

# User-Generated Ratings and the Evaluation of Credibility and Product Quality in Ecommerce Transactions

**Andrew J. Flanagin**

**Miriam J. Metzger**

**Rebekah Pure**

**Alex Markov**

Department of Communication, University of California, Santa Barbara  
Santa Barbara, CA 93106

[flanagin@comm.ucsb.edu](mailto:flanagin@comm.ucsb.edu), [metzger@comm.ucsb.edu](mailto:metzger@comm.ucsb.edu), [rebekah.pure@gmail.com](mailto:rebekah.pure@gmail.com), [markov@umail.ucsb.edu](mailto:markov@umail.ucsb.edu)

## Abstract

*Although extremely popular, electronic commerce transactions often lack information that has traditionally served to ensure trust and credibility among exchange partners. The capacity of digital media to aggregate information and connect individuals to one another, however, offers new potential for determining information quality and credibility. To understand people's perceptions of the credibility of commercial web site information and the factors they find important in their evaluative processes, a nationally representative survey and quasi-experiment were conducted. Survey results showed that while people engage in ecommerce regularly, they do not contribute consumer-generated information very often. They do, however, rely heavily on ratings to evaluate the credibility of commercial information they find online. Experimental results further indicate that people attend to aggregate product ratings, but not to the number of ratings when evaluating the quality of products sold online. We conclude with implications of these findings for system designers.*

## 1. The Prevalence and Risk of Online Commercial Transactions

People are increasingly relying on web-based commercial information for electronic commerce ("ecommerce") transactions that range from small personal items to home purchases [1]. Retail ecommerce sales in the U.S. currently constitute roughly four percent of total retail sales, which translates to almost 40 billion dollars annually [2]. Indeed, the number of Americans who have purchased a product online has steadily increased since 2000, and a majority (66%) now report having

made at least one online purchase [3]. An even larger percentage (93%) has used the Internet for ecommerce-related activities, including researching information about a product they are thinking of buying, with more than a quarter of Americans reporting they do this on a daily basis [3]. In fact, over the last decade the number of people either researching or buying a product or service online has nearly doubled, a trend that holds true across a wide range of ecommerce-related activities [1].

Given this level of online commercial activity, it is important to understand people's perceptions of the credibility of commercial web site information and the factors they find important in their evaluative processes. This is particularly critical given that, despite its popularity, online commercial transactions often lack elements that have traditionally served to ensure trust and credibility among parties. As Hawes and Lumpkin [4] note, different levels of risk accompany various "patronage modes," with risk increasing as the consumer is separated from the physical presence of the retail store. Indeed, information asymmetries between buyers and sellers, privacy concerns, system security shortcomings, immature legal protection mechanisms, low investment in infrastructure, the inability to sufficiently inspect goods prior to purchase, and fraud are among the many risks faced by online consumers today [5-7].

Research has shown that risk may be reduced by relevant commodity information [8-10], through vicarious experience that may serve as the basis of personal trust [11], and from system or structural features that serve to bolster trust in parties by means of external assurances [11, 12]. In addition, Bailey et al. [12], note that feedback profiles, achievement awards and other information about sellers, third-party payment options and insurance, and product

descriptions can reduce uncertainty and increase trust in unknown others in commercial transactions. Thus, although online commercial transactions create a risky environment for users, they may be mitigated to a degree by certain features designed to promote greater trust among parties [13].

Accordingly, this study investigates ecommerce activity by assessing consumers' engagement with and use of ratings information as a means of evaluating the credibility of commercial information on the Web. It also examines consumers' use of ratings in their product quality evaluations, which is an important, yet overlooked, variable in both the ecommerce and credibility literatures. To do so, we present distinct yet complementary results of a national survey of U.S. Internet users and an embedded quasi-experiment. Results provide a solid foundation upon which system scientists can build tools and applications that will be both effective and valuable to online consumers. Moreover, understanding *how* users cognitively process ratings information will be useful when developing systems to facilitate effective decision-making in ecommerce-related activities.

