RESEARCH-IN-BRIEF

Commenting quality
Effects of user comments on perceptions of journalistic quality

Qualität kommentieren
Die Wirkung von Nutzerkommentaren auf die Wahrnehmung journalistischer Qualität

Anna Sophie Kümpel & Nina Springer
Anna Sophie Kümpel, Institut für Kommunikationswissenschaft und Medienforschung, Ludwig-Maximilians-Universität München, Oettingenstr. 67, 80538 München; Kontakt: kuempel(at)ifkw.lmu.de

Nina Springer, Institut für Kommunikationswissenschaft und Medienforschung, Ludwig-Maximilians-Universität München, Oettingenstr. 67, 80538 München; Kontakt: nina.springer(at)ifkw.lmu.de
Commenting quality
Effects of user comments on perceptions of journalistic quality

Anna Sophie Kümpel & Nina Springer

Abstract: The (participatory) opportunities provided by interactive features on news sites have been widely theorized and investigated in recent years. However, user comments’ effects on the perception of journalistic quality have only begun to be examined. To investigate those effects, the present study deployed two 2 × 2 experiments with a between-subject design, thereby exposing participants (N = 224) to a constructed online news article (covering a potential military intervention of the German armed forces against ISIS) and corresponding user comments. Comments varied in terms of (I) support for the issue described in the article as well as the (II) addressing of journalistic quality criteria (accuracy, impartiality). Results indicate that user comments indeed had considerable effects on readers’ quality assessments. If anonymous users praised the quality of the article in their comments, journalistic quality was perceived to be significantly higher. Implications for news media and media effects research are discussed.

Keywords: Journalistic quality, user comments, media effects, quantitative experiment.
1. Introduction

User comments are considered to be the most common form of user participation in online journalism (Hermida, 2011; Richardson & Stanyer, 2011), challenging our understanding of theorizing media effects in online environments (Springer, Engelmann, & Pfaffinger, 2015). Because news and user-generated information are nowadays inseparable, media effects research must take these interdependencies into account and focus on the wider embedding of media content instead of investigating it in isolation. The very same news article may have different effects on readers depending on the number of likes, shares, and – not least – the number and especially content and valence of user comments. This study examines the influence of user comments on readers’ assessments of journalistic quality. The relevance is twofold. First, if comments indeed affect quality assessments, this effect has implications for media images and, more generally, trust in media outlets. Second, these changes in media images and trust may affect selection decisions and thus determine whether readers are inclined to turn to certain media outlets after all (cf. Wolling, 2009). In this study, under a controlled experimental setting, we investigate how the positive and negative addressing of two journalistic quality criteria (accuracy respectively impartiality) as well as issue-related evaluations in user comments influence how people perceive the quality of a corresponding news article.

2. Effects of user comments on quality assessments

Scholars who studied user comments from a media effects perspective focused on different dependent variables, ranging from risk perceptions (Anderson, Brossard, Scheufele, Xenos, & Ladwig, 2014), third-person perceptions (Houston, Hansen, & Nisbett, 2011), and perceptions of media bias or the climate of opinion (Friemel & Dötsch, 2015; E.-J. Lee, 2012), to perceptions of credibility (J. Lee, 2014; J. Lee & Lim, 2014; Marchionni, 2014; Thorson, Vraga, & Ekdale, 2010). In general, all studies indicate that comments – at least for a short period of time – are able to shape attitudes and perceptions. However, to our knowledge, the effects on the assessments of journalistic quality have not been studied sufficiently. Experimental studies show that enabling comments does not lead readers to perceive journalism as more trustworthy (Marchionni, 2014), and that the sheer presence of comments can, on the contrary, even lead to a more negative assessment of journalistic quality (Ash, Hettinga, & Halpern, 2009; Prochazka, Weber, & Schweiger, 2016).

But which factors are relevant with respect to journalistic quality – and are readers actually able to detect it? Through comparison of studies that develop catalogues of journalistic quality criteria, we have identified five dimensions that can – albeit sometimes differently labelled – be found in all of them: relevance, comprehensibility, diversity, impartiality, and accuracy (e.g., Jungnickel, 2011; Schatz & Schulz, 1992; Urban & Schweiger, 2014; Wellbrock & Klein, 2014). In the present study, we decided to solely examine the effects of accuracy and impartiality.
due to a) the need for reduction in experimental research, b) the overall relevance of these criteria identified in previous studies on journalistic quality (e.g. Arnold, 2008; Wellbrock & Klein, 2014), and – most importantly – c) the fact that they were most often addressed in actual user comments on our chosen issue.1 Studies focusing on readers’ quality perceptions (e.g., Jungnickel, 2011; Urban & Schweiger, 2014) suggest that recipients are generally able to spot differences regarding these two quality criteria, although accuracy seems – due to the need for background knowledge – harder to evaluate than impartiality.

