

Influence of user ratings, expert ratings and purposes of information use on the credibility judgments of college students

Sook Lim and Nick Steffel

Master of Library and Information Science Program, St. Catherine University, Saint Paul, Minnesota, United States

Abstract

Introduction. This study examined whether user ratings, expert ratings and the purpose of the use of a book on a user-generated site influenced the credibility of the book. It also examined whether the effects of user ratings and expert ratings on credibility judgments of the book varied according to the purpose of information use. In addition, this study examined whether the co-existence of user ratings and expert ratings influenced people's credibility judgments of user ratings, and vice versa, on a user-generated content site.

Method. Data were collected using both an experiment and a survey. The experiment took the form of a 2x2x2 factorial design, involving user ratings of a book (high or low ratings), expert ratings (high or low ratings) and purposes (academic or casual reading). A total of 233 college students participated in the study.

Analysis. Analysis of variance (ANOVA) was used to test the research hypotheses and answer the research questions of the study. Correlation analyses were performed for the exploratory findings.

Results. Expert ratings influenced the credibility of the book while user ratings did not. The effects of user ratings and expert ratings did not vary according to the purpose of information use regarding the credibility of the book. However, when expert ratings were high, the respondents perceived higher user ratings as more credible than lower ones. On the other hand, when expert ratings were low, the respondents perceived lower user ratings as more credible than higher ones.

Conclusions. Expert opinions are important in the credibility judgments of the book and the credibility of user ratings on a user-generated site. The results provide practical implications for user-generated content designers and librarians.

CHANGE FONT

Introduction

Web 2.0 technologies enable anyone to contribute to content. User-generated content can provide useful information and diverse perspectives through the experiences and situated knowledge of numerous individuals. However, it can also contain inaccurate information because of its unfiltered publishing mechanism. As a result, the quality or credibility of user-generated content is uncertain and users often face challenges discerning credible from non-credible information.

Researchers point out that the uncertain quality of information and the abundance of information on the web lead people to become highly dependent on cues in assessing such information (Fu and Sim, 2011). People use a variety of heuristic methods that are triggered by these cues. Among such methods, the two endorsement-based ones are widely discussed in the credibility literature, which are social endorsement and authority heuristics. First, in the participatory Web 2.0 environment, social endorsement is a particularly powerful heuristic cue that influences credibility judgments (Flanagin, Metzger, Pure, and Markov, 2011; Metzger, Flanagin, and Medders, 2010). This heuristic operates in a manner such that people are inclined to believe information if others do so, often without their own critical assessments of the information (Metzger and Flanagin, 2013). This judgment rule is known as the bandwagon heuristic, on which collective endorsement and popularity are based (Fu and Sim, 2011; Sundar, 2008). In fact, regardless of their official credentials, laypeople contribute their expertise and experiences to such user-generated content as user ratings and comments, which become experiential authority in this participatory information environment (Pure et al., 2013).

Second, the authority heuristic is another endorsement-based heuristic (Sundar, 2008). This heuristic operates in the following manner: people believe if a topical expert or official authority endorses a site or information, then this site or information should be credible (Sundar, 2008). With respect to authority as a cue on the Internet, and despite certain claims of the diminished importance of formal authority in credibility judgments (Lankes, 2008), previous research has shown that authority cues still influence credibility judgments concerning user-generated content or search results (Flanagin and Metzger, 2013; Schwarz and Morris, 2011; Sundar, Xu, and Oeldorf-Hirsch, 2009). In fact, traditional experts who have recognised expertise provide their specialized knowledge using a variety of Web 2.0 technologies as well. As a result, both the opinions of laypeople and experts often co-exist on the Web. In addition, Lee and Sundar (2013) contend that no particular cue is more privileged than others in credibility judgments. This evidence indicates that neither formal authority nor social endorsement cues alone influence credibility judgments. It is likely that users utilise noticeable cues jointly as well as individually in evaluating the credibility of such information.

Additionally, previous studies have shown that the purpose of information searching can affect an individual's credibility judgment (Metzger, 2007; Rieh and Hilligoss, 2008; Stavrositu and Sundar, 2008). In other words this literature suggests that the effect of social endorsement and authority on credibility judgments may vary according to the purpose of information.

Based on the literature, this study examined: 1) whether authority, as well as social endorsement is manifested in credibility judgments; 2) whether people use authority and social endorsement cues jointly; 3) whether the effect of authority and social endorsement cues differs according to the purpose of information use on credibility judgments; and 4) whether the co-existence of authority and social endorsement cues influences people's credibility judgments of social endorsement or authority itself. The study examined user ratings and expert ratings for social endorsement and authority respectively. The study employed academic and non-academic reading for the purposes of information use.

More specifically, this study explored the following questions:

RQ1. Do user ratings influence the credibility judgments of a book on a user-generated content site?

RQ2. Do expert ratings influence the credibility judgments of a book on a user-generated content site?

RQ3. Does the effect of user ratings on credibility judgments of a book differ according to expert ratings on a user-generated content site?

RQ4. Does the effect of user ratings and expert ratings on credibility judgments of a book on a user-generated content site vary according to the purpose of information? In other words, is there a three-way interaction among user ratings, expert ratings and the purpose of information regarding the credibility judgments of a book on a user-generated content site?

RQ5-6. How do user ratings and expert ratings manifest themselves in the credibility judgments of expert ratings or user ratings, given the purpose of information use? In particular, do the credibility judgments of user ratings vary according to expert ratings (RQ 5) and vice versa (RQ6) on a user-generated content site?

The main contributions of this study are as follows: 1) this study provides new knowledge of the role of expert ratings as a confirmatory factor of the credibility of user ratings. This is the most interesting finding and the important contribution of this study to the credibility literature concerning user-generated content. This study shows that authority still matters in the age of the wisdom of crowds. Not only do expert ratings influence users' credibility judgments of a book on a user-generated content site, but they also work as a confirmatory or benchmarking factor of the credibility of user ratings. However, the credibility of expert ratings was not affected by user ratings. 2) The results provide practical implications for user-generated content designers and librarians. That is, it would be useful for user-generated content designers to offer certain features for expert opinions or reviews along with user ratings. In addition, librarians can provide their opinions, reviews or a list of expert sources, which can be helpful and informative to users.