## 2. User-Generated Content as a Risk Mitigation Mechanism

The capacity of digital media to aggregate information and to connect individuals to one another offers new potential for determining information quality and credibility [14]. The potential for peer-to-peer information assessment, for example, can be seen in numerous web-based applications, ranging from "credentialing" activities present on eBay or Amazon, to topical discussion forums in which individuals provide firsthand experiences and opinions on a host of topics and products. The ubiquity of user-generated evaluative content online is seen in the proliferation of sites like Epinions ([www.epinions.com](http://www.epinions.com)), which offers a repository of consumer ratings on a range of products. Other sites offer extensive collections of ratings in more specialized domains, be they technical (CNET, [www.cnet.com](http://www.cnet.com)), cinematic (IMDb, [www.imdb.com](http://www.imdb.com)), academic (Rate My Professors, [www.ratemyprofessors.com](http://www.ratemyprofessors.com)), or social (Hot or Not, [www.hotornot.com](http://www.hotornot.com)).

Such "user-generated content" makes it easier for individuals to harness collective intelligence to help them assess and evaluate information they find online. In the words of Madden and Fox [15, pp. 2], so-called Web 2.0 tools and applications can "replace

the authoritative heft of traditional institutions with the surging wisdom of crowds." Indeed, recent studies have found that social and group-based information evaluation processes are very important as Internet users frequently turn to online ratings and reputation systems to help them assess information or its source [14, 16]. Bottom-up assessments of information quality constructed through collective or community efforts (e.g., wikis, ratings and reputation systems, or social networking applications) serve as a form of metadata about information found online, and may be emerging as new arbiters of credibility, authority, and trust in that context.

Ecommerce is among the many domains in which user-generated content has dramatically affected individuals' knowledge capacity. In the context of ecommerce transactions, risk can be mitigated by reducing uncertainty and increasing trust, which can be achieved via a host of strategies for information exchange among interested consumers [17]. For example, trust in online auctions is typically gained through people sharing their experiences with the product or party in question via electronic means such that "vicarious experience might [even] allow for the problem-free experiences of others to become the basis for one's own trusting behavior" [11, pp. 399]. Shimp and Bearden [9] found that consumers rely on product usage experience or word-of-mouth assurances to reduce the level of risk associated with product purchases [see also 8], and additional research shows that personal trust is developed as a result of feedback mechanisms that enable members of a user community to share their collective experiences [see 18 for a related argument].

The few trust cues that are available in online interactions become particularly salient and important in forming attitudes and determining behaviors, which is consistent with several theories of computer-mediated communication [e.g., the SIDE model, 19, 20, the "hyperpersonal" perspective, 21, and social information processing, 22]. Thus, the structure of the Internet "substitutes a much better distribution of what information there is for the much more limited, but more reliable information of traditional retail markets" [7, pp. 316], which enables consumers to reduce uncertainty and therefore risk about sellers or products, even in the absence of firsthand experience with them. In essence, the aggregated, user-generated information shared among unknown or anonymous online exchange partners serves to mitigate risk and bolster the trust that is potentially tenuous in online commercial transactions.

## 2.1. Commercial Product Ratings

Commercial product ratings are a particular instance of user-generated content that has exploded in recent years [1, 23]. Although ratings span a host of venues, and have been applied to everything from people's opinions, to assessments of whether messages are spam, to ratings of fellow raters, their predominant usage in ecommerce is to assess product quality, which is the focus of our quasi-experiment, as described later.

Changes in the nature of both web content and its providers now require research exploring how individuals use *social* and *group-based* information to arrive at credibility decisions. While some studies have examined the impact of product reviews and ratings on purchase decisions [24-28], almost no research has sought to understand how individuals employ or access social collaborative efforts such as ratings to evaluate the quality of commercial information online. To fill this void, we assess the extent to which consumers engage in ecommerce transactions, and the degree to which they rely on product ratings in evaluating the credibility of commercial information online. Thus, the first research questions of this study are:

**RQ1:** To what extent are users engaging in ecommerce activities, including rating products and services online?

**RQ2:** What is the relative importance of ratings information in helping users establish the credibility of commercial information online?

Existing research on consumer-generated content and credibility with regard to ecommerce primarily examines ratings and reviews (i.e., consumer testimonials). An important focus in this literature is on how reviewers establish credibility, and how consumers evaluate reviewers' credibility. For example, Hu, Liu, and Zhang [29] found that when consumers read online reviews, they pay attention to contextual information such as reviewers' reputation and exposure. Reviewers have been shown to establish their expertise through assertions of their experience with products, and through careful use of proper and appropriate language [30]. However, in spite of negative structural indicators such as spelling errors that have been shown in past research to diminish perceived credibility, reviewers are still rated as credible when they assert reasons why their opinions should be trusted [30]. Consumers tend to care most about review quality under conditions of high personal interest in the product [31], and they focus on information about the reviewer's identity more than the review itself when evaluating the credibility of the review, as well as the product [32].

