



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2013

Discussions in the comments section: Factors influencing participation and interactivity in online newspapers' reader comments

Weber, Patrick

Abstract: Posting comments on the news is one of the most popular forms of user participation in online newspapers, and there is great potential for public discourse that is associated with this form of user communication. However, this potential arises only when several users participate in commenting and when their communication becomes interactive. Based on an adaption of Galtung and Ruge's theory of newsworthiness, we hypothesized that a news article's news factors affect both participation levels and interactivity in a news item's comments section. The data from an online content analysis of political news are consistent with the hypotheses. This article explores the ways in which news factors affect participation levels and interactivity, and it discusses the theoretical, normative, and practical implications of those findings.

DOI: <https://doi.org/10.1177/1461444813495165>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-89817>

Journal Article

Accepted Version

Originally published at:

Weber, Patrick (2013). Discussions in the comments section: Factors influencing participation and interactivity in online newspapers' reader comments. *New Media Society*:1-17.

DOI: <https://doi.org/10.1177/1461444813495165>

Discussions in the comments section: Factors influencing participation and interactivity in online newspapers' reader comments

Patrick Weber

University of Zurich, CH

Abstract

Posting comments on the news is one of the most popular forms of user participation in online newspapers, and there is great potential for public discourse that is associated with this form of user communication. However, this potential arises only when several users participate in commenting and when their communication becomes interactive. Based on an adaption of the theory of newsworthiness (Galtung and Ruge, 1965), we hypothesized that a news article's news factors affect both participation levels and interactivity in a news item's comments section. The data from an online content analysis of political news are consistent with the hypotheses. This article explores the ways in which news factors affect participation levels and interactivity, and it discusses the theoretical, normative, and practical implications of those findings.

Keywords

online discussion value, user-generated content, user comments, interactivity, news factors, participatory journalism, participation, public sphere

Corresponding author

Patrick Weber, Institute of Mass Communication and Media Research, University of Zurich, Andreasstrasse 15, 8050 Zurich, CH. Email: p.weber@ipmz.uzh.ch

Accepted for publication in *New Media & Society*. Please cite as:

Weber, P. (in press). Discussions in the comments section: Factors influencing participation and interactivity in online newspapers' reader comments. *New Media & Society*.

Introduction

Based on the implementation of Web2.0 technologies in online newspapers, we are witnessing an increase in user-generated content in online news that supplements professionally generated journalistic content. There are a variety of website features that enable user participation, with commentary and debate spaces currently being the most widely offered participation opportunities in online newspapers (Domingo et al., 2008). Commenting on the news is the most common form of participation in contemporary news use, and it is certainly one of the most common forms of citizen engagement online (Emmer et al., 2011).

Reader comments are especially interesting from the normative perspective of deliberative digital democracy, which revolves around the question of whether internet-based interactive technologies support 'the extension of a deliberative democratic public sphere of rational communication and public opinion formation' (Dahlberg, 2011: 589). Online mass media such as news websites are particularly important forums in the public sphere because they have the ability to communicate collectively relevant issues to large audiences and to facilitate the formation of public opinion (Gerhards and Schäfer, 2010). As in traditional offline mass media, the processes of selecting information, evaluations and views, and synthesizing and representing them in publicly accessible content and thereby shaping public opinion has generally been the domain of professional communicators. However, by providing commentary and debate spaces,

online newspapers create the opportunity for active communication that is easy and accessible for ordinary users in these important forums of the public sphere. This opportunity opens up prospects for the visible and public discursive processing of news issues by readers. However, the potentials for quality discourse emerge only when a substantial amount of users participate in commenting on a news item and when there is a certain degree of interactivity among the users' comments, i.e. the degree to which users' comments refer to prior message transmissions (Kioussis, 2002). Therefore, the general questions pursued here address the ways in which news content might affect participation levels and interactivity in an article's comments section. From the perspective of digital deliberative democracy, this question is particularly relevant for political news, and we seek to discover which features of a political news event trigger or limit participation and interactivity.

Literature review: Reader comments in online newspapers

The bulk of the research on reader comments is conducted in studies on participatory journalism where comments are treated as one of several modes of participation. In general, this research focuses on the appraisals of user participation by journalists, the adoption of participation features on news websites, the management of user-generated content and the reasons for the adoption of different management strategies (Diakopoulos and Naaman, 2011b, Nielsen, 2012 and Reich, 2011 explicitly address comments). Other areas of interest are the content of reader comments (e.g.

Diakopoulos and Naaman, 2011a; Douai and Nofal, 2012; McCluskey and Hmielowski, 2012) and the influence of reader comments on users' perceptions and judgments (e.g. Lee and Jang, 2010). Additionally, information about the readers and writers of comments is beginning to accumulate (Bergström, 2008; Diakopoulos and Naaman, 2011b; Emmer et al., 2011; Hujanen and Pietikäinen, 2004; Springer, 2011), and the research explores the influence of different comment management strategies on commenting (Weber, 2013).

The research on the influence of journalistic content on reader comments is scarce, but it has already identified certain news content features that affect the amount of commentary. Tsagkias et al. (2009) found that the exclusiveness of topics, persons and places covered in a news article, as well as the use of distinctive and well-known names for these entities, predict the amount of comments more effectively than formal features such as article length. Focusing on story sentiment, Diakopoulos and Naaman (2011a) found that articles dealing with positive topics tend to have fewer comments per article. Boczkowski and Mitchelstein (2011) explored the role of topics and showed that the most-commented articles tend to focus on public affairs articles (in particular politics) and that this effect is intensified during periods of heightened political activity. Weber (2012a) showed that news events with a high social impact and a determinable temporal structure, and if they are controversial, negative and involve the own nation, attract more comments, whereas an unexpected news event, especially one that is

covered using a highly factual presentation style, will decrease the amount of commentary. Furthermore, in international news, the status and proximity of a covered country will influence the number of reader comments per article (Weber, 2012b).