Literature review

Concept of credibility

In the 1950s Carl Hovland and his associates defined credibility as expertise and trustworthiness (as cited in [Self, 2009](#), p. 438). Since then, other scholars have widely adopted this concept of credibility ([Flanagin & Metzger, 2008](#); [Jensen, 2008](#); [Lim, 2013](#); [Rieh, 2010](#); [Rieh & Danielson, 2007](#); [Tseng & Fogg, 1999](#); [Wang, Walther, Pingree, & Hawkins, 2008](#)). Expertise refers to a communicator's competence or qualifications to know the truth about a topic, while trustworthiness refers to a communicator's inclination or motivation to tell the truth ([Jensen, 2008](#); [Wang et al., 2008](#)). The two dimensions of expertise and trustworthiness are not always perceived together ([Rieh, 2010](#)). Nonetheless, [Self \(2009\)](#) contends that credibility judgments are ultimately a subject matter in the new interactive media environment where users make situated judgments. Based on this literature, this study defines credibility as an individuals' assessment of whether information is believable based on their knowledge, experience and situation.

Relevant studies

In the participatory user-generated content environment, Internet users aggregate their experience and build their collective expertise regardless of their formal credentials ([Pure et al., 2013](#)). User ratings, comments or reviews are typical forms that display such collective experiential expertise. The literature shows that such individuals' aggregated first-hand experiences or opinions about a topic or situation influence people's credibility or quality judgments ([Metzger and Flanagin, 2011](#); [Flanagin et al., 2014](#); [Xu, 2013](#)). More specifically, [Xu's \(2013\)](#) study shows that people perceived a news story with a high number of social recommendations as more credible than that with a low number. Similarly, people perceived a product with high user ratings as having a better quality than that with low ones ([Flanagin et al., 2014](#)). In this environment, people tend to perceive information as credible if others do so ([Metzger et al., 2010](#)). In other words, social endorsement is a factor affecting the credibility judgments of user-generated content. This literature leads to the following hypothesis:

H1. User ratings affect the credibility judgments of a book.

The social endorsement heuristic has both benefits and drawbacks. According to [Messing and Westwood \(in press\)](#), social endorsement cues convey more social relevance, interest and utility in selecting news stories than do source cues. As a result, social endorsement cues are more powerful heuristic cues than source cues. This is because people are more likely to trust a news story recommended by people they like than a story without a recommendation. Furthermore, a source (e.g., New York Times) can include a variety of contents, while the recommended news stories are likely to be relevant to the reader. Similarly, in a user-generated content site such as YouTube, laypeople's videos are more likely to match viewers' expectations, and are therefore more likely to be accepted by them ([Paek, Hove and Jeong, 2011](#)). For instance, in their study regarding child abuse prevention public service announcements on YouTube, [Paek and her colleagues \(2011\)](#) found that more favourable attitudes toward user-generated public service announcement than towards professional ones. In addition, social endorsement cues have other benefits. [Messing and Westwood's \(in press\)](#) study reveals that social endorsement leads to reduced partisan selectivity of news stories.

However, the social endorsement-based heuristic does not always result in positive outcomes. Social endorsement can lead to the blind acceptance of others' opinions without one's own scrutiny, and tends to induce popularity bias over quality ([Metzger and Flanagin, 2011](#); [Sundar et al., 2009](#)). Furthermore, it is likely that unfiltered user-generated content sites run the risk of containing inaccurate or misleading information, which can result in serious consequences to users. For instance, a review study by [Madathil \(in press\)](#) shows that some public health-related YouTube videos contain inaccurate and misleading information that contradicts mainstream medical standards (e.g., vaccinations). More seriously, such videos received higher mean ratings and more views than videos supporting standard guidelines. However, interestingly, Internet users seem to have scepticism about user ratings or reviews at least regarding commercial online sites ([Flanagin et al., 2014](#); [Reichelt, Sievert and Jacob, 2014](#)). It is therefore likely that Internet users do not entirely depend on social endorsement cues for assessments of the credibility of user-generated content. It is probable that users employ other available cues to signify relatively certain credibility or quality in credibility judgments. Sources with credentialed expertise or signals of formal authority can be such cues.

Authority is traditionally one of the most important criteria for assessing the credibility of information ([Rieh and Hilligoss, 2008](#); [Sundar, 2008](#)). People are inclined to believe that information from an authoritative agency (e.g., US Department of Health and Human Services) or a person with credentialed expertise (e.g., MD) should be trusted. [Sundar et al. \(2009\)](#) show that although online users are generally influenced by peer endorsement

cues, authority cues can influence credibility judgments when peer endorsement and authority cues are inconsistent. In a similar vein, Lim's (2013) study shows that there is a positive correlation between professors' endorsements and college students' perceived credibility of Wikipedia, although the correlation between peer endorsement and students' perceived credibility of Wikipedia was much higher. In addition, Schwarz and Morris (2011) show that augmenting search results with the visualization of additional features, such as overall popularity and expert popularity helps users make more accurate credibility assessments. In particular, information about the visitation patterns of topical experts appears to be effective. This literature leads to the following hypothesis:

H2. Expert ratings affect the credibility judgments of a book.

