

*Full Length Research Paper*

# Web news readers' comments: Towards developing a methodology for using on-line comments in social inquiry

Natalie Henrich<sup>1\*</sup> and Bev Holmes<sup>2</sup>

<sup>1</sup>Centre for Health Evaluation and Outcome Sciences, Vancouver, Canada.

<sup>2</sup>Michael Smith Foundation for Health Research, Vancouver, Canada.

Accepted 26 November, 2012

**Readers of on-line news are increasingly posting comments in response to news articles and the comments of other readers. These comments are a rich source of qualitative data that reflect public opinions and provide insight into how decisions are made and beliefs formed. Analyzing this unique data source requires a well thought out methodological approach to meet the goals of the study. This article explores opportunities and challenges associated with using on-line comments in qualitative studies, and encourages a debate among researchers about these issues toward finessing approaches for using these unique data most effectively.**

**Key words:** On-line news, posted comments, methodology, inclusion criterion, public deliberation, participatory journalism.

## INTRODUCTION

Readers' comments posted in response to on-line news articles are a form of participatory journalism that gives the public a medium for expressing their perspectives on current issues (Chung, 2008; Manosevitch and Walker, 2009). Comments have the potential to increase our understanding of public opinions, how the public makes decisions and how beliefs are formulated, yet comments have only rarely been used as data. A challenge to using readers' comments in research is confirming a methodology for analyzing these data. Methodological decisions are critical due to both the uniqueness and relative newness of these data, and the various purposes an analysis might fulfill. These purposes include measuring the range and common instances of reactions to an article (Schuth et al., 2007), understanding factors that predict the use of such interactive features in online newspapers (Chung, 2008), and exploring the extent to which online reader comments represent a form of public

deliberation (Manosevitch and Walker, 2009).

We analyzed comments posted in response to news articles dealing with the H1N1 vaccine during the 2009 H1N1 pandemic (Henrich and Holmes, 2011). Our intent was to better understand what was acknowledged as resistance by the general public to getting vaccinated, by exploring factors that might influence their decision making. The study analysed 1,796 comments posted in response to 12 articles on websites of three major Canadian news sources. Articles were selected based on topic and number of comments.

In our study, we found ourselves creating elements of our methodology. Though largely drawing on standard qualitative methods, we found that aspects of the study did not conform to traditional approaches. We recognized an opportunity for the research community to embrace this area of social inquiry and collectively develop and evolve a range of procedures, depending on the goal of the analysis, for working with what is acknowledged as a rich new data source (Lee and Yoon, 2010; Reich, 2009; Sooyoung and Youngshin, 2009). Drawing on our experience, we have identified a number of methodological

---

\*Corresponding author. E-mail: [nhenrich@cheos.ubc.ca](mailto:nhenrich@cheos.ubc.ca).

areas that are ripe for discussion and development, including depth vs. breadth of the sample under study, demographics of those commenting, inclusion criteria (for both news sites and comments) and access to comments.

### **On-line comments**

Commenting in response to online news provides an opportunity for readers to interact with each other or express their agreement or disagreement with the content of the article or other comments. “Commenters” (Schuth et al., 2007) may also correct what they believe to be misinformation in other comments, provide support or share anecdotes, and provide ‘facts’ and links to websites with additional relevant content. The back and forth among commenters creates a conversation in which ideas are debated and negotiated. Readers of comments often have the option of indicating whether they agree or disagree with a comment by simply clicking on the agree/disagree links that usually appear with each comment. Comments cannot be taken as representative of the views of the general population. However, due to the high number of comments available on certain articles, they can reflect the perspectives of a large segment of the population.

Comments yield real-time insights into public attitudes on issues, the factors that influence decision making on an issue, and the particular content that most strongly influences these decisions. These comments and responses serve as a gauge of public opinion that is immediate, spontaneous and (presumably) honest. Because comments seem to influence the opinions of other readers the perspectives conveyed in the comments may come to reflect wider public opinion (Lee and Yoon, 2010; Park and Lee, 2007). Given that almost all on-line news articles on major news sites allow for posting of comments, readers can choose to comment on those topics that are of greatest personal interest. Hence the volume of comments on a topic may act as an indicator of the relative importance or passion the public has for different issues. Due to the immediacy, anonymity and largely unmoderated nature of comments, comments tend to be more impulsive, shallow and aggressive than traditional forms of audience participation (Reich, 2009), which may mean that the comments are providing a truer insight into people’s opinions than those expressed in other contexts. Manosevitch and Walker (2009) suggest that comments “provide more diverse and authentic public deliberation” than traditional letters to the editor. Sooyoung and Youngshin (2009) assert that comments provide a more realistic setting for understanding the public response as opposed to interviews and experimental studies that create artificial contexts for responding to an issue. Data derived from comments are entirely participant driven and presumably reveal the

issues that matter to the commenters.