The present paper extends this line of research by taking a more discriminating view of reader comments. The focus of extant research is on explaining the aggregate number of reader comments that, in principle, results from a) the number of readers participating in commenting and b) the interactivity among reader comments, or rather, the number of comments posted by each participating reader. The aggregate number of comments increases with both measures, and thus far, it is unknown by which mechanism – increased participation or interactivity – journalistic content influences the amount of commentary. Both mechanisms could be equally, but also differently, affected by the same content features and, moreover, individual characteristics of journalistic content could only affect one of them. To provide a theoretical basis for the exploration of these relationships, in the following section, we will provide a comprehensive account of relevant news content features and relate them to the two measures.

News factors and user comments

The theory of newsworthiness (Galtung and Ruge, 1965) provides a comprehensive set of news characteristics. Based on the principles of the psychology of perception, Galtung and Ruge (1965) derived the qualities of events that influence the

likelihood for them to become news. They proposed 12 of these qualities, referred to as news factors, among them *frequency* (temporal structure of an event), *meaningfulness* (as indicated by the proximity and impact of a news event), *unexpectedness*, *continuity* (the relation of an event to established issues), *power* and *influence*, reference to (often prominent) *persons*, and *negativity*, as indicated by damage or conflict. In the literature, there have been several differentiations of this taxonomy adhering to the core concepts (Eilders, 2006).

Galtung and Ruge (1965) assumed that news factors operate in the whole news system, serving both as journalistic selection criteria (in the decision to publish a news item about an event) and as criteria that influence audience response to these news items (ranging from selective exposure to the selective memory of news). They propose a positive influence of each news factor on selection for the first stages of news diffusion (i.e. journalistic selection), an additive impact of news factors on selection, as well as the ways in which news factors are replicated and accentuated in each representation of the news event in the process of news diffusion. Today, there is ample empirical evidence of the influence of news factors on journalistic selection as well as on the audience's selective media use and selective retention of news items (cf. Eilders, 2006).

So why investigate the influence of the news factors on participation levels and interactivity? An established finding of the research on the theory of newsworthiness is that news factors do influence journalistic selection and shape the presentation of news

events (Eilders, 2006). In this respect, news factors are typical content features of news. Furthermore, the literature offers a comprehensive catalogue of news features and has already found systematic relationships between news factors and the overall number of comments. Finally, in the discourse on media performance, some factors are considered to be criteria for the (societal) relevance of media content (McQuail, 1999: 197ff), and, from the perspective of deliberative digital democracy, it is of value to determine if a higher amount of participation and interactivity in readers' comments can be expected with important political issues.

With regard to the amount of participation, a hypothesis on the influence of the news factors will be derived from following assumptions. These assumptions are grounded in Eilders' (2006) conceptualization of news factors as collective relevance indicators, i.e. as perceived content features that *concurrently* lead people to assign relevance to a reported event (Shoemaker and Cohen, 2006 provide a similar approach). According to this approach, news factors influence both the allocation of cognitive resources (i.e. focal attention) and the activation of existing knowledge during the reception of news, and they thereby generate relevance assignment for three reasons: First, it is assumed that the allocation of attention and processing resources on certain stimuli became automatic through the course of evolution because it was associated with a relative increase in fitness. Hence, the automatic assignment of relevance to events that are associated with *damage* or *conflict*, are *unexpected* and have *factual*

consequences is a result of biological evolution. Second, with respect to socialization it is assumed that people assign relevance based on their social identity and focus their attention on events that are significant for their role in society or events that are connected to collective values and norms. Therefore, relevance is assigned to events that are related to the own *national or social group*, events where *powerful* persons or institutions are involved or that have *positive or negative consequences* on the societal level. Finally, it is assumed that events that are related to a user's prior knowledge and experience can be contextualized and processed more fluently, which affects relevance attribution. Hence, for events that are associated with issues that have already been *established* in preceding news coverage, the assignment of relevance is much more likely.

Based on this conceptualization of news factors, we assume that perceived relevance is a prerequisite for commenting. Only news that is deemed relevant and involves the user will increase the willingness to comment on it (Diakopoulos and Naaman, 2011b; Springer, 2011). This assumption is also supported by research showing that involvement increases the willingness to express one's opinion (Lasorsa, 1991; Shamir, 1997). As *collective* relevance indicators, news factors lead people to assign relevance concurrently and involve users beyond their individual interests. Therefore, they should encourage more users to voice their opinions and thereby influence the amount of participation in commenting on the aggregate level.

The second assumption is that existing knowledge and experience (including having an opinion on an issue) are crucial to the motivation to comment (Diakopoulos and Naaman, 2011b; Springer, 2011). Because there is a systematic relationship between some news factors and the likelihood of prior knowledge from previous news reporting, these news factors should influence the willingness to comment and will therefore affect the overall amount of participation. This proposition finds support in research showing that knowledge and media use influence the willingness to speak publicly (Scheufele and Eveland, 2001; Shamir, 1997). Therefore, we propose the following hypothesis:

H1: The news factors of a news item influence the level of participation in commenting in an article's comments section.