Recent empirical studies have shown that credibility evaluations depend on a number of factors such as the volume of ratings or number of followers. For instance, Flanagin and Metzger's (2013) study reveals that when the volume of ratings is low, expert ratings are perceived as more credible than non-expert ratings on a movie site. However, at a high volume of ratings, non-expert ratings are preferred to those of experts'. In a similar line of research, but demonstrating different findings, Lee and Sundar (2013) found that users perceived the content of tweets to be more credible when a professional source (high authority) with many followers tweets, compared to one with fewer followers. However, tweets from a low authority source with many followers were considered less credible than tweets from a layperson with fewer followers. Yet, their study did not find significant effects for either the authority cue (e.g., professional source) or the bandwagon cue (e.g., number of followers). Based on their study, Lee and Sundar (2013) conclude that one positive cue is not favoured over the other. Instead, cues have an additive effect, and the consistency of cues is important in processing information. Similarly, based on their literature review, Pure and his colleagues (2013) report that credentialed expertise (formal authority) is complemented by laypeople's experiential authority in evaluating the credibility of information. This evidence indicates that it is likely that Internet users will utilise noticeable cues jointly in evaluating the credibility of information, thus leading to the following hypothesis:

H3. The effect of user ratings on credibility judgments differs according to expert ratings.

Previous studies have shown that the purposes of information searching or use affect user credibility judgments (Metzger, 2007; Rieh and Hilligoss, 2008; Stavrositu and Sundar, 2008). For instance, Stavrositu and Sundar (2008) found that Internet use for information was positively correlated with Internet credibility, while Internet use for entertainment was unrelated to Internet credibility. More recently, Austin *et al.* (2012) found that people tend to use traditional media for educational purposes during crises, while they tend to use social media for insider information. In addition, people perceive traditional media as more credible than social media for crisis information. Despite different contexts, these findings suggest that people process authority and social endorsement cues differently regarding credibility and depending on the purpose of information use.

H4. There is a three-way interaction among user ratings, expert ratings and purpose of information.

The literature shows that consumers of online information are aware that user ratings can be biased, and user ratings may not be a reliable indicator of the quality of a product or service (Dou *et al.*, 2012; Flanagin *et al.*, 2014; Kusumasondjaja, Shanka and Marchegiani, 2012). For instance, research has shown that Internet users perceive ratings as non-credible when ratings are excessively positive (Flanagin *et al.*, 2014; Reichelt, Sievert and Jacob, 2014). Similarly, consumers tend to perceive a negative online review as more credible than a positive online review (Kusumasondjaja, Shanka and Marchegiani, 2012; Pan and Chiou, 2011). In addition, Dou and his colleagues (2012) remark that Internet users of commercial information seem to doubt the possible marketing intentions behind the online reviews of products. This literature shows that Internet users indeed have scepticisms about user ratings or reviews. As a result, it is likely that the co-existence of expert ratings and user ratings can affect their judgments regarding the credibility of user ratings themselves. In addition, Flanagin and Metzger (2013) show that user ratings triumph over expert ratings under the condition of a high volume of ratings. This evidence suggests that user ratings can influence the credibility judgments of expert ratings as well when people are exposed to both ratings simultaneously. Based on this literature, this study further explores the following research questions:

RQ5 & 6. How do user ratings and expert ratings manifest themselves in the credibility judgments of expert ratings and user ratings, given the purposes of information use? In other words, do the credibility judgments of user ratings vary according to expert ratings (RQ 5) and vice versa (RQ6)?

Methodology

Participants

The data were collected at a small private university in the Midwestern United States in fall 2013. The population of the study consisted of female undergraduate students at the university. The study participants were students who lived in the residence halls on campus. A total of 233 students participated in the study. The students had the opportunity to enter a random draw to win a prize of a \$30 gift card.

Data collection methods

This study employed both an experiment and a web survey regarding students' credibility judgments about a book, user ratings and expert ratings. The experiment was embedded in the survey.

The design overview

The experiment took the form of a 2x2x2 factorial design, involving user ratings of a book (high or low ratings), expert ratings (high or low ratings) and purposes (academic or casual reading). The design tested the effects of user ratings (for social endorsement), expert ratings (for authority) and purposes of information use on credibility of the book, user ratings and expert ratings.

Procedure

The participants were directed to the study's website via a written URL included in the solicitation email. On the study site, the participants were asked to read a consent form, prior to their participation in the study. The consent form described the standard protocols of research involving human subjects. In addition, the participants were told that they would be asked to judge the credibility of a site that they would be viewing. After giving informed consent online, the participants were randomly assigned to one of the two following conditions: they were asked to imagine that they were looking for a book to either write a term paper for a course or to spend their spare time when they viewed a site on the next screen. Then, the participants were randomly assigned to one of four screens on a modified book page of Goodreads by a computerized program. They were instructed to read the site they were viewing. Then all participants were directed to a questionnaire that they completed online. Once directed to the questionnaire, the participants were not able to view the site again.

Experimental conditions

A Goodreads book page, titled "Healthy Eating, Healthy World," was selected (at <http://www.goodreads.com/book/show/11494318-healthy-eating-healthy-world>) as the basis for creating the screens of the experiment. The subject of the book was selected, as the topic is controversial enough to raise questions about its credibility, which is suggested by previous studies (Hu and Sundar, 2010). Based on this site, four modified screens, reflecting the experimental variations in user and expert ratings, were created: high user and high expert ratings; high user and low expert ratings; low user and high expert ratings; and low user and low expert ratings. The experimental materials included expert ratings that were not included on the original page to test the effects of expert ratings on credibility judgments. A mean value of 4.7 and 1.6 were created for the high and low ratings respectively, for both user and expert ratings. Another independent variable of purpose was manipulated by asking the participants to imagine that they were looking for a book, either for academic purposes or casual reading.

A total of 375 ratings and 27 ratings were used for the number of user and expert ratings, respectively. These numbers were selected and modified based on Flanagin and Metzger's (2013) study. This study kept the same volumes for all screens to avoid any volume effects, while maintaining the number of ratings on the site at a realistic level.

The overall design of the pages looks similar to a Goodreads book page. Certain elements of the original page were modified for this study to eliminate any possible effects of such elements on credibility judgments. For instance, the publisher, rating details and rating reviews were removed from the page. The description of the book was modified so that it simply described the book's contents. The title and author of the page were modified to avoid evoking prior knowledge of the book. A fictional title and a gender-neutral name were replaced with the

originals. All other elements, including one image for each screen were held constant (See Figure 1 for a screen shot).