In fact, reviewers are voted as most "helpful" on Amazon when they provide detailed information about their offline identity [32]. Furthermore, consumers pay more attention to reviewer characteristics when review volume is very high, perhaps as a way to heuristically process information when they cannot effectively process all reviews [32].

Considering research that focuses on product ratings specifically, Mackiewicz [30] examined the ratings of 640 online products and found that nearly half of all products rated received 5 stars (i.e., the highest ranking), suggesting a positivity bias in ratings. Similarly, Resnick and Zeckhauser [7] have noted that eBay user feedback is overwhelmingly positive and that net feedback ratings reported by eBay may encourage overly optimistic assessments of others [see also 33], and Hu, Pavlou, and Zhang [34] found that 53 percent of products show a bimodal ratings distribution, indicating that for most products ratings are only very positive and very negative. This suggests that the average rating of a product does not necessarily reflect its true quality, but rather the opinions of consumers who were either very pleased or very disappointed with it. Thus, Hu et al. [34] argue that ratings on the whole are biased, and therefore should not be viewed as credible sources of consumer information.

Nevertheless, ratings have been shown to be important in people's purchasing decisions. For example, negative user ratings have a disproportionate, and negative, influence on bid prices in online auctions [35]. Similarly, Chevalier and Mayzlin [26] examined consumer ratings on Amazon and on the Barnes and Noble website and found that while most ratings were positive, 1-star reviews had a greater impact than 5-star product reviews. Additionally, they found that a book's average rating score was correlated with sales, although the directionality of this effect was not tested. Duan et al. [27], however, found that movie box office sales were positively correlated with the volume of online ratings and reviews, but not with the ratings themselves, suggesting that perhaps popularity predicts an increase in sales and more online ratings, rather than online ratings influencing sales.

A majority of the research on consumer-generated content with regard to ecommerce has focused on reviewers and their reviews, with an emphasis on the impact of consumer-generated information on sales, rather than on ratings themselves. Additionally, the research that has examined the impact of ratings has done so correlationally rather than experimentally, which

inhibits researchers from making causal claims about the impact of ratings. Finally, researchers in this area suggest that reviews and ratings might have different effects depending on other factors involved in the decision process, but little research has investigated this to date. For example, Shen [36] argues that prior experience with the product or brand may influence the extent to which reviews matter in the purchasing decision.

Given the lack of research on how people understand and use ratings information, and given the recent attention paid in studies of web-based information credibility to the influence of others' opinions in credibility assessments [14, 16], we focused on the way in which product ratings inform people's quality assessments of commercial products. Past research has overlooked perceptions of product quality as a variable of interest, choosing instead to focus on product sales, for example. This study treats perceptions of product quality as a distinct outcome of information credibility assessment in the context of online commercial information, albeit one that is fundamentally related to purchase decisions.

Product ratings consist of two separate but related dimensions. First, ratings typically span some scale or range between endpoints indicating low and high quality. Often, heuristic cues, such as "stars" representing a numeric average rating, serve to summarize such aggregated ratings data, in order to facilitate information processing. Second, because ratings are tallied among members of some user group (e.g., those who have firsthand experience with a product and therefore are qualified to render an assessment of it), there is typically also an indication of the number of raters providing assessments.

Theoretically, each of these dimensions is a necessary, but insufficient, indicator of product quality. For example, one disgruntled consumer has a very large impact on the overall rating of a product when there are only four ratings and a very small impact when there are over 1,000, so decoupling the number of ratings from the average rating could reflect a critical deficiency in people's ability to correctly interpret the meaning of online ratings. Therefore, we endeavored to assess the effects of both the nature and number of ratings on perceived product quality, as stated in Research Question 3:

**RQ3:** To what degree do people attend to the relevant and appropriate online credibility cues (i.e., the nature and number of peer product ratings) in making assessments of product quality?

### 3. Method

#### 3.1. Sample and General Procedure

Data for this study were collected as part of an online survey conducted by the professional research firm Knowledge Networks, which maintains a probability-based panel of participants that is representative of the entire U.S. population. The sample from this study was drawn from a combination of random digit dialing and address-based sampling methods which allows Knowledge Networks to reach cell-phone only homes, do-not-call listed homes, and homes that use call-screening that normally would be missed by random digit dialing methods alone.