A non-directional hypothesis is proposed here for good reason. Beyond the assumptions that can be made with regard to a general relationship between news factors and participation levels, a positive relationship cannot always be reasonably assumed. For example, according to the conceptualization outlined above, *unexpected* events should be considered to be more relevant due to the automatic assignment of relevance, which contributes to the likelihood of a positive relationship. On the other hand, with unexpected events, users' knowledge of and prior experience with the event should be too low (due to a lack of prior news reporting) to provide additional information on an issue. This situation makes a negative relationship between

unexpectedness and the level of participation equally likely. Therefore, the following research question is posed to explore the impact of each factor:

RQ1: How do individual news factors influence the level of participation in commenting in an article's comments section?

To derive a hypothesis about the influence of news factors on the interactivity of reader comments, we refer to the same assumptions explicated above: Posting more than one comment requires the user to keep track of the development of an issue and the other users' comments. The reader must revisit the news site, re-read the comments and take the time to prepare and post another comment. Without a certain degree of involvement with the news issue, and without having some knowledge about the issue, it seems unlikely that a reader will make such an effort. Because news factors are supposed to influence involvement and to be related to the likelihood of prior knowledge, they might also affect readers' continued commenting and, by extension, the number of times a user posts comments in an article's comments section.

H2: The news factors of a news item influence interactivity in an article's comments section.

Again, the current state of research does not allow for deriving directional hypotheses for single news factors, and therefore the following research question is posed to explore the effects of single factors:

RQ2: How do single news factors influence interactivity in an article's comments section?

Methods

For this study, the hypotheses' testing is based on a content analysis of political news in three online newspapers.

Sampling

The media sample was selected to satisfy two criteria. 1) The selected news websites must offer article-related comment functionality. 2) The selected news websites must differ in their comments management strategies, so that respective differences can be controlled for. The media sampling frame consisted of a ranking of the most important German websites based on how much of the German population they reach. The ranking is based on technical measurements on the websites and a representative telephone survey (AGOF, 2008). This list was reduced to online newspapers because they are the most important websites for German consumers of online news (van Eimeren and Frees, 2008). From that list, the three highest ranking online newspapers that satisfied the criteria were selected: *Welt.de* (Website A, 3.72 million unique users per average week; no registration required for commenting, unmoderated comments sections), *FOCUS online* (Website B, 3.06m unique users; no registration required, pre-publication assessment of the comments by moderators) and

sueddeutsche.de (Website C, 2.16m unique users; registration required, post-publication moderation of the comments sections).

The time frame for sampling was October and November 2008. For each month, one week was constructed to randomly determine the access day. For each access day, one access time was determined at random.

At each of the 14 access times, software with an implemented web crawler was used to access the website and to produce offline copies. The units to sample were news items, and their relevant characteristics were news factors, features of the related comments and the degree of prominence with which the news item was linked on the website at the time of access. The news items were culled from the contents of single web pages, as was the information on the comments. Indicators of how prominently a news item was linked were the contents of a website's homepage and the overview page of the politics section. Therefore, the homepage and the politics section overview page were accessed and downloaded by the web crawler. A sampling frame for all pages containing news articles was obtained by letting the web crawler explore the websites to a depth of one link starting from these two pages. The resultant article sampling frame consisted of 3753 pages and provided the basis for the coding process. Starting from these pages, the web crawler was allowed to explore the websites again to an infinite depth, but only within the boundaries of each page's server-side subdirectory, and store all linked pages.

A thousand pages were randomly selected for coding from the article sampling frame. Nine hundred and forty of them were randomly allotted to 20 coders. The remaining 60 were given to all coders to enable reliability testing. Reliability was measured by computing average pairwise intercoder agreements (PA) using Holsti's method and Cohen's *kappa* (κ ; cf. Neuendorf, 2002).

Variables

Four groups of variables were coded: Firstly, properties of each page: The *website* of origin (PA = 1, κ = 1) and the *page type* (PA = .96, κ = .92). For each article, the *time of publication* was coded by recording year, month, day, hour, and minute separately (average PA = .96, κ = .95).

Secondly, a news item's teaser characteristics on the homepage (HP) and the politics section overview page (PSOP) were coded. Coders determined whether there was a *news teaser*, and, if so, coded the following teaser characteristics: *placement* (main content area of the page vs. outside main content area; PA_{HP} = .95, κ _{HP} = .76, PA_{PSOP} = .98, κ _{PSOP} = .66), the presence and relative size of a *picture/illustration* (PA_{HP} = .95, κ _{HP} = .75, PA_{PSOP} = .85, κ _{PSOP} = .67), amount of *text* (detailed text vs. single-line teaser; PA_{HP} = .95, κ _{HP} = .75, PA_{PSOP} = .87, κ _{PSOP} = .74) and accentuation of the *headline* (highlighted by contrasting color or typeface vs. not highlighted; PA_{HP} = .96, κ _{HP} = .75, PA_{PSOP} = .87, κ _{PSOP} = .74). These teaser features are known to influence

users' selection of news in online newspapers (Knobloch et al., 2003; Zillmann et al., 2004), and they are therefore central control variables.