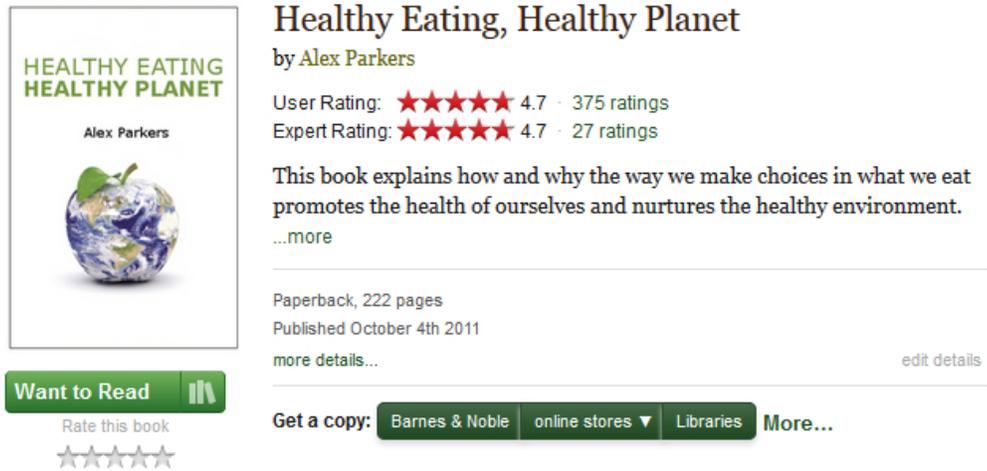


Figure 1: Example stimulus page

Measurement of dependent variables

Credibility was defined as an individuals’ assessment of whether information is believable based on their knowledge, experience and situation. The concept of credibility was examined in three ways: credibility of a book, and credibility of user ratings and expert ratings. The items for credibility of user ratings and expert ratings were modified based on the literature of credibility (Cassidy, 2007; Gaziano and McGrath, 1986; Hilligoss and Rieh, 2008; Lim, 2013; Meyer, 1988; Tseng and Fogg, 1999; Tsfati and Cappella, 2005). All items for these variables were rated on a 7-point scale with the anchors “strongly disagree” and “strongly agree” (Table 1).

Conceptual variables	Survey items	Cronbach’s α
Credibility of the book	<p>The quality of the book appears to be good.</p> <p>The book appears to be credible.</p>	$\alpha=0.773$
Credibility of user ratings	<p>I trust the user ratings information.</p> <p>I found the user ratings information to be credible.</p> <p>I would rely on the user ratings information provided to help me decide whether or not to read the book.</p> <p>I am confident that the user ratings</p>	$\alpha=.939$

	<p>information provided is an accurate reflection of the book's quality.</p> <p>The user ratings are reliable.</p> <p>The user ratings are trustworthy.</p> <p>The user ratings are believable.</p>	
Credibility of expert ratings	<p>I trust the expert ratings information.</p> <p>I found the expert ratings information to be credible.</p> <p>I would rely on the expert ratings information provided to help me decide whether or not to read the book.</p> <p>I am confident that the expert ratings information provided is an accurate reflection of the book's quality.</p> <p>The expert ratings are reliable.</p> <p>The expert ratings are trustworthy.</p> <p>The expert ratings are believable.</p>	$\alpha = .953$

Table 1: Dependent Variables

Findings

The findings are organized into three subsections and by the research questions. The first subsection presents the characteristics of the participants and the descriptive statistics. The second subsection presents the results of the experiment,

corresponding to RQ1 through RQ6. Along the way, the results of the hypothesis testing are reported. Finally, the third subsection reports the other exploratory findings.

Descriptive statistics

Participants. The mean age of the participants was 19.48 (N=180, SD. 1.61). All but one respondent (99.5%, N=190) were female. A majority of the respondents were Caucasian (76.2%, N=144), followed by Asian (7.9%, N=15), African-American (5.3%, N=10), Hispanic (5.3%, N=10) and other including mixed races (5.3%, N=10). The respondents were distributed across majors. Approximately 35.6% (N=83) of the respondents were declared science majors, followed by the social sciences (26.2%, N=61) and the humanities or arts (21%, N=49). Approximately 8.2% (N=19) of the respondents had not yet decided on their majors. Finally, a majority of respondents were first-year students (39.6%, N=76) and sophomores (18.9%, N=44), and one-third of the respondents were juniors (16.7%, N=32) and seniors (17.2%, N=33). The rest were classified as other (3.6%, N=7).

Frequency of viewing user ratings and comments. With respect to user ratings viewed in the past 3 months relative to when this study was conducted, among the 198 respondents, 40.4% (N=80) were occasional users who used ratings between 1 and 5 times. Approximately 28.8% (N=57) were frequent users, with a frequency of more than 11 times. Another 25.8% (N=51) of the respondents viewed ratings moderately, showing a frequency of between 6 and 10 times. The rest (5.1%, N=10) had not viewed ratings during the past three months relative to when the study was conducted.

In addition, among the 196 respondents, one-third often (33.7%, N=66) read user comments, with a frequency of more than 11 times. Another one-third (31.1%, N=61) of the respondents were occasional users who read comments between 1 and 5 times. Approximately 29.1% (N=57) of the respondents read user comments moderately, showing a frequency of between 6 and 10 times. The rest (6.1%, N=12) had not read user comments during the past three months relative to when the study was conducted. The data show that the majority of the respondents (over 94 %) viewed user ratings or comments at least occasionally in the past 3 months relative to when the study was conducted (Table 2).

Posting comments. Among the 198 respondents, a majority of respondents were non-contributors (54%, N= 107) or only occasional contributors (32.3%, N= 64) with respect to commenting on sites in the past three months relative to when the study was conducted. Only a few respondents were frequent contributors (6%, N=12). The data echo Reichelt and his colleagues' (2014) remark that the majority of online users simply consume online content, and only a few people contribute to it.