The survey was fielded in late-2009 and included 3,991 adults in the U.S. who use the Internet. Subjects took the survey from wherever they typically accessed the Internet, at their leisure, in order to maintain as naturalistic an environment as possible. Descriptive data on usage behaviors, described more fully in the Descriptive Results subsection below, come from this entire sample. These data were weighted such that results are generalizable to all adult Internet users in the United States.

#### 3.2. Experimental Conditions

A subsample of these respondents ( $N = 2,139$ ) was randomly selected to participate in a quasi-experimental portion of the survey, which directly assessed the effect of user-generated ratings of commercial information online. To gauge the degree to which people attend to the nature and number of user-generated product ratings as credibility cues for commercial information online, and the factors that influence their evaluations, each of these subjects was presented with one screenshot from a set of product pages on Amazon.com, followed by questions about the product they had viewed. Product pages were modified to maintain comparability across items and to meet technical requirements for the study. Three different products were shown (a digital camera, an electric toothbrush, and rolling luggage) in order to gauge stimulus generalizability. Because analyses showed that people's interest in each product varied (as assessed by the question "How interested are you in buying this type of product?"), we statistically controlled for interest in the product in all subsequent analyses. Additionally, we controlled

for the perceived importance of the product manufacturer, since past research [35] has shown that prior brand experience may influence the extent to which reviews matter in the purchasing decision.

In addition to the product, two factors were varied systematically by altering information on the web page screenshots used in the study: (a) the *number* of ratings provided about the product and (b) the *average* rating provided about the product by other users. Specifically, the pages showed the number of user ratings as 4, 16, 102, or 1002 and average “star” ratings (on a 1-5 scale, where 5 is the best rating) of 1.6, 2.23, 3.0, 3.68, 4.4, 4.84, or 5.0. In this manner, 84 different page image stimuli were presented in the experiment, representing each possible combination of number of ratings, average ratings, and product. All other content was held constant across all pages. Figure 1 shows an example page with a digital camera presented as receiving an average rating of 3.68, across 102 total ratings.

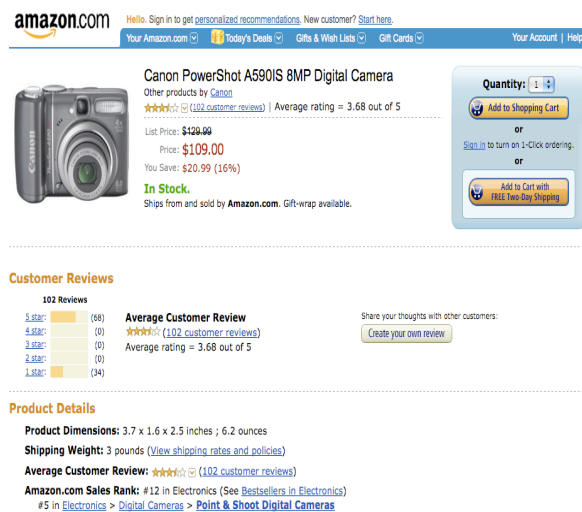


Figure 1. Example product web page

Because analyses showed that the dependent variable did not depend on the specific product, product was collapsed into one factor for all analyses. The quasi-experiment therefore took the form of a 4 (*number* of ratings: 4, 16, 102, or 1002) by 7 (*average* rating: 1.6, 2.23, 3.0, 3.68, 4.4, 4.84, or 5.0) factorial design, where subjects were randomly assigned to one condition. Although the conditions were not exhaustive, we believe them to be representative of common types of ratings distributions online. The dependent variable of *product quality* was assessed by the question "On a scale of 0 to 10, how would you rate the quality of

this product?" with response categories ranging from 0 = "The quality is *much worse* than average" to 10 = "The quality is *much better* than average," with the scale midpoint of 5 defined as "The quality is *about average*" ( $\bar{X} = 6.79$ ,  $SD = 2.31$ ). The results of the quasi-experiment were used to evaluate Research Question 3, and in conjunction with questions from the rest of the survey, to gauge participants' use of ratings as a credibility cue.

## 4. Results

### 4.1. Descriptive Results

The first research question concerned the degree to which people engage in ecommerce activities, including generating ratings. Respondents reported that overall they buy things online "sometimes," as demonstrated by a mean score of 2.83 ( $SD = 1.04$ ) to the question "How often do you buy something online?" (on a 5-point scale ranging from 1 = never to 5 = very often). People tend to rate things online less often than this, with an average score of 1.85 ( $SD = 1.01$ ) on the same scale. This means that, on average, users report that they rate something between "rarely" and "never."