Comments also play a role in shaping the attitudes of the general public. Thus, comments may become more reflective of general attitudes as readers – even those who do not comment, or “lurkers” (Manosevitch and Walker, 2009)– integrate the perspectives of the comments into their own views. Studies show that people modify their beliefs and behaviours based on how they think other people are responding to media, and readers interpret comments as a good gauge of public opinion despite their non-representativeness (Lee and Yoon, 2010; Park and Lee, 2007). Even though the rating of an article (for example, number who agree/disagree) is more representative of public opinion than comments (because more people rate an article than comment on it), readers’ opinion of public attitudes about a topic are more influenced by the comments than the numbers who agree or disagree (Lee and Yoon, 2010). A study using manipulated article content and associated comments found that readers were more influenced by the comments than the article (Yang, 2008).

It is difficult to estimate the percentage of the population that participates in posting comments. According to the Project for Excellence in Journalism, in 2009 approximately 7% of all Americans posted comments about news stories, as did 20% of the most technologically oriented internet users (The Project for Excellence in Journalism, 2009). Although, a minority of the public actively posts comments, these comments are read by large swaths of the population. A study in South Korea found that 84.3% of Internet news users read others’ postings at least once a week (Na and Rhee, 2008).

The Internet is providing increasing opportunities for online journalism involving the public; online reader comment forums represent one of the newest ways for readers to be heard (Santana, 2011). The likelihood is that more and more people will join these forums to debate issues of importance to citizens. Such interactions offer the potential to not only enhance collaborative knowledge generation (Shanahan, 2010), but to study how knowledge is generated and shared. The methodology to support such study is in its infancy, but rapidly evolving. We offer considerations below, based on our recent experience with comment analysis, to add to the discussions among researchers who are using this unique new data source.

## **METHODOLOGICAL CONSIDERATIONS**

### **Demographic depth vs. breadth**

In selecting news articles on a study’s topic, numerous articles could be selected from a single news source in order to ensure that commenters from that source are well represented.

Another approach is to select a smaller number of articles from numerous sources, which would broaden the population of commenters that would be represented but reduce the under-

standing of the commenters from any one source. If multiple news sources are used, there is a question about how to interpret differences in comments across sources. It is difficult to determine if differences are attributable to demographic differences across sources, or differences in the articles that were included from each source. Specific studies of newspapers and their comments – for the purposes of analyzing the extent to which overall reader profiles are similar to overall commenter profiles (see below) – would be useful in this regard in the longer term. In the shorter term, it will be important that the source and comment selection reflects the goals of the study, and that limitations related to generalizability are clearly spelled out.

### **Uncertainty of commenters' demographics**

Reader profile for news sites provides an indication of the population from which commenters on each site come, but it does not tell us if the profile of commenters parallels the profile of readers. Although different news sites may have different reader profiles, the self-selection of commenters may make the commenters from each site more similar than one would expect given the readers' profiles. If commenters are more similar than readers, then differences in comments across news sites may not be attributable to demographic attributes but perhaps to other factors. Without more information of commenters' demographics, resolving the issue of attribution remains problematic. Chung (2008), noting that the characteristics of audiences who engage with online newspapers are absent in the literature, conducted a survey of online readers who participate in interactive features of news sites and found respondents' characteristics in accord with the user profile of the participating newspaper. More work in this area would greatly contribute to the understanding of online comments and to our ability to produce useful analyses.

### **Article inclusion criterion**

An inclusion criterion is needed for determining which articles from among all the articles dealing with a topic will be included in a study. A potential criterion could be to select the articles with the most comments. This selection could be based on the assumption that the number of comments is indicative of the importance of an article to readers with readers commenting on articles that most resonate with them. However, it is not clear that number of comments and importance to readers necessarily correlate. It may be the case that readers are more likely to post comments if they disagree with a news story or feel discontent about an issue. Therefore, selecting articles based on the number of comments could distort our understanding of public opinions. Another possible criterion is to include articles on specific aspects of a topic that are identified *a priori* as important by the researchers, although this may exclude articles that are important to the public if the researchers' and the public's perspectives differ. Articles could be randomly selected although this runs the risk of missing stories that were of particular importance to the public and/or researchers. Debate on article inclusion criteria is clearly needed. A start would be for researchers who study online comments to include in their manuscripts reflections on this aspect of the methodology and the effect it may have had on the results.

### **Comment inclusion criterion**

An inclusion criterion is needed for determining which comments to include from among all the comments posted in response to an article. Some possible criteria are to include all comments, to include one comment per user name in order to avoid biasing the

analysis by over representing the opinions of users who post multiple times, or to include comments that relate to the article or the topic of interest. From an analysis perspective, the question of including or excluding off-topic comments becomes significant. Consider a hypothetical scenario. If we report that out of 1,000 comments posted in reply to an article, 200 dealt with the safety of vaccines then 20% of comments were on that topic. However, if 300 of the comments were unrelated to vaccine safety and we remove them from the dataset then we could report that out of 700 comments on vaccine safety, 200 dealt with the safety of vaccines and thus 29% of comments were on that topic. It is uncertain which calculation better reflects the importance of safety of vaccines to the population of commenters. As above, our recommendation is that researchers involved in the study of online comments think carefully about comment inclusion criteria, share their inclusion rationale with readers, and reflect on the extent to which different criteria may have generated different results.