Thirdly, news factors were coded for every news item. The *proximity* of a news event was defined as the involvement of the own nation and was measured based on the country where the event took place (PA = .90, κ = .89) and the country that was mostly affected by the event (PA = .91, κ = .89). Events taking place outside Germany and not involving Germany have the lowest proximity. Events located in Germany and only affecting Germany have the highest proximity. The variable *frequency* coded whether the reported issue had a determinable beginning and end or not (PA = .76, κ = .53). *Prominence* denotes reference to well-known persons (PA = .83, κ = .46). *Power* was defined as the sphere of political influence of the persons or institutions associated with the event (local, regional, national, international; PA = .73, κ = .42). *Personalization* was coded when the news story was framed in terms of individual experiences (PA = .76, κ = .34). *Controversy* denotes the level of conflict (in communication at most vs. non-violent conflict vs. violent conflict; PA = .78, κ = .56). An event was coded as *unexpected* when it was explicitly stated that the event was unexpected (PA = .71, κ = .27) and it was coded as having an *impact* when it had consequences for one or more social categories or groups (PA = .68, κ = .38). *Damage* denotes negative (PA = .71, κ = .47) and *success* positive consequences (PA = .80, κ = .44) of an event. The level of a news item's *facticity* was determined by coding whether the article provided mainly

analysis and interpretation (low facticity) or reported concrete actions and events (high facticity, $PA = .77$, $\kappa = .41$). *Continuity* was defined as reference to established issues. To enable a respective coding decision, a list of news issues for each week between 01 September 2008 and 30 November 2008 was compiled using the archives of the news websites. For each issue that was covered on all three websites, the date of introduction into the media agenda was listed. On the basis of that list, the coders were able to decide if a news event belonged to an issue that was on the list (i.e., established) or not ($PA = .74$, $\kappa = .41$).

Fourthly, the *number of comments* an article received was recorded per news item ($PA = .99$, $\kappa = .99$), and, for each comment, the *name of the comment's author* was recorded ($PA = .93$).¹ Note that the coding unit for this variable was the individual comment.

The reliability tests show that the PAs exceed .7, which is the critical value in content analysis (Frey et al., 2000), for nearly all of the variables coded. *Kappa* is a highly conservative measure and, as expected, the values are lower. However, all values are above zero, which indicates systematic above chance agreement for all variables. Moreover, nearly all *kappas* were above .40 what indicates at least fair to good agreement beyond chance (cf. Neuendorf, 2002: 143; Fleiss, 1981: 218). Also note that most of the news factors have only few codes. In that case, also lower *kappas* indicate substantially high agreement (Bakeman et al., 1997).

Indices

The study's dependent variables were computed from the coded comment features. The *level of participation* in commenting is measured by the number of individual commentators that posted a comment in an article's comments section. This measure was obtained by aggregating the comments per article and counting the number of individual author names. Irrespective of how many comments they posted, this index counts the number of readers that participated in commenting. To determine the degree of *interactivity* in an article's comments section, we applied an operationalization used to measure interactive discussions in online forums (e.g. Birch and Weitkamp, 2010): the average number of comments individual users posted on an article (which is the quotient of the number of comments per news item divided by the number of individual commentators). This index captures the individual commentators' continued comments, it measures how much commenting extends beyond single postings, and it fits the purpose of the study to separately analyze the mechanisms that drive the overall number of comments per article.

The coded teaser characteristics of a news item were submitted to a principal component analysis ($KMO = .809$, Bartlett $\chi^2 = 10468.8$, $df = 28$, $p < .001$) that yielded two components with eigenvalues greater than one that account for 88.7% of the variance. There is a clear factor structure (main loadings $\geq .71$, all cross loadings $< .3$). The first factor comprises the coded HP-teaser characteristics and the second the

characteristics of the PSOP-teasers. Standardized factor scores were computed to serve as indices of the prominence of the linking of a news item on the homepage (*HP-prominence*, Cronbach's $\alpha = .97$) and the politics section overview page (*PSOP-prominence*, $\alpha = .92$).

Finally, the time lag between access and the time of publication of a news item was computed for each article as a measure of *exposure* time.

Analysis methods

The data were analyzed using the regression approach with the study's dependent variables as outcomes and news factors and control variables (*website*, *HP-prominence*, *PSOP-prominence*) as covariates. All regression models included a constant and *exposure*. The model for the interactivity outcome additionally controlled for the amount of participation.

With the properties of our outcomes (non-normally distributed, highly skewed, and with a limited range of outcomes, see results section), the traditional linear regression model using the ordinary least squares method (OLS) for parameter estimation is not appropriate (Hardin and Hilbe, 2007: 17). The variable *level of participation* represents the number of individual commentators and is thus a count variable. OLS-Regression analysis of count data typically violates central assumptions, resulting in inefficient estimators and invalid *p*-values of the regression coefficients (cf. Cameron and Trivedi, 1998). The negative binomial regression model using maximum

likelihood procedures (ML) for parameter estimation is a convenient standard method that accounts for the properties of empirical count data (Cameron and Trivedi, 1998: 60) and is thus used for hypothesis testing.

Similar problems arise when conducting OLS-regressions with the study's second dependent variable, *interactivity*, due to its distributional characteristics. Within the framework of generalized linear models (GLzM, see Hardin and Hilbe, 2007), regression models can easily be adjusted to such data situations. When faced with nonnegative responses that are greater than zero, have a high initial peak, a rapid drop and a long right tail (as is the case with the *interactivity* outcome), the gamma distribution is preferred to model the outcome (Hardin and Hilbe, 2007: 89ff). Therefore, hypothesis testing was based on a gamma model using ML-estimation.