Despite the low frequency of actually contributing rating information themselves, user-generated information appears to be an important credibility cue to people when shopping online. The second research question probed the relative importance of ratings information in helping users establish the credibility of commercial information online. A randomly selected subset of all respondents ( $N = 766$ ) was asked to imagine they were buying something on the Internet and to assess the importance of various factors for determining the credibility of the information they found. Among these, people indicated that others recommending a website or information source was somewhere between "somewhat important" to "important" in their decision ( $\bar{X} = 3.45$ ,  $SD = 1.00$ ; on a 5-point scale ranging from "not at all" to "very" important), while the existence of "high ratings, positive comments, or good reviews" was even more important to them ( $\bar{X} = 3.71$ ,  $SD = 1.04$ ) in determining credibility. Relative to the other credibility cues, ratings were ranked highly, and were only less important than website security, and the currency and completeness of the information given on commercial websites to users in establishing credibility. Ratings were judged to be significantly more important than such

credibility cues as source reputation and expertise; familiarity with and popularity of website; and webpage design, accuracy of the information provided, and ease of use. Table 1 shows these results.

**Table 1. Means and standard deviations for cues used to determine credibility of commercial information online**

	$\bar{X}$	<i>SD</i>
the website seems safe and secure	4.19	.96
the information on the website is up-to-date	4.02	.90
the information is very complete	3.91	.91
there are high ratings, positive comments, or good reviews	3.71 <sub>a</sub>	1.04
the website is easy to use	3.71 <sub>a</sub>	1.08
the information seems reasonable to you	3.70 <sub>a</sub>	.88
the information is well written and you see no typing mistakes	3.64 <sub>b</sub>	1.14
you get more than just one person's opinion	3.56 <sub>b</sub>	1.03
you have heard good things about the information source or website creator	3.53 <sub>b</sub>	1.04
experts believe the information (like your doctor, teacher, etc.)	3.49 <sub>b</sub>	1.02
the information is from an expert on the topic	3.48 <sub>b</sub>	1.02
others recommend the website or information source	3.45 <sub>b</sub>	1.00
you have heard of the source or information creator before	3.31 <sub>c</sub>	1.06
the information on the website is similar to information on other websites	3.29 <sub>c</sub>	1.10
people you know, such as friends and family, believe the website or information source	3.28 <sub>c</sub>	1.13
the information you find is similar to what you already think	3.21 <sub>c</sub>	1.05
a lot of other people use the website	3.10	1.20
the website looks good	3.00	1.14
you just like the website	2.64	1.11

*Note:* Means with common subscripts do not differ significantly from one another.

Relatedly, when examining only those people who participated in the quasi-experiment, participants indicated that customer reviews and ratings were important in making online purchasing decisions ( $\bar{X}$  = 3.84, *SD* = 1.00, on a 5-point scale where higher values indicate greater importance). Ratings were nearly equally important as information about who makes the product ( $\bar{X}$  = 3.77, *SD* = 1.01) and product details such as color, size, and the like ( $\bar{X}$  = 3.79, *SD* = 1.03). Price information, however, was the most important factor ( $\bar{X}$  = 4.43, *SD* = .76) and who sells the product was the least important factor ( $\bar{X}$  = 3.33, *SD* = 1.07) in online purchasing decisions, perhaps indicating a lack of loyalty to any particular vendor.

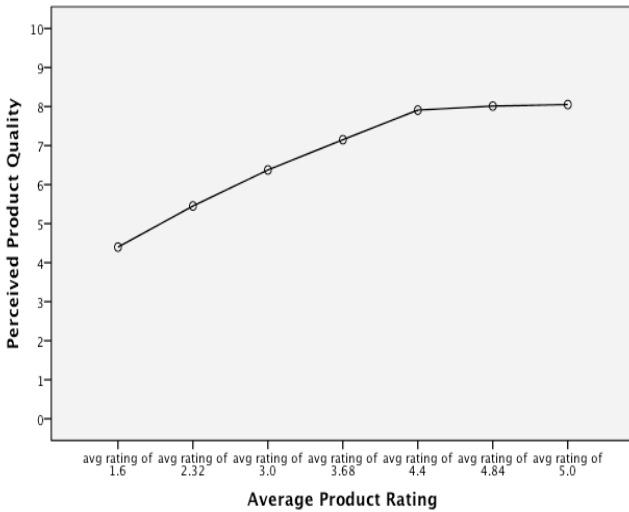
#### 4.2. Quasi-Experiment Results

In order to assess the degree to which people attend to the relevant and appropriate cues (i.e., the number and nature of ratings) in determining product quality as posed in the last research question, a 4 (*number* of ratings) by 7 (*average* rating) ANCOVA analysis was performed, with interest in buying the product and the importance of the product manufacturer as the covariates, and the perceived quality of the product as the dependent measure. As mentioned earlier, the three product types were collapsed for this analysis.