### **Time limited access to comments**

News sites do not leave articles and the associated comments online indefinitely. This limits access to the data to the window during which the information is posted and may prohibit use of these data for exploring research questions retrospectively. There may be ways to access comments that have been taken off-line, although we were unsuccessful in retrieving "historical" comments and we are uncertain as to what happens to comments when they are removed from a news site. Our efforts to find out where these comments go and what is done with them was also unsuccessful, which made us curious as to editors' and publishers' rationale for including a comment function with their online news. Are they using the public's comments to find out what interests the public the most, or to encourage debate among readers? Or have they simply been "thrust into the Web 2.0 world" (Santana, 2011) without having thought about the possibilities and ramifications? Conversations with news editors about the shared opportunities offered by news comments could prove useful in this regard.

### **Using public consensus data**

Many sites allow readers to indicate whether they agree or disagree with posted comments. It is necessary to establish an approach for incorporating data on agreement and disagreement with comments into analyses because they provide a measure of representativeness of individual comments to a larger audience. We suggest that collaboration between qualitative and quantitative scientists/statisticians could lead to some promising mixed methods studies using online comments.

## **FUTURE DIRECTIONS**

If the content of comments is going to be used to help shape editorial, policy, and public health strategies, a range of methodologies for effectively and validly using the comments would ideally be collectively developed, evolved and shared. Given the paucity of research using comments, it appears that comments are not being considered as more than a platform for people to vent or to swap opinions with other commenters. Another view is that perhaps comments simply are not on the radar of those who could benefit from the insights contained within comments. In addition to the methodological

challenges of working with a new type of data is the challenge of raising awareness about the existence and value of these data.

Apart from the study of comments themselves, a further area of potential research is the role of comments in participatory journalism. We believe that studies involving commenters, such as in focus groups or interviews, would shed light on motivations for on-line participation, how commenters would like their comments to be used, and what they see as the benefits and challenges of expressing themselves in this format.

Work has started on understanding how media perceives and uses comments (Reich, 2009; Thurman, 2008), but studies to date suggest that as yet online news publications are not maximizing the use of the interactive features available via the Internet (Chung, 2008). A fruitful line of inquiry would be to learn more about media's use and attitudes towards comments in particular and participatory journalism in general; as well as how participatory journalists and traditional media can collaborate to amplify public voices.

Finally, we suggest that research involving online comments could also be expanded to include policy makers – who reportedly seek public input into decisions that affect the public – and explore how they use or could use the public's comments to inform their decisions.

## ACKNOWLEDGEMENT

This article draws on research funded by Genome British Columbia.

## REFERENCES

- Chung D (2008). Interactive features of online newspapers: Identifying patterns and predicting use of engaged readers. *J. Comput. Mediat. Comm.* 13:658-679.
- Henrich N, Holmes B (2011). What the Public Was Saying about the H1N1 Vaccine: Perceptions and Issues Discussed in On-Line Comments during the 2009 H1N1 Pandemic. *PLoS One* 6(4):e18479.
- Lee EJ, Yoon JJ (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. *Commun. Res.* 37(6):825-846.
- Manosevitch E, Walker D (2009). Readers comments to online opinion journalism: A space of public deliberation. 10th International Symposium on On-line Journalism. Austin, Texas.
- Na EK, Rhee JW (2008). A study on readers' comments: Changes in the online news usage pattern and the implications for discursive public. Seoul, Korea: Korea Press Foundation.
- Park N, Lee K (2007). Effects of Discussions in Online News Forum on Corporate Reputation. *Public Relat. Rev.* 33(3):346-348.
- Reich Z (2009). Weaving the Thread: Gatekeeping and Filtering Strategies for User Comments. Conference Papers -- International Communication Association: 1.
- Santana A (2011). Online readers' comments represent new opinion pipeline. *News Res. Jour.* 32(3):66-81.
- Schuth A, Marx M, de Rijke M (2007). Extracting the discussion structure in comments on news-articles. 9th ACM International Workshop on Web Information and Data Management (WIDM 2007) pp. 97-104.
- Shanahan M (2010). Changing the meaning of peer-to-peer? Exploring online comment spaces as sites of negotiated expertise. *J. Sci. Commun.* 9(1):A01.
- Sooyoung C, Youngshin H (2009). Netizens' Evaluations of Corporate Social Responsibility: Content Analysis of CSR News Stories and Online Readers' Comments. *Public Relat. Rev.* 35:147-149.
- The Project for Excellence in Journalism (2009). Citizen-based media. The State of the News Media: Journalism.org.
- Thurman N (2008). Forums for citizen journalists? Adoption of user generated content initiatives by online news media. *New Media Soc.* 10(1):139-157.
- Yang HS (2008). The effects of the opinion and quality of user postings on Internet news readers' attitude toward the news issue. *Korean J. Comm. Stud.* 52:254-280.