journal of Social Psychology* 29(2): 121–134.

Suler J (2004) The online disinhibition effect. *Cyberpsychology & Behavior* 7(3): 321–326.

Tajfel H and Turner JC (1986) The social identity theory of intergroup behaviour. In: Worchel S and Austin WG (eds) *Psychology of Intergroup Relations*. Chicago, IL: Nelson-Hall Publishers, pp. 7–24.

Varjas K, Talley J, Meyers J, et al. (2010) High school students' perceptions of motivations for cyberbullying: an exploratory study. *The Western Journal of Emergency Medicine* 11(3): 269–273.

Walther JB (2011) Theories of computer-mediated communication and interpersonal relations. In: Knapp ML and Daly JA (eds) *The Handbook of Interpersonal Communication*. 4th ed. Thousand Oaks, CA: SAGE, pp. 443–479.

Zaichkowsky JL (1985) Measuring the involvement construct. *Journal of Consumer Research* 12(3): 341–352.

Zerback T and Fawzi N (2016) Can online exemplars trigger a spiral of silence? Examining the effects of exemplar opinions on perceptions of public opinion and speaking out. *New Media & Society* 19(7): 1034–1051.

Zhao X, Lynch JG and Chen Q (2010) Reconsidering Baron and Kenny: myths and truths about mediation analysis. *Journal of Consumer Research* 37(2): 197–206.

Ziegele M and Quiring O (2013) Conceptualizing online discussion value: a multidimensional framework for analyzing user comments on mass-media websites. In: Cohen EL (ed.) *Communication Yearbook 37*. New York: Routledge, pp. 125–153.

Zillmann D (2002) Exemplification theory of media influence. In: Bryant J and Zillmann D (eds) *Media Effects: Advances in Theory and Research*. Mahwah, NJ: Lawrence Earlbaum Associates, pp. 19–41.

Author biographies

Marius Johnen is a PhD student at the Institute of Marketing at University of Hamburg, Germany. His research focuses on interactions between consumers and brands in online environments as well as branding strategies in social media.

Marc Jungblut is PhD student at the Department of Communication Studies and Media Research at Ludwig-Maximilians-Universität Munich, Germany. His primary research interests are political and strategic communication, public diplomacy and conflict communication.

Marc Ziegele, PhD, is a postdoctoral researcher at the Department of Communication at the University of Mainz, Germany. His research interests concern online participation research, research on changes in journalism, online privacy behavior, analyses of electronic-word-of-mouth, and theory development within interactive online environments.