Results showed no main effect for the number of ratings, but a significant main effect for average rating ( $F[6, 2091] = 196.46, p < .001, partial \eta^2 = .36$ ) and an interaction effect between average rating and the number of ratings that approached, but did not quite reach, significance ( $p = .056, partial \eta^2 = .01$ ). Notably, the majority of the variance explained was due to the main effect for average rating, and not to the marginally significant interaction among the independent variables.

As the average ratings increased, so too did subjects' assessment of product quality, with ratings ranging from a mean value of 4.28 (*SD* = 2.33) when the average star rating was 1.6, to 8.23 (*SD* = 1.95) when the star rating was 5.0. These results are illustrated in Figure 2. Follow-up tests indicated that nearly all mean values of the seven individual star ratings levels were statistically different from one another at the  $p < .001$  level. The exceptions to this occurred only at the high end of the star ratings scale, where ratings of 4.4, 4.84, and 5.0 did not differ on perceived product quality from one another. This suggests a "ceiling effect," whereby ratings have

diminished positive effects on perceived product quality as they approach the top of the ratings scale.



**Figure 2. Product quality ratings by average rating**

## 5. Discussion

Results of our study indicate that while Internet users in the U.S. engage in ecommerce transactions with some frequency, they tend to provide ratings on their commercial endeavors rather infrequently. Moreover, they report that user-generated information in the form of ratings and recommendations from others is important to them in their credibility assessments, although experimental evidence from our study indicates that such information is used suboptimally. Overall, it appears that ecommerce is an important dimension of people’s web use and that user-generated commercial information serves to mitigate the risk inherent in this reduced cues environment by providing a sort of metadata, although not always in the most effective manner.

More specifically, results of Research Question 1 indicated that although people are engaging in ecommerce, they are not providing consumer-generated content in the form of ratings to the same extent. This is consistent with research from 2004 and 2008 showing that, on the whole, people are not taking full advantage of the opportunity to participate in Web 2.0 tools [23], indicating a relative invariance in participation over time despite increasing opportunities for providing consumer-generated feedback. These results are also consistent with a great deal of past research demonstrating that

individuals are content to free-ride on the contributions of others to information public goods [37, 38], particularly under conditions of low visibility, when responsibility is diffused [39]. From a systems perspective, one challenge thus becomes producing tools that encourage users to rate products easily, consistently, and habitually, in spite of the temptation to free-ride on others’ efforts.

Research reviewed earlier also suggests that those who provide product ratings do so because they are either very happy with the product or very unhappy with the product, and thus ratings are likely to be polarized [34], which could potentially diminish their usefulness to consumers, and thus their credibility as a source of metadata. This suggests that free-riding in this situation becomes especially problematic, since extreme views are disproportionately being reported. Again, systems science solutions aimed at engaging a wider variety of consumers’ opinions become critical under these circumstances.

With regard to the second research question, results showed that information from others and consumer-generated product ratings are important in helping people determine the credibility of commercial information online. People place most importance on evaluating whether commercial website information is secure, up-to-date, and complete when determining the credibility of online commercial information, but next they rely on product ratings, comments, and reviews (among other factors) to make decisions about credibility and whether or not to purchase a product. This finding is consistent with Metzger et al. [16], who found that people are more likely to believe information about a product, and indeed to buy a product, when the product has high ratings.