Relying on ratings to make a decision. Two-thirds (66.3%, N=130) of the respondents at least sometimes relied on ratings information to make a decision in the past three months relative to when the study was conducted. However, the rest (33.6%, N=66) rarely or never relied on ratings (Table 2).

Action	No. of uses	N	%
Viewing ratings in the past 3 months	0	10	5.1
	1-5	80	40.4
	6-10	51	25.8
	11-15	25	15.6
	More than 15	32	16.2
Reading user comments in the past 3 months	0	107	54.0
	1-5	14	32.3
	6-10	15	7.1
	11-15	7	3.5
	More than 15	5	2.5
Posting comments or reviews in the past 3 months	0	107	54.0
	1-5	14	32.3
	6-10	15	7.1
	11-15	7	3.5
	More than 15	5	2.5

Relying on ratings in the past 3 months	Never	22	11.2
	Rarely	44	22.4
	Sometimes	81	41.3
	Often	36	18.4
	Very often	13	6.6

Table 2: Participants' use of online content

Results of hypothesis testing

The experiment was intended to answer RQ1 through RQ6. The research questions from RQ1 through RQ2 were answered by examining the effects of user ratings (RQ1) and expert ratings (RQ2) on credibility judgments, the joint effects of user ratings and expert ratings on credibility judgments (RQ3), and the joint effects of user ratings, expert ratings and the purpose of information use on credibility judgments (RQ4). A three-way Analysis of variance (ANOVA) was performed on the dependent variable of credibility of a book to answer RQ1 through RQ4, which corresponded to research hypotheses H1 through H4, respectively. Two three-way ANOVAs were performed on other dependent variables involving the credibility of user ratings and the credibility of expert ratings to answer RQ5 and RQ6. A two-tailed test under $\alpha=0.05$ was performed to test the hypotheses.

Manipulation checks for the experimental conditions. A series of chi-square tests was performed to assess the effectiveness of the manipulations regarding user ratings, expert ratings and purposes. Results indicated that user ratings differed as expected, with participants having high user ratings noticing higher user ratings more correctly than incorrectly ($N=90$ and $N=15$ respectively, $p<0.00$). Expert ratings differed as intended, with participants having high expert ratings noticing higher expert ratings more correctly than incorrectly ($N=82$ and $N=27$, respectively, $p<0.00$). Finally, the manipulations of reading purposes were successful, with those being instructed for the academic purpose viewing the screen as intended ($N=76$ versus $N=24$, $p<0.00$), and with those being instructed for casual reading viewing the screen as intended ($N=70$ versus $N=30$, $p<0.00$).

RQ1 through RQ4. Only expert ratings were significant ($F(1, 222) = 19.87$, $p<0.00$, Mean square error=107.623). Neither main effects for user ratings or purposes, nor interaction effects were significant. That is, the higher the expert ratings, the higher the credibility of the book (RQ2). However, user ratings did not affect the respondents' credibility judgments of the book (RQ1). The effect of user ratings on the book's credibility did not differ according to expert ratings either (RQ3). Finally, there was no three-way interaction among user ratings, expert ratings and the purpose of information either (RQ4). That is, there were no joint-effect of user ratings, expert ratings and the purpose of information on credibility judgments.

The results indicate that H2 was supported, while H1, H3, and H4 were not. Table 3 presents the mean values of the book's credibility across user ratings, expert ratings and purposes.

Purpose	Expert ratings	Users' ratings						Total
		High ratings			Low ratings			
		M	SD	N	M	SD	N	
Academic	Low	9.53	2.87	26	9.51	2.97	27	9.52 (N=53)
	High	11.25	1.52	35	11.13	1.64	29	11.20 (N=64)
	Total	10.52	2.34	61	10.35	2.49	56	10.44 (N=117)
Casual	Low	9.72	3.14	22	9.59	2.72	27	9.65 (N=49)
	High	10.28	2.22	25	11.25	1.55	39	10.87 (N=64)
	Total	10.02	2.68	47	10.57	2.24	66	10.34 (N=113)
Total	Low							9.58 (N=102)
	High							11.03 (N=128)
	Total	10.30	2.49	108	10.47	2.35	122	N=230

Table3: Expert ratings, user ratings and purposes on the dependent variable, *credibility of the book*

RQ5 and RQ6. Interaction effects between expert and user ratings on credibility of user ratings and expert ratings were examined to find out whether the co-existence of expert and user ratings affected participants' credibility judgments of user or expert ratings themselves. There was an interaction effect between expert and user ratings on the credibility of user ratings ($F(1, 220)=4.66, p<.032, MSE=290.92$). That is, when expert ratings were high, the respondents perceived higher user ratings as more credible than lower ones. However, when expert ratings were low, the respondents perceived lower user ratings as more credible than higher ones (RQ5). On the other hand, the credibility judgments of the expert ratings did not vary according to the user ratings (RQ6).

Taken all together, the credibility judgments of user ratings tend to be subjective in the presence of expert ratings, while the credibility judgments of expert ratings are not affected by the presence of user ratings. In other words, the respondents used expert ratings as a confirmatory or benchmarking factor in their credibility judgments of user ratings, while the credibility of the expert ratings was not influenced by the presence of the user ratings. This is one of the most interesting findings of this study (Figure 2, Table 4).