The hypotheses were tested by comparing the fit of different models for each of the dependent variables. Starting from a baseline model that only includes the control variables, the news factors were added to the model. Both the baseline and the news factors model were evaluated by a Wald test.² The hypotheses were tested by assessing improvements in model fit in the news factors over the baseline models using likelihood ratio tests (LR-tests) between nested models (cf. Long and Freese, 2006: 99ff). Additionally, a comparison between the Bayesian information criteria (BIC) was conducted to assess the strength of the evidence for the news factors model based on the guidelines suggested by Raftery (1996).

Results

The coding yielded complete data for $n = 558$ news items. On average, 22.12 users participated in commenting on a news item ($SD = 25.78$), 10.75% of the news articles remained uncommented upon, 50% received comments from at most 13 readers, and the maximum number of users participating in commenting on a single news article was 152. To test whether the news factors of an article contribute to an explanation of this high variability in the amount of participation, we first estimated the baseline model for this outcome (Wald $\chi^2 = 350.19$, $df = 4$, $p < .01$ for the control variables). Next, the news factors were added to the baseline model and the Wald test indicates that they have an additional influence (Wald $\chi^2 = 156.93$, $df = 12$, $p < .01$). Consistent with H1, an LR-test comparing the baseline model and the news factors model shows that adding the news factors resulted in a highly significant increase in model fit (LR $\chi^2 = 145.45$, $df = 12$, $p < .01$). A difference of 69.55 in the BICs provides very strong support for the news factors model, implying that the news factors of a news item do affect the number of users who participate in commenting in an article's comments section.

Regarding the $n = 498$ news items that received at least one comment, 1.16 comments per user were posted on average ($SD = 0.28$). In 40.56% of these articles, the participating readers posted just one comment, and in over 60% of the cases, the number of comments per user did not exceed an average of 1.1. This finding indicates that the degree of interactivity in most articles' comments sections remains rather low.

An extremely high interactive commenting with an average number of comments per participating user of two or more occurred in less than 2% of the articles' comments sections and thus seems to be an exception. To test whether the news factors of an article contribute to an explanation of the degree of interactivity, we first estimated a baseline model for interactivity as the outcome (Wald $\chi^2 = 3573.92$, $df = 5$, $p < .01$ for the control variables).³ Next, the news factors were added and the Wald test indicates that they have an additional influence (Wald $\chi^2 = 254.77$, $df = 12$, $p < .01$). An LR-test shows that adding the news factors resulted in a highly significant increase in model fit (LR $\chi^2 = 80.13$, $df = 12$, $p < .01$), and, consistent with H2, an absolute difference of 5.60 in the BICs provides positive support for the news factors model. We conclude that an article's news factors indeed do affect interactivity in the related comments section.

The research questions were designed to assess the singular contribution of each news factor in a multivariate context. Table 1 shows the results of both regression analyses and lists the hypothesis test for each covariate.

[Table 1 here]

The factor *proximity* positively affects both the amount of participation and the degree of interactivity in an article's comments section. The more a news event focuses on the own nation, the more readers comment on it and the more frequently the individual commentators post a comment on the same story. This effect implies that participation and interactivity can be expected to be much higher with national than

with international news. Three additional news factors affect both outcomes: News events that have a manifest *impact* on a defined social group or category are commented on by more users and commenting is more interactive than with events that have no consequences or affect only single individuals. *Facticity* has a significant negative effect. Providing a mere factual report on an event without further interpretation or situational analysis decreases the amount of participation as well as interactivity. Additionally, *continuity* proves to be an influential factor. News articles that continue issues that are already on the media agenda receive comments from more users (but note that there is an only marginally significant relationship), and the participating users tend to post more than one comment.

According to our analyses, neither outcome is significantly related to *prominence*, *personalization*, *unexpectedness* or *success*.

Frequency is a significant predictor of the number of readers who post comments but not of interactivity: In the case of news events having a determinable beginning and end, the level of participation is higher than with reports of long-term or unfinished issues.

By contrast, the remaining three factors affect only interactivity. *Power* decreases interactivity. On issues that are associated with highly powerful persons or institutions (e.g. heads of state, national parties or international organizations), participating users tend to post fewer comments, whereas commenting becomes more interactive when a

news issue features institutions whose sphere of influence is restricted to a municipal or regional level (e.g. citizens' groups or federal state parliaments). For the factor *controversy*, the model estimates a marginally significant positive effect, implying that with increasing conflict levels reported in news items ranging from a clash of opinions to violent conflict, commenting can be expected to increase in interactivity. The same is true for the factor *damage*. Reporting of negative events including those that involve physical damage or even bodily harm positively affects readers' continued commenting.

Table 1 also shows that all control variables exert significant influences on both outcomes. Firstly, the negative coefficients for the websites suggest that the less liberal a website's comment management strategy, the lower are the participation and interactivity levels in the comments sections. Both pre-publication moderation (Website B) and the requirement of registration plus post-publication moderation (Website C) reduce participation and decrease interactivity. Secondly, both outcomes are influenced by how prominently a news item is linked. An attention-grabbing placement and presentation of a news item's teaser on the homepage or the politics section overview page is associated with higher levels of participation and increased interactivity.

Discussion

Posting comments on news articles is currently one of the most popular forms of user participation in online-content generation. From the perspective of deliberative digital democracy (Dahlberg, 2011), there is great potential for public discourse

associated with this form of user communication. However, in the context of online newspapers, this potential only emerges when a number of users participate in commenting and when users repeatedly post comments to the point at which communication in the article's comments section becomes interactive. Extant research that identified features of journalistic content affecting user comments exclusively focused on the quantity of commentary. Because that outcome results from a) the number of users commenting and b) the number of comments posted per user, this paper disentangled those issues and investigated how the content of a news item might affect a) the amount of participation and b) the level of interactivity in an article's comments section.