Results of the quasi-experiment (Research Question 3) showed that people carefully attend to average product ratings when evaluating product quality, but do not focus on the number of ratings provided. In fact, even when the number of ratings was hugely disparate (i.e., 4 ratings compared to 1002 reviews) within an average rating, there was no difference in perceived product quality. This is clearly suboptimal for consumers, since by ignoring such data users neglect a great deal of relevant information useful in making informed decisions. This is consistent, however, with recent research that finds users routinely employ cognitive heuristics, or mental shortcuts, in judging the credibility of information online as a way to cope with the overwhelming task of methodically evaluating all of the information cues available on a website or set of websites returned from a given search query [16]

Additionally, data from this study show that once the average rating for a product reached a certain level (4.4 out of 5 stars in our study), a ceiling effect was evident, such that ratings beyond this threshold did not translate to greater perceptions of product quality. This again suggests that consumers are neglecting information that could help them to further discern product quality more accurately. To address these information processing shortcomings, systems scientists could focus on creating tools that encourage users to process the full range of relevant information, rather than just focusing on certain aspects of aggregate ratings data. Findings of the quasi-experiment, however, should be interpreted with the recognition that the conditions of in the current study were not exhaustive, and future research may benefit from examining a wider range of products under more diverse ratings conditions.

### 5.1. Implications for Systems Scientists

Results of our study suggest two main courses of action for researchers interested in developing systems to facilitate effective decision-making in ecommerce-related activities. At a more micro level, it is important to recognize that a disconnect occasionally exists in users' minds between the average rating of a product and the number of ratings that product receives, as demonstrated by the results of Research Question 3. To address this, system designers could build mechanisms into ecommerce web sites that make these discrepancies more salient. For instance, increasing the font size of the average rating on the web page based on the number of ratings from which it is derived, such that ratings from a larger pool of reviews appear more prominent, might draw consumers' attention to a gap where it exists. Similarly, web sites could flag products that have not yet received a certain critical mass of reviews, in order to warn consumers that the credibility of information about the product may be suspect, much like how short Wikipedia entries may be flagged as "stubs." Alternatively, rather than averaging the ratings into one rating score, commercial web sites could display ratings in the form of histograms, so that consumers can see the distribution of user ratings, including bi-modal product ratings. In this way consumers may be encouraged to incorporate *both* dimensions of rating systems into their processing, ensuring that average ratings are placed in their proper context.

The second, more holistic, approach to improving these systems recognizes that most product ratings are bimodal [34], and people's unwillingness to rate products calls the credibility of these aggregate

ratings into question. System scientists motivated to capitalize on the knowledge shared by large groups of people should therefore encourage users to rate products more regularly, perhaps by better incorporating the rating process into the ecommerce experience, and thus making it easier and more efficient for users to supply ratings. For instance, ecommerce web sites, such as Amazon.com, could greet consumers with a splash screen displaying their recently purchased products and encourage them to rate the products before continuing on to the web site. Additionally, these sites might offer incentives for users to rate the products they purchase by providing coupons or discounts on future purchases when users submit ratings. Moreover, they could devote a section of their web page to profiling popular reviewed items, in order to capture what products people are interested in and provide information to consumers about what a large group of individuals are reviewing. Finally, some evidence suggests that people can be motivated to contribute to ratings systems by being made aware of their rating behavior relative to other users [40], which could easily be programmed into ratings presentation. All of these strategies would help to ensure that the average ratings for these products are coming from a large and diverse set of reviews, and perhaps encourage consumers to opt into the rating system themselves.

In the end, the credibility of any information, and particularly that of ratings systems and other types of user-generated content, lies at the intersection of the quality of the information provided by the source (in this case, a large base of users), and on the quality of recipients' proper interpretation of that information. Most research has focused on the former, but system designers stand to benefit from research that provides a better understanding of both ends of this spectrum.

## 6. Acknowledgments

The authors thank the John D. and Catherine T. MacArthur Foundation for their generous support of this work.

## 7. References

- [1] J. Horrigan. (2008, June 8, 2010). *The Internet and consumer choice*. Available: <http://www.pewinternet.org/Reports/2008/The-Internet-and-Consumer-Choice.aspx>
- [2] C. Bureau, "Statistical Abstract of the United States: 2010," C. Bureau, Ed., ed. Washington, DC, 2010.
- [3] J. Horrigan. (2008, June 8, 2010). *Online shopping*. Available:



- <http://www.pewinternet.org/Reports/2008/Online-Shopping.aspx>
- [4] J. Hawes and J. Lumpkin, "Perceived risk and the selection of a retail patronage mode," *Journal of the Academy of Marketing Science*, vol. 14, pp. 37-42, 1986.
- [5] Z. Lee, *et al.*, "The effect of negative buyer feedback on auction prices in Internet auction markets," in *21st International Conference on Information systems*, Brisbane, Australia, 2000, pp. 286-287.
- [6] A. Miyazaki and A. Fernandez, "Consumer perceptions of privacy and security risks for online shopping," *Journal of Consumer Affairs*, vol. 35, pp. 27-44, 2001.
- [7] P. Resnick and R. Zeckhauser, "Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system," in *Advances in Applied Microeconomics: A Research Annual*. vol. 11, M. Bayne, Ed., ed Amsterdam: Elsevier Science, 2002, pp. 127-157.
- [8] T. Roselius, "Consumer rankings of risk reduction methods," *Journal of Marketing*, vol. 35, pp. 56-61, 1971.
- [9] T. Shimp and W. Bearden, "Warranty and other extrinsic cue efforts on consumer risk perceptions," *Journal of Consumer Research*, vol. 9, pp. 38-46, 1982.
- [10] S. Tan, "Strategies for reducing consumers' risk aversion in Internet shopping," *Journal of Consumer Marketing*, vol. 16, pp. 163-178, 1999.
- [11] J. Boyd, "The rhetorical construction of trust online," *Communication Theory*, vol. 13, pp. 392-410, 2003.
- [12] B. P. Bailey, *et al.*, "Trust in cyberspace," in *Human Factors and Web Development*, J. Ratner, Ed., ed London: Erlbaum, 2003, pp. 311-321.
- [13] U. Brinkman and M. Siefert, "Face-to-interface - The establishment of trust in the Internet: The case of e-auctions," *Journal of Sociology*, vol. 30, pp. 23-47, 2001.
- [14] A. Flanagin and M. Metzger, "Digital media and youth: Unparalleled opportunity and unprecedented responsibility," in *Digital Media, Youth, and Credibility*, M. Metzger and A. Flanagin, Eds., ed Cambridge: MIT Press, 2008, pp. 5-27.
- [15] M. Madden and S. Fox. (2006 *Riding the waves of Web 2.0*. Available: <http://pewresearch.org/pubs/71/riding-the-waves-of-web-20>
- [16] M. Metzger, *et al.*, "Social and heuristic approaches to credibility evaluation online," *Journal of Communication*, vol. 60, pp. 413-439, 2010.
- [17] A. Flanagin, "Commercial markets as communication markets: Uncertainty reduction through mediated information exchange in online auctions," *New Media & Society*, vol. 9, pp. 401-423, 2007.
- [18] J. Boyd, "In community we trust: Online security communication at eBay," *Journal of Computer-Mediated Communication*, vol. 7, 2002.
- [19] R. Spears and M. Lea, "Social influence and the influence of the "social" in computer mediated communication," in *Contexts of Computer-Mediated Communication*, M. Lea, Ed., ed London: Harvester- Wheatsheaf, 1992.
- [20] R. Spears and M. Lea, "Pancea or panopticon? The hidden power in computer-mediated communication," *Communication Research*, vol. 21, pp. 427-459, 1994.
- [21] J. Walther, "Computer-mediated communication: Impersonal, Interpersonal, and hyperpersonal interaction," *Communication Research*, vol. 23, pp. 3-43, 1996.
- [22] J. Walther, "Interpersonal effects in computer-mediated interaction: A relational perspective," *Communication Research*, vol. 19, pp. 53-90, 1992.
- [23] L. Rainie and P. Hitlin. (2004, June 3, 2010). *Use of online rating systems*. Available: <http://www.pewinternet.org/Reports/2004/use-of-online-rating-systems.aspx>
- [24] D. Godes and D. Mayzlin, "Using online conversations to study word-of-mouth communication," *Marketing Science*, vol. 23, pp. 545-560, 2004.
- [25] P. Chen, *et al.*, "The impact of online recommendations and consumer feedback on sales.," *Proceedings of the International Conference on Information Systems*, pp. 711-724, 2004.
- [26] J. A. Chevalier and D. Mayzlin, "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research (JMR)*, vol. 43, pp. 345-354, 2006.
- [27] W. Duan, *et al.*, "Do online reviews matter?--An empirical investigation of panel data," *Decision Support Systems*, vol. 45, pp. 1007-1016, 2008.
- [28] Y. Liu, "Word of mouth for movies: Its dynamics and impact on box office revenue," *Journal of Marketing*, vol. 70, pp. 74-89, 2006.
- [29] N. Hu, *et al.*, "Do online reviews affect product sales? The role of reviewer characteristics and temporal effects," *Information Technology and Management*, vol. 9, pp. 201-214, 2008.
- [30] J. Mackiewicz, "Reviewer Bias and