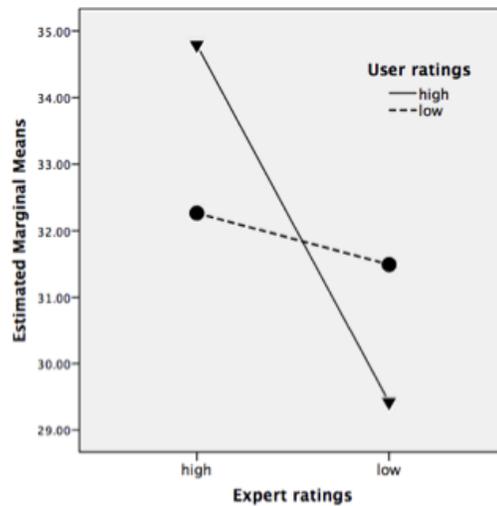


Figure 2: Credibility of user ratings

Purpose	Expert	Users						Total (N)
		High use rate			Low use rate			
		M	SD	N	M	SD	N	
Academic	Low	33.12	5.32	25	26.11	8.98	27	31.03 (N=52)
	High	32.97	8.23	36	35.50	8.03	28	34.07 (N=64)
	Total	33.03	7.13	61	32.36	9.02	55	32.71 (N=116)
Casual	Low	29.89	7.01	22	29.73	9.76	26	29.79 (N=48)
	High	31.56	8.83	25	34.07	6.38	39	33.09 (N=64)
	Total	30.76	7.99	47	32.33	8.12	65	31.67 (N=112)
Total	Low							30.44 (N=100)
	High							33.58 (N=128)
	Total	32.04	7.56	108	32.35	8.51	120	N=228

Table4: Expert ratings, user ratings and purposes on the dependent variable, *credibility of user ratings*

Exploratory findings

This study had a few exploratory findings. Over half (54.4%, N=106) of the 195 respondents perceived user ratings as biased (a mean of 3.32 on a 7 point-scale regarding their agreement on the non-biasedness of user ratings). Interestingly, there was no correlation between the use of user ratings and the participants' perceptions of the non-biasedness of user ratings. However, there were positive correlations between the use of ratings and the participants' perceptions of usefulness ($r=.381, p<0.00$); helpfulness ($r=.386, p<0.00$); and informativeness ($r=.371, p<0.00$).

In addition, this study explored whether contributors of user-generated content were likely to perceive user ratings as more credible than those who did not contribute, given the literature showing that those who contribute more often to user-generated content are more likely to perceive user ratings as more credible and are more likely to rate movies similarly to user ratings, compared to those who contribute less ([Flanagin and Metzger, 2013](#)). This study shows that there was no correlation between contributing to ratings and the perceived credibility of user ratings. However, the data show that those who contributed to ratings were more likely to perceive user ratings as useful, compared to those who did not ($r=.231, p<.001$). In addition, those who posted ratings tended to post user comments as well ($r=.781, p<0.000$).

Taken all together, a considerable number of respondents reported that user ratings are likely to be biased. Interestingly their perceptions of user ratings regarding bias did not prevent them from using user ratings, which is consistent with reports in other studies ([Flanagin et al., 2014](#); [Reichelt, Sievert and Jacob, 2014](#)). The findings show that the use of user ratings was associated with their utilitarian values.

Discussion

The findings show that expert ratings influenced the credibility of the book of the study. The result may be interpreted to mean that expert opinions are still important to users, at least for a book of this study. The result is also consistent with the previous research, which shows that authority cues are critical in evaluating credibility judgments for health information ([Lee and Sundar, 2013](#)). In other words, the respondents might have put more weight on expert ratings than on non-expert ratings for the subject (e.g., healthy eating and the environment) of this experiment in their credibility judgments.

Furthermore, it is worth noting to note that expert ratings were successful as a confirmatory factor in assessing the credibility of user ratings. However, user ratings had no influence on the credibility judgments of expert ratings. The results indicate that the respondents were more confident with authority-based endorsements (e.g., expert ratings) than with social endorsements (e.g., user ratings) in their credibility judgments. Despite the need for further research by employing different samples, the results indicate that the expert opinions are still important in users' credibility judgments. These findings provide practical implications for user-generated content designers and librarians. Expert ratings work as a notable cue to users, which can be helpful to them in evaluating the credibility of user-generated content. Therefore, it is recommended that user-generated content designers offer added features to solicit expert opinions. In addition, librarians could guide expert sources for evaluating the credibility of user-generated content. For instance, it may be useful for librarians to actively create certain features in user-generated content so as to offer their own collective opinions or comments (e.g., librarians' ratings, reviews or choices) for general users or students. Or librarians could provide a list of expert sources or certain expert cues for users. Traditionally, librarians have been the gatekeepers of reliable information, and they can continue to play a similar role in the participatory Web 2.0 environment. This can be particularly beneficial to college students who are frequent users of a variety of user-generated content ([Kim, Sin and Tsai, 2014](#)).

On the other hand, this study did not show the respondents' reliance on user ratings in their credibility judgments, which is inconsistent with the previous studies described above. Despite the need for cautious interpretation, the result may mean that students tend not to be influenced by user ratings when they are simultaneously exposed to expert ratings on the subject (healthy eating and the environment) or the object (a book) of the study. The result may however have something to do with the sample of the study that was drawn from a women's college. Indeed, previous studies have shown that female students perceive user-generated content such as Wikipedia as less credible than their male counterparts ([Lim and Kwon, 2010](#); [Lim, 2012](#)). By the same token, it may be possible that female students tended to rely more on authoritative cues (e.g., expert ratings) in evaluating the credibility of a book in the study. Further research is needed to examine the influence of expert opinions on credibility judgments by employing both sexes.

The study shows no effect with respect to the purpose of information use or no joint-effect among the purpose of information, expert ratings and user ratings on credibility judgments. The results may be interpreted to mean that the purposes of information use do not affect the way people process user ratings or expert ratings in evaluating the credibility of the book of this study. Or, it may be possible that the respondents consider the subject of the book (healthy eating and the environment) as a serious topic, about which the credibility of its information matters, regardless of the purpose of information use. As a result, this may have resulted in not having any effect on credibility judgments.

The exploratory findings of this study are consistent with the literature in that the respondents are aware of or concerned about the possibility of the bias of user ratings. Nonetheless, they find user ratings useful and use them anyway (Flanagin *et al.*, 2014; Reichelt, Sievert and Jacob, 2014). The literature shows that electronic word of mouth (eWOM) satisfies the utilitarian functions of products or services (Reichelt, Sievert and Jacob, 2014). This study shows that user ratings serve such electronic word of mouth functions. In addition, the results may be interpreted to mean that users' perceptions of the bias of user ratings may not be necessarily negative regarding the way they evaluate or use such information. Indeed, such information seems to serve as a piece of information for their credibility judgments. This is consistent with the evidence that online consumers tend to perceive a negative online review as more credible than a positive online review (Kusumasondjaja, Shanka and Marchegiani, 2012; Pan and Chiou, 2011). Further research is needed to examine why and how people use social information that they perceive as biased.