Since none of the existing experimental studies focusing on quality assessments actually varied how quality criteria were addressed in user comments, we want to test whether comments with explicit reference to journalistic quality are able to affect readers’ quality assessments. Studies investigating the influence of positively and negatively valenced evaluations in user comments generally suggest that a positive evaluation in user comments is reflected in participants’ assessment of the subject or issue in question. Lee and Lim (2014), for example, found that positively valenced comments about a political candidate led participants to perceive said candidate as more trustworthy than those participants that were exposed to negatively valenced comments (see also J. Lee, 2014). Thus, we assume:

H1: A news article that is accompanied by user comments that positively address the accuracy/impartiality of the article is judged to be of higher quality than a news article that is accompanied by user comments that negatively address the accuracy/impartiality of the article.

Besides, we are interested whether the sheer valence of comments addressing the issue of the article might influence quality assessments as well. We assume that the expression of consent or dissent in comments might induce a kind of spillover effect on perceptions of the article, simply through indicating (dis-)agreement. Due to people’s tendency to use heuristics when making judgments (Gigerenzer & Gaissmaier, 2011), the availability of valenced issue evaluations in the user comments might steer their quality assessments in the respective (positive or negative) direction. Hence, we propose:

H2: A news article that is accompanied by user comments that endorse the issue of the article is judged to be of higher quality than a news article that is accompanied by user comments that do not endorse (= criticize) the issue of the article.

3. Method

To investigate the effects of user comments on readers’ assessments of journalistic quality, we conducted two 2 × 2, post-test only experiments with a between-subject design. We exposed participants to a constructed online news article (same for all participants) as well as to corresponding user comments that varied in terms of (I) support for the issue described in the article (military intervention of the
German armed forces against ISIS) as well as the (II) addressing of journalistic quality criteria. Support was varied on two factor levels, differentiating between endorsement and refusal of the military intervention. The addressing of quality criteria was also varied on two factor levels, differentiating between praising (positive addressing) and criticizing (negative addressing). To keep the engagement time with the questionnaire as short as possible, one group of participants only saw comments focusing on accuracy (henceforth accuracy group), while the other group of participants only saw comments focusing on impartiality (henceforth impartiality group).

In the experimental procedure, participants were randomly assigned to the experimental conditions and exposed to the article that was supposedly published on Spiegel Online. The article was written as balanced as possible, focusing on arguments for both perspectives – support for/opposition against the intervention – in equal depth. On the next page, five user comments were presented (see Table 1). Two of them addressed the issue of the article, both either endorsing or refusing it. Another two of them addressed the article’s impartiality (resp. its accuracy), both either praising or criticizing it. In all groups, an irrelevant comment (addressing the availability of the article on Twitter) was added to increase external validity as well as to fit the actual website design in which five comments are displayed.

In December 2014, a total of 248 German participants took part in the two online experiments. They were recruited both by the authors, the participants in the seminar, and postings in Facebook university groups. The final sample consisted only of participants who had spent at least 15 seconds on both of the pages presenting the stimulus material (n = 224). Demographic characteristics of the sample included age (M = 25.09, SD = 9.57), gender (66.1% female), and educational level (34.8% with a university degree, 58.5% with a higher education entrance qualification and 6.7% with lower educational qualifications).

Adapting operationalization by Jungnickel (2011) and Urban and Schweiger (2014), we used two statements to assess both perceptions of accuracy (“The article does not contain erroneous information” and “The journalists’ statements are without contradictions.”) and impartiality (“The journalist reports in a neutral and unbiased manner” and “All perspectives are covered in equal depth.”) on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). Measures also included assessment of prior knowledge about the issue in question, Internet use and the use of commenting...
features (active, passive), as well as various sociodemographic variables.