Contributions

To advance knowledge on the interplay between professional journalistic content and reader comments, a comprehensive account of news content characteristics was used (Galtung and Ruge, 1965), and news factors were hypothesized to affect both the amount of participation and interactivity. Consistent with these hypotheses, the study found that news factors contribute to explaining the number of readers who post at least one comment and to explaining the average number of comments posted per commentator.

Four factors were related to both the amount of participation and interactivity. Based on the conceptualization of news factors as collective relevance indicators

(Eilders, 2006) and on the assumption that involvement is a prerequisite for commenting, these effects can be interpreted as follows. The effects of *proximity* and *impact* are assumed to result from socialization: News events pertaining to the own social group (be it based on nationality or a more specific social attribute) are more likely to appeal to the social identity of a user and are therefore inherently more involving, increase the willingness to comment in a higher proportion of readers and thereby increase participation levels overall. Moreover, these news factors seem to motivate readers to accept the costs of repeated commenting, which in turn increases interactivity. The positive effect of *continuity* can be interpreted by referring to the assumption about the role of knowledge: For news events that perpetuate an issue that is already on the media agenda, it is much more likely that readers have some knowledge about the topic from prior news and have probably developed an opinion. Because such knowledge motivates readers to comment (Diakopoulos and Naaman, 2011b), a positive impact on participation levels and interactivity is observed on the aggregate level. Similarly, articles offering analysis and journalistic appraisals of a news event (and thereby being low in *facticity*) can easily create dissonance on the part of the user who might have a different understanding of the issue at hand (Springer, 2011). Therefore, on the aggregate level, a negative relationship is observed.

Our analyses showed that the temporal dimension of a news event (*frequency*) only affects participation levels. We presume that for news events that have a clearly

determinable beginning and end (high *frequency*), knowledge might be less relevant for an understanding of the event because it provides a self-contained meaning and thus allows more users to express their views – irrespective of actual knowledge levels.

Power, controversy and *damage* are related solely to interactivity. News stories featuring a potentially damaging or controversial event receive more attention (i.e. they are allocated more processing resources, Eilders, 2006), and they are thus more involving, what also motivates repeated commenting. A rather unexpected finding is the negative effect of *power*. The emergence of interactive discussions is more likely with political events of limited reach. We suppose that such events are more likely to be rooted in people's everyday lives, meaning that people probably have direct experience with the issue. Therefore, such news events are more involving and thus have a higher discussion value.

Altogether, the current study complements existing research by showing that news factors influence the overall amount of reader comments on an article by affecting both participation levels and interactivity and by clarifying the various ways in which news factors affect the range of mechanisms.

Implications

The theory of newsworthiness as proposed by Galtung & Ruge (1965) covers all of the processes of traditional news diffusion from real world events to the formation of a personal image. The opportunities for participatory news use offered by Web2.0-

technologies make possible new stages of news diffusion initiated by users, e.g. the capacity to actively discuss news events in online environments. The results presented here suggest that the traditional model can be extended to such user-generated content, i.e. reader comments. News factors also influence selectivity on this new stage of news diffusion (selective participation) and affect the interactivity of users' communication. However, the results also imply that the traditional theory's hypotheses cannot simply be transferred to these processes. The negative influence of some news factors in particular suggests that the single news factors do not simply combine additively to affect this stage. Thus, the traditional selection and additivity hypothesis require a new specification for user-driven communication in the online-news context.

Although we cannot say anything about the quality of the users' discourse in the comments, from a normative point of view, it seems noteworthy that participation and the emergence of interactive discussions are influenced by news characteristics that are indicators of the (societal) relevance of media content (McQuail, 1999: 197ff). For example, a discursive treatment of political issues that affect a social group is desirable for the formation of public opinion. As indicated by the positive effect of the factor *impact* in our study, we find that the potential for quality discourse does indeed arise in online newspapers' comments sections when a political issue is social relevant.

The influence of *facticity* on both the level of participation and interactivity illustrates the crucial role that contextualization plays for initiating public discussions in

the comments section. Our results suggest that the perspectives on the news issues offered by journalists play a key role in making them accessible for a discursive treatment. This shows that even though online newspapers provide ordinary users with an opportunity to voice their views, professional news workers retain substantial power in channeling the active discourse in these forums of the public sphere.

The effect of *controversy* deserves some critical elaboration. One potential of the online public sphere is seen in its capability of invigorating rational discussions between people with opposing views (Wojcieszak and Mutz, 2009). If one assumes that controversial issues are likely to involve readers with opposing views, a positive effect of *controversy* could indicate the facilitation of discussions among those with different opinions. Our results suggest that comments sections do not contribute to an improved public sphere in this respect. This is, however, only a very tentative interpretation. One has to bear in mind that we coded higher *controversy* only for issues that involve conflicts beyond differences in opinions, including violent conflict such as war.

Seen from a normative point of view, the differences in the level of participation and interactivity between our sample's websites suggest that the potential for quality discourse is also determined by the features of the websites, in particular by a website's comment management strategy. This illustrates that an online newspaper's design features are important boundary conditions of deliberation (Wright and Street, 2007) and that a restrictive comment management might thwart the users' discourse in the

comments sections at a very basic level. However, unrestricted comments sections do not guarantee quality discourse. Moreover, unlike many private discussion forums, comments sections are embedded in editorial products for which the publishers bear responsibility. Therefore, to realize the potential for deliberation in these public forums, the challenge is to keep the hurdles to participation low and at the same time to implement a comment management strategy that facilitates quality discourse.