This study has certain limitations, and a few suggestions for further research have emerged from the current study. First, this research employed a convenience sample from a women's college at a small university. This is one of the main limitations of the study. As a result, the findings of the study may not be generalizable to either the population of students at the study site or the population of college students in general. Male students may be different from their female counterparts with respect to the research variables concerning user-generated content. Further research is needed to examine the influence of expert opinions on credibility judgments by employing both sexes. Second, the study employed only one subject area (healthy eating and the environment) as the experimental stimulus. People may perceive user ratings and expert ratings differently for other subjects. Further research is needed to examine a variety of subject areas.

Third, this study did not examine the volume of ratings as a variable. Instead, fixed numbers of low and high volumes were used for all of the experimental conditions. These numbers might have a compounding effect on credibility judgments. Given some inconsistent findings regarding the effect of user ratings volumes (e.g., Lee and Sundar, 2013; Flanagin and Metzger, 2013; Flanagin *et al.*, 2014), this variable needs to be further examined. Finally, it appears that the exploratory data show the importance of the utilitarian functions of user ratings in their use. Further research is needed to examine whether (and if so, under what conditions) the utilitarian values of user-generated content are more important than the content's credibility to users, and why, if such is the case.

Conclusions

This study examined the influence of authority, social endorsement and the purpose of information use on credibility judgments. In particular, expert ratings and user ratings were employed for authority and social endorsement, respectively. Academic versus casual use was manipulated to examine the purposes of use. This study shows that expert ratings influence the credibility of a book and the credibility of user ratings. However, user ratings did not influence the credibility judgments of the book or the credibility of expert ratings. The purpose of use did not jointly vary with user ratings or expert ratings on credibility judgments. The findings suggest that it would be helpful for designers to create certain features to solicit expert opinions, which can enhance users' credibility judgments of user-generated content. Librarians can help users of user-generated content users by providing their professional opinions or by providing a list of expert sources.

Acknowledgements

The authors appreciate Heidi Anderson-Isaacson, Director of Residence Life at St. Catherine University, for her great help with the data collection for this study.

About the authors

Sook Lim is an Associate Professor in the Master of Library and Information Science Program at St. Catherine

University. Her research centers on human information behaviour concerning user-generated content such as Wikipedia. She is particularly interested in credibility issues. For further information about her, visit her website at <http://sooklim.org>. She can be contacted at slim@stkate.edu.

Nick Steffel is an Instructional Technology Support Coordinator with the Master of Library and Information Science Program at St. Catherine University. His professional background is in information technology; he also earned his MLIS from St. Catherine University in 2012. He can be contacted at: njsteffel@stkate.edu.

References

- Austin, L., Liu, B. B. & Jin, Y. (2012). How audiences seek out crisis information: exploring the social-mediated crisis communication model. *Journal of Applied Communication Research*, 40(2), 188-207.
- Cassidy, W.P. (2007). Online news credibility: an examination of the perceptions of newspaper journalists. *Journal of Computer-Mediated Communication*, 12(2). 478-498.
- Dou, X., Walden, J. A., Lee, S. & Lee, J. (2012). Does source matter? Examining source effects in online product reviews. *Computers in Human Behavior*, 28(5), 1555-1563.
- Flanagin, A., Metzger, M., Pure, R., & Markov, A. (2011). User-generated ratings and the evaluation of credibility and product quality in ecommerce transactions. In *Proceedings of the 44th Hawaii International Conference on System Sciences*, (10 p.). Washington, DC: IEEE Computer Society.
- Flanagin, A., Metzger, M., Pure, R., Markov, A. , & Hartsell, E. (2014). Mitigating risk in ecommerce transactions: perceptions of information credibility and the role of user-generated ratings in product quality and purchase intention. *Electronic Commerce Research*, 14(1),1-23.
- Flanagin, A. & Metzger, M. (2013). Trusting expert versus user-generated ratings online: the role of information volume, valence, and consumer characteristics. *Computer in Human Behavior*, 29(4), 1620-1634.
- Flanagin, A. J., & Metzger, M. (2008). Digital media and youth: unparalleled opportunity and unprecedented responsibility. M. Metzger, & A. Flanagin (Editors), *Digital media, youth, and credibility* (pp. 5-28). Cambridge, MA: MIT Press.
- Fu, W. W. & Sim, C. C. (2011). Aggregate bandwagon effect on online videos' viewership: value uncertainty, popularity cues, and heuristics. *Journal of the American Society for Information Science and Technology*, 62(12), 2382-2395.
- Gaziano, C. & McGrath, K. (1986). Measuring the concept of credibility. *Journalism Quarterly*, 63(3), 451-462.
- Hilligoss, B. & Rieh, S. Y. (2008). Developing a unifying framework of credibility assessment: concept, heuristics, and interaction in context. *Information Processing and Management*, 44(4), 1467-1484.
- Hu, Y. & Sundar, S. S. (2010). Effects of online health sources on credibility and behavioral intentions. *Communication Research*, 37(1), 105-132.
- Jensen, J. D. (2008). Scientific uncertainty in news coverage of cancer research: effects of hedging on scientists' and journalists' credibility. *Human Communication Research*, 34(3), 347-369.
- Kim, K., Sin, S. J. & Tsai, T. (2014). Individual differences in social media use for information seeking. *The Journal of Academic Librarianship*, 40(2), 171-178.
- Kusumasondjaja, S. Shankar, T. & Marchegiani, C. (2012). Credibility of online reviews and initial trust: the roles of reviewer's identity and review valence. *Journal of Vacation Marketing*, 18(3), 185-195.
- Lankes, R. D. (2008). Credibility on the Internet: shifting from authority to reliability. *Journal of Documentation*, 64(5), 667 – 686.
- Lee, J., & Sundar, S. (2013). To tweet or to retweet? That is the question for health professional on twitter. *Health Communication*, 28(5), 509-524.
- Lim, S. (2013). College students' credibility judgments and heuristics concerning Wikipedia. *Information Processing & Management*, 49(2), 405-419.
- Lim, S. (2012). Gender differences concerning Wikipedia: a follow-up study. *Proceedings of the American Society for Information Science and Technology Annual Meeting*, 49(1). Retrieved from <https://www.asis.org/asist2012/proceedings/Submissions/323.pdf> (Archived by WebCite® at <http://www.webcitation.org/6WpbKLFA1>)
- Lim, S. & Kwon, N (2010). Gender differences in information behavior concerning Wikipedia, an unorthodox information source? *Library and Information Science Research*, 32(3), 212–220.
- Madathil, K. C., Rivera-Rodriguez, A., J., Greenstein, J. S., & Gramopadhye, A. K. (in press). [Healthcare information on YouTube: a systematic review](#). *Health Informatics Journal*. Retrieved from http://www.researchgate.net/publication/261138299_Healthcare_information_on_YouTube_A_systematic_review (Archived by WebCite® at <http://www.webcitation.org/6WpbZyloy>)
- Messing, S. & Westwood, S. (in press). [Selective exposure in the age of social media: endorsements trump partisan source affiliation when selecting news online](#). *Communication Research*. Retrieved from <http://web.stanford.edu/~seanjw/papers/CRsocialNews.pdf> (Archived by WebCite®)