As a treatment check, we asked participants to rate statements that described the content of the user comments on the same Likert scale. Perception of accuracy in comments was measured with the statement “If the comments addressed the quality of the article, they praised it as well researched and factually correct” (TC\textsubscript{A}) and perception of impartiality in comments with the statement “If the comments addressed the quality of the article, they praised it as comprehensive and impartial” (TC\textsubscript{IP}). As predicted, participants in the accuracy group who read comments that praised the articles’ accuracy agreed significantly more to TC\textsubscript{A} \((M = 4.21, SD = 0.87)\) than those who read comments that criticized its accuracy \((M = 1.28, SD = 0.64)\), \(t(106) = 20.03, p < .001\). Accordingly, participants in the impartiality group who read comments that praised the articles’ impartiality agreed significantly more to TC\textsubscript{IP} \((M = 4.19, SD = 0.78)\) than those who read comments that criticized its impartiality \((M = 1.31, SD = 0.54)\), \(t(104) = 21.98, p < .001\).4

---

4 We also conducted a treatment check to investigate whether participants perceived the issue endorsement in user comments as intended. We did so by assessing their agreement to the statement “If the comments addressed the military intervention against ISIS, they supported it”. As intended, participants who read the comments that endorsed the intervention agreed significantly more to the statement \((M = 3.79, SD = 0.84)\) than those who read the comments that opposed the intervention \((M = 1.71, SD = 0.82), t(212) = 18.269, p < .001\).
4. Results

To test our hypotheses, we conducted multifactorial analyses of covariance (ANCOVA). The dependent variables were participants’ accuracy and impartiality assessments; the factors were the addressing (positive, negative) of accuracy respectively impartiality and the endorsement of the issue (endorsing the issue, not endorsing the issue) in user comments. We controlled for participants’ gender, age, prior knowledge about the issue, usage of Spiegel Online, and their frequency of comment reading and posting.

In the accuracy group, a significant main effect of the addressing of accuracy on both accuracy (AA) and impartiality assessments (IA) could be observed, $F_{AA}(1, 102) = 38.89, p < .001, \eta_p^2 = .28; F_{IA}(1, 102) = 25.55, p < .001, \eta_p^2 = .20$ (see Table 2): Participants that read the news article that was accompanied by user comments that positively addressed the accuracy of the article judged it to be of higher quality in terms of both accuracy and impartiality than participants that read the news article that was accompanied by user comments that negatively addressed the accuracy of the article. Thus, H1 is supported. Additionally, for the assessment of accuracy, a significant main effect of issue endorsement ($F(1,102) = 5.53, p = .021$, $\eta_p^2 = .05$).

---

Table 2: Main and interaction effects between the addressing of accuracy/impartiality and issue endorsement in user comments

<table>
<thead>
<tr>
<th>Accuracy Group</th>
<th>Accuracy +</th>
<th>Accuracy -</th>
<th>Accuracy addressing (AA)</th>
<th>Issue endorsement (IE)</th>
<th>AA*IE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Issue + (n=30)</td>
<td>Issue - (n=29)</td>
<td>F</td>
<td>$\eta_p^2$</td>
<td>F</td>
<td>$\eta_p^2$</td>
</tr>
<tr>
<td>Accuracy assessment</td>
<td>4.2</td>
<td>3.6</td>
<td>3.0</td>
<td><strong>38.9</strong>*</td>
<td>.28</td>
</tr>
<tr>
<td>Impartiality assessment</td>
<td>3.2</td>
<td>2.7</td>
<td>2.4</td>
<td><strong>25.6</strong>*</td>
<td>.20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Impartiality Group</th>
<th>Impartiality +</th>
<th>Impartiality -</th>
<th>Impartiality addressing (IA)</th>
<th>Issue endorsement (IE)</th>
<th>IA*IE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Issue + (n=27)</td>
<td>Issue - (n=29)</td>
<td>F</td>
<td>$\eta_p^2$</td>
<td>F</td>
<td>$\eta_p^2$</td>
</tr>
<tr>
<td>Accuracy assessment</td>
<td>4.1</td>
<td>3.9</td>
<td>3.6</td>
<td><strong>4.2</strong></td>
<td>.04</td>
</tr>
<tr>
<td>Impartiality assessment</td>
<td>3.3</td>
<td>2.3</td>
<td>2.4</td>
<td><strong>37.5</strong>*</td>
<td>.27</td>
</tr>
</tbody>
</table>

Notes: A plus (“+”) indicates that the quality criterion was positively addressed resp. that the issue was endorsed in user comments, while a minus (“-”) indicates that the quality criterion was negatively addressed resp. that the issue was not endorsed in user comments. Both accuracy and impartiality assessments are indices based on 2 items (scale from 1 to 5, the higher the value, the higher the perceived journalistic quality in terms of accuracy resp. impartiality).