Finally, we briefly address some practical implications. By enabling user participation, news organizations seek to increase performance measures such as click rates or re-visits (Vujnovic, 2011). Our study revealed some of the characteristics in news content that engage readers in commenting, an activity that clearly drives performance measures. What should be of interest to practitioners is that some of these characteristics refer to the presentation style of news (*facticity*), while others refer to event characteristics that can be emphasized in order to optimize a news story for active audience participation.

Limitations and directions for future research

One limitation of the current study is the difficulty of determining the appropriateness of the interpretations of the effects found in the study. These explanations are principally open to empirical investigation. However, they are not accessible by content analysis. The limitations of this method were accepted in favor of the possibility it provided to simultaneously test for influences of a high number of

variables. Future research could be dedicated to the empirical assessment of the proposed explanations and the exploration of the relationships between content features, involvement and commenting behavior.

Furthermore, the decisions made during sampling set the limits for making generalizations. Strictly speaking, the results obtained here can be generalized to commenting on political news in the selected online newspapers during the sampled time period. Therefore, future research should include replications of the study based on longer time frames, with different sets of online newspapers and possibly a range of news that extends beyond political.

On a conceptual level, research on user comments in online newspapers can be extended in different ways. The current research focuses on three groups of variables that affect commenting: individual factors, media conditions and features of professionally generated journalistic content. Future research should also take into consideration another factor. Users not only react to the journalistic content with their comments, but they also react to the comments of other users. As a result, discussion structures emerge in an online newspaper's comment space. Future research should try to identify distinct patterns and typical dynamics in users' news discourse as well as the factors that affect it. Finally, more empirical assessments of the quality of the users' discourse in newspapers' comments sections would be helpful to assess whether it

makes a relevant contribution in the public sphere and to identify the factors that affect the quality of the discussions in the comments section.

¹ Computation of κ was not feasible for this variable because it has no numerical values but is a alphabetical character string. Because of the very high number of individual author names we did not assign them numerical values.

² The Wald test assesses the null hypothesis that the coefficients of the added variables are simultaneously equal to zero, implying that they would add no predictive value to the preceding model (cf. Long and Freese, 2006: 99ff).

³ For GLzM using a gamma family specification for the mean and variance function, the most common link functions are the reciprocal-, identity- and log-link (Hardin and Hilbe, 2007: 89ff). Using these specifications, we initially evaluated three news factor models for the interactivity outcome following the suggestions of Hardin & Hilbe (2007: 115). The specification with the identity link performed best (lowest deviance and lowest BIC) and is thus used for hypothesis testing.

References

AGOF (2008) *internet facts 2008-II*. Available at:

<http://www.agof.de/index.download.d55c96c3272b95a7a210cd9ce332ae49.pdf>.

Bakeman R, McArthur D, Quera V and Robinson BF (1997) Detecting sequential patterns and determining their reliability with fallible observers. *Psychological Methods* 2(4): 357–370.

Bergström A (2008) The reluctant audience: Online participation in the Swedish journalistic context. *Westminster Papers in Communication and Culture* 5(2): 60–79.

Birch H and Weitkamp E (2010) Podologues: conversations created by science podcasts. *New Media & Society* 12(6): 889–909.

Cameron AC and Trivedi PK (1998) *Regression analysis of count data*. Cambridge: Univ. Press.

Dahlberg L (2011) Re-constructing digital democracy. *New Media & Society* 16(6): 855–872.

Diakopoulos N and Naaman M (2011a) Topicality, time, and sentiment in online news comments. In: Tan D, Amershi S, Begole B, Kellogg WA and Tungare M (eds) *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems*. New York: ACM, 1405–1410.

- Diakopoulos N and Naaman M (2011b) Towards quality discourse in online news comments. In: Hinds PJ, Tang JC, Wang J, Bardram JE and Ducheneaut N (eds) *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*. New York: Association for Computing Machinery, 133–142.
- Domingo D, Quandt T, Heinonen A, Paulussen S, Singer JB and Vujnovic M (2008) Participatory Journalism Practices in the Media and beyond. *Journalism Practice* 2(3): 326–342.
- Douai A and Nofal HK (2012) Commenting in the online Arab public sphere. *Journal of Computer-Mediated Communication* 17(3): 266–282.
- Eilders C (2006) News factors and news decisions. *Communications. The European Journal of Communication Research* 31(1): 5–24.
- Emmer M, Vowe G, Wolling J and Seifert M (2011) *Bürger online*. Konstanz: UVK.
- Fleiss JL (1981): *Statistical methods for rates and proportions*. New York: Wiley.
- Frey LR, Botan CH and Kreps GL (2000) *Investigating communication*. Boston: Allyn and Bacon.
- Galtung J and Ruge M (1965) The structure of foreign news. *Journal of Peace Research* 2(1): 64–91.
- Gerhards J and Schäfer MS (2010) Is the internet a better public sphere? *New Media & Society* 12(1): 143–160.