<http://www.webcitation.org/6WpcJF89O>

- Metzger, M. J. (2007). Making sense of credibility on the Web: models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, 58(13), 2078-2091.
- Metzger, M. J. & Flanagin, A. J. (2011). Using web 2.0 technologies to enhance evidence-based medical information. *Journal of Health Communication*, 16(1), 45-58.
- Metzger, M. & Flanagin, A. (2013). Credibility and trust of information in online environments: the use of cognitive heuristics. *Journal of Pragmatics*, 59(Part B), 210-220.
- Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*, 60(3), 413-439.
- Meyer, P. (1988). Defining and measuring credibility of newspapers: developing an index. *Journalism Quarterly*, 65(3), 567-574.
- Paek, H., Hove, T., & Jeong, H. (2011). Peer or expert? The persuasive impact of YouTube public service announcement producers. *International Journal of Advertising*, 30(1), 161-188.
- Pan, L. & Chiou, J. (2011). How much can you trust online information? Cues for perceived trustworthiness of consumer-generated online information. *Journal of Interactive Marketing*, 25, 67-74.
- Pure, R. A., Markov, A. R., Mangus, J. M., Metzger, M. J., Flanagin, A. J., & Hartsell, E. H. (2013). Understanding and evaluating source expertise in an evolving media environment. In T. Takševa (Ed.), *Social software and the evolution of user expertise: future trends in knowledge creation and dissemination* (pp. 37-51). Hershey, PA: Information Science Reference.
- Reichelt, J., Sievert, J. & Jacob, F. (2014). How credibility affects eWOM reading: the influences of expertise, trustworthiness, and similarity on utilitarian and social functions. *Journal of Marketing Communications*, 20(1-2), 65-81.
- Rieh, S. Y. (2010). Credibility and cognitive authority of information. M. J. Marcia J. Bates, & M. N. Maack (editors), *Encyclopedia of library and information science* 3rd ed. (pp. 1337-1344). New York, NY: CRC Press.
- Rieh, S. & Hilligoss, B. (2008). College students' credibility judgments in the information seeking process, In M. J. Metzger & A. J. Flanagin (Eds), *Digital media, youth, and credibility* (pp. 49-72). Cambridge, MA: The MIT Press.
- Rieh, S. Y., & Danielson, D. R. (2007). Credibility: a multidisciplinary framework. *Annual Review of Information Science and Technology*, 41, 307-364.
- Schwarz, J., & Morris, M. R. (2011). Augmenting Web pages and search results to support credibility assessment. In *Proceedings of the 29th International Conference Extended Abstracts on Human Factors in Computing Systems (CHI'11)* (pp. 1245-1254). New York, NY: ACM Press.
- Self, C. C. (2009). Credibility. In D. W. Stacks, & M. B. Salwen (Eds), *An integrated approach to communication theory and research*, 2nd ed. (pp. 435-456). New York, NY: Routledge.
- Stavrositu, C. & Sundar, S. (2008). If Internet credibility is so iffy, why the heavy use? The relationship between medium use and credibility. *CyberPsychology & Behavior*, 11(1), 65-68.
- Sundar, S. S. (2008). The MAIN model: a heuristic approach to understanding technology effects on credibility. In M. J. Metzger, & A. J. Flanagin (Eds.), *Digital media, youth, and credibility* (pp. 73-100). Cambridge, MA: MIT Press.
- Sundar, S. S., Xu, Q., & Oeldorf-Hirsch, A. (2009). Authority vs. peer: how interface cues influence users. *Proceedings of the 27th International Conference Extended Abstracts on Human Factors in Computing Systems (CHI'09)* (pp. 4231-4236). New York, NY: ACM Press.
- Tseng, S., & Fogg, B. J. (1999). Credibility and computing technology. *Communications of the ACM*, 42(5), 39-44.
- Tsfati, Y., & Cappella, J. N. (2005). Why do people watch news they do not trust? the need for cognition as a moderator in the association between news media skepticism and exposure. *Media Psychology*, 7, 251-171.
- Xu, Q. (2013). Social recommendation, source credibility, and recency: effects of news cues in a social bookmarking website. *Journalism & Mass Communication Quarterly*, 90(4), 757-775.
- Wang, Z., Walther, J. B., Pingree, S., & Hawkins, R. P. (2008). Health information, credibility, homophily, and influence via the Internet: Websites versus discussion groups. *Health Communication*, 23(4), 358-368.