* $p < .05$; ** $p < .01$; *** $p < .001$
could be observed, thereby supporting H2 for accuracy assessments. However, for assessments of impartiality, no main effect of issue endorsement was found. Instead, the ANCOVA showed a significant interaction effect between the addressing of impartiality and issue endorsement, $F(1,102) = 5.02, p = .027, \eta^2_p = .05$. As Figure 1 shows, participants that were exposed to a mixed valence comment thread consisting of user comments that positively addressed the articles accuracy but evaluated the issue of the article negatively, perceived the article’s impartiality to be the highest. Furthermore, for assessments of accuracy, a significant effect of the covariate frequency of comment posting emerged, with further inspection of the data suggesting that participants, who indicated to regularly post own comments, perceived the article to be less accurate, $F(1,102) = 4.33, p = .040, \eta^2_p = .04$.

In the impartiality group, a significant main effect of the addressing of impartiality on both accuracy (AA) and impartiality assessments (IA) could be observed, $F_{AA}(1, 102) = 4.18, p = .043, \eta^2_p = .04; F_{IA}(1, 102) = 37.52, p < .001, \eta^2_p = .27$ (see Table 2), thereby supporting H1. Again, participants that read the news article that was accompanied by user comments that positively addressed the impartiality of the article judged it to be of higher quality in terms of both impartiality and accuracy than participants that read the news article that was accompanied by user comments that negatively addressed the impartiality of

Figure 1: Interaction effect between accuracy addressing and issue endorsement on impartiality assessments (accuracy group)

Notes: A plus (“+”) indicates that the quality criterion was positively addressed resp. that the issue was endorsed in user comments, while a minus (“-“) indicates that the quality criterion was negatively addressed resp. that the issue was not endorsed in user comments. Impartiality assessment is an index based on 2 items (scale from 1 to 5, the higher the value, the higher the perceived journalistic quality in terms of impartiality).
the article. The predicted main effect of issue endorsement could not be observed for either AA or IA, thus contradicting H2 in the impartiality group. Apart from that, the ANCOVA showed a significant effect of the covariate prior knowledge about the issue for accuracy assessments in the impartiality group, $F(1,102) = 4.08$, $p = .046$, $\eta_p^2 = .04$: Participants, who indicated to have more knowledge about the military intervention, perceived the article to be more accurate.

5. Discussion

Our results show that user comments that addressed the journalistic quality criteria impartiality and accuracy had a systematic effect on readers’ quality assessments. If user comments praised the news article as being impartial and balanced, it actually led people to perceive the article to be of higher quality in terms of impartiality and – to a smaller degree – in terms of accuracy as well. The same effect could be observed for the assessment of accuracy. If user comments highlighted the article to be without erroneous information or contradictions, it was perceived as being more accurate and, once more, as being more impartial. Thus, it seems that compliments or acknowledgements about coverage in user comments generally lead readers to perceive a news article as being of higher quality, regardless of whether the comments explicitly address the article’s impartiality or its accuracy. In the context of credibility assessments, Yale and colleagues (2015) obtained similar results and found that people tend to evaluate news credibility in a heuristic manner, thus making nuanced distinctions between different types of credibility un-

likely. Hence, our assumption that a news article with user comments positively addressing the accuracy/impartiality of the article is judged to be of higher quality than a news article with user comments negatively addressing the accuracy/impartiality of the article, proved to be correct in both groups under investigation and for both kinds of quality assessments.