- Hardin JW and Hilbe JM (2007) *Generalized linear models and extensions*. College Station, TX: Stata Press Publ.
- Hujanen J and Pietikäinen S (2004) Interactive uses of journalism. *New Media & Society* 6(3): 383–401.
- Kiousis S (2002) Interactivity: a concept explication. *New Media & Society* 4(3): 355–383.
- Knobloch S, Hastall M, Zillmann, D, and Callison C (2003) Imagery effects on the selective reading of internet newsmagazines. *Communication Research* 30(1): 3–29.
- Lasorsa DL (1991) Political Outspokenness: Factors Working Against the Spiral of Silence. *Journalism Quarterly* 68(1/2): 131–140.
- Lee E and Jang YJ (2010) What do others' reactions to news on internet portal sites tell us? *Communication Research* 37(6): 825–846.
- Long JS and Freese J (2006) *Regression models for categorical dependent variables using Stata*. College Station, Tex.: Stata Press.
- McCluskey M and Hmielowski J (2012) Opinion expression during social conflict. *Journalism* 13(3): 303–319.
- McQuail D (1999) *Media performance*. London: Sage.
- Neuendorf KA (2002) *The content analysis guidebook*. Thousand Oaks: Sage.
- Nielsen C (2012) Newspaper journalists support online comments. *Newspaper Research Journal* 33(1): 86–100.

- Raftery AE (1996) Bayesian model selection in social research. *Sociological Methodology* 25(1): 111–163.
- Reich Z (2011) User comments. The transformation of participatory space. In: Singer JB, et al. (eds) *Participatory journalism*. West Sussex: Wiley-Blackwell, 96–117.
- Scheufele DA and Eveland WP (2001) Perceptions of 'public opinion' and 'public' opinion expression. *International Journal of Public Opinion Research* 13(1): 25–44.
- Shamir J (1997) Speaking up and silencing out in face of a changing climate of opinion. *Journalism & Mass Communication Quarterly* 74(3): 602–614.
- Shoemaker PJ and Cohen, AA (2006) *News around the world*. New York: Routledge.
- Springer N (2011) Suche Meinung, biete Dialog?. In: Wolling J, Will A and Schumann C (eds) *Medieninnovationen..* Konstanz: UVK, 247–264.
- Tsagkias M, Weerkamp W and Rijke M de (2009) Predicting the volume of comments on online news stories. In: Cheung D (ed.) *Proceedings of the ACM Eighteenth International Conference on Information and Knowledge Management, Hong Kong, China*. New York: ACM, 1765–1768.
- van Eimeren B and Frees B (2008) Internetverbreitung. *Media Perspektiven* (7): 330–344.
- Vujnovic M (2011) Participatory journalism in the marketplace. In: Singer JB, et al. (eds) *Participatory journalism*. West Sussex: Wiley-Blackwell, 139–154.

- Weber P (2012a) Der virtuelle Stammtisch. Determinanten interpersonal-öffentlicher Kommunikation auf Nachrichtenwebsites. In: Ockenfeld M, Peters I and Weller K (eds) *Social Media und Web Science*. Frankfurt: DGI, 457–459.
- Weber P (2012b) Nachrichtenfaktoren & User Generated Content: Die Bedeutung von Nachrichtenfaktoren für Kommentierungen der politischen Berichterstattung auf Nachrichtenwebsites. *Medien & Kommunikationswissenschaft* 60(2): 218–239.
- Weber P (2013) Nachrichtenfaktoren und Kommentierungsintensität von Online-Nachrichten. In: Pfetsch B, Greyer J and Trebbe J (eds) *MediaPolis - Kommunikation zwischen Boulevard und Parlament*. Konstanz: UVK, 229–245.
- Wright S and Street J (2007) Democracy, deliberation and design: The case of online discussion forums. *New Media & Society* 9(5): 849–869.
- Wojcieszak ME and Mutz, DC (2009) Online groups and political discourse. *Journal of Communication* 59 (1): 40–56.
- Zillmann D, Chen L, Knobloch S and Callison C (2004) Effects of lead framing on selective exposure to internet news reports. *Communication Research* 31(1): 58–81.

Table 1

Coefficients estimated in the regression analyses (SEs)

Variable	Level of participation	Interactivity
Proximity	0.43** (0.06)	0.45** (0.07)
Frequency	0.57** (0.14)	0.17 (0.17)
Power	0.07 (0.09)	-0.76** (0.18)
Prominence	-0.06 (0.23)	0.18 (0.26)
Personalization	0.03 (0.20)	0.14 (0.18)
Controversy	0.14 (0.12)	0.20# (0.10)
Unexpectedness	-0.29 (0.19)	-0.23 (0.30)
Impact	0.44** (0.17)	1.46** (0.22)
Success	-0.02 (0.15)	0.08 (0.18)
Damage	0.18 (0.13)	0.49** (0.13)
Facticity	-1.07** (0.21)	-1.35** (0.22)
Continuity	0.30# (0.16)	0.48* (0.21)
Website B	-0.57** (0.19)	-1.60** (0.33)
Website C	-2.37** (0.19)	-1.51** (0.37)
HP-prominence	0.63** (0.08)	0.49** (0.08)
PSOP-prominence	0.47** (0.07)	0.58** (0.08)

Level of participation		-0.01** (0.00)
Constant	2.97** (0.38)	4.88** (0.50)

Note. The model for *level of participation* was estimated by a negative binomial regression model (n = 558, dispersion parameter $\alpha = 2.36$, $SE = 0.13$). The model for *interactivity* was estimated by a gamma regression (link = identity, n = 498, scale parameter = 0.60).

** p<.01, * p<.05, # p<.10