However, the assumption that the expression of consent or dissent in comments (in the form of issue endorsement) might induce a kind of spillover effect on perceptions of the article could only be confirmed for accuracy assessments in the accuracy group. The news article that was accompanied by user comments that endorsed a military intervention against ISIS was judged to be of higher accuracy than a news article that was accompanied by user comments that refused such an intervention. While it is not entirely clear why this effect was only observed for accuracy assessments, we suspect that this might be explained by the fact that “it is difficult for recipients to evaluate articles for accuracy because they need background knowledge on the topic to formulate adequate judgements” (Urban & Schweiger, 2014, p. 833). In our case, this background ‘knowledge’ might simply have been a cue derived from the valence of user comments. Thus, the positive valence (induced through expressing support/endorsement) in the comments might have been carried over to the perception of the news article’s accuracy, thereby positively influencing participants’ assessment. As our analysis has shown, accuracy assessments were also influenced by participants’ frequency of comment posting: Users who post own
comments more frequently, perceived the article to be less accurate. Motives connected to posting comments (cf. Springer et al., 2015; Ziegele et al., 2013) indicate that commenters are more critical and questioning in nature, further suggesting that they might judge an article’s accuracy more sceptically. However, this result has to be considered with caution since the distribution of comment posting frequency is highly positively skewed, in that 89.7% of our participants never post own comments and only 10.3% (at least rarely) do so. In the group of participants that read comments focusing on impartiality (impartiality group), participants’ accuracy assessment was, however, neither influenced by the addressing of the article’s issue nor participants’ frequency of comment posting. Instead, the accuracy assessment was influenced by participants’ prior knowledge about the issue. We assume that – due to the fact that no cues about the article’s accuracy were available in this group – participants had to rely more (or even solely) on their available knowledge. Because the article was composed, we were able to ensure that it was indeed without erroneous information or contradictions: Participants with more knowledge might actually have been able to recognize this accuracy.

Besides, for assessments of impartiality in the accuracy group, a rather interesting interaction effect between accuracy addressing and issue endorsement emerged. Somewhat counterintuitively, participants that read a comment thread that consisted of user comments that positively addressed the article’s accuracy but evaluated the issue of the article negatively, perceived the article’s impartiality to be the highest. Since impartiality involves the attempt to cover diverse ideas and opinions with professional detachment, the presentation of differently valenced and dissident user comments alongside the article probably spread to participants’ perception of the article’s impartiality. However, we certainly require more research to fully understand the mechanisms behind news users’ quality assessments.

Overall, our study provides evidence that comments left by anonymous users on news websites can serve as meaningful cues for people’s own assessments. If journalistic work can be distorted or degraded by negative comments, or, on the other hand, appears more valuable, factual, or well researched due to positive comments, it seems increasingly crucial for online editorial departments to develop strategies for dealing with comments. One approach might be to actively participate in user discussions. Meyer and Carey (2014) have shown that having an active moderation in the comment section actually increases users’ willingness to post comments, and, supposedly, more thoughtful and respectful ones as well. At the same time, this feature offers the possibility for journalists to respond to criticism early on and to “take ‘emotional and physical ownership’ of the online environments they control” (p. 225).

Of course, our study is not without limitations. Following recommendations for online surveys to minimize the need for vertical scrolling as well as to control the reading time for both the article and the user comments, we presented the article and the user comments on two sites of the questionnaire. The consequent highlighting of the user comments, however, might...
have increased their influence (recency effect). Besides, trying to make the content of the manipulated comment threads as realistic as possible, we somewhat abandoned internal validity in favour of external validity. Hence, the wording and arguments differed between the experimental conditions—especially for the factor of issue endorsement. However, as the treatment check has shown, participants nevertheless perceived the comments as intended. To keep it simple, we furthermore disregarded the “interpersonal nature” (Ziegele, Breiner, & Quiring, 2014, p. 24) of user comments. The commenters in the stimulus do not refer to each other or even challenge or question certain opinions expressed by their fellow commenters. Future studies should thus investigate whether a comment thread with disputing and arguing commenters influences (quality) assessments differently. Additionally, it would be valuable to examine the behavioral consequences of the identified shifts in quality assessments. Do they actually affect future selection decisions or are perceptions of quality only a small element in the process of selecting media (e.g., Wolling, 2009)? It cannot be denied that user comments are meaningful cues in the environment of online journalism. Hence, they should not be treated as some kind of irrelevant attachment to news but rather as an inherent part of nowadays news production and consumption processes.

Note: We would like to thank Gianna Banke, Tabea Eppelein, Michael Fischbacher, Markus Gropper, Corinna Hacker, Natascha Herr, Julia Herzer, Sebastian Huber, Laura Michalowski, David Rausch, Anna Schmitt, Leonie Teisner, and Julius zur Nedden, who participated in a seminar we conducted at LMU Munich and eagerly worked with us to prepare and conduct the experiments. Furthermore, we would like to thank the anonymous reviewers and the SCM editorial staff for their constructive comments and suggestions.

References


Thorson, K., Vraga, E., & Ekdale, B. (2010). Credibility in context: How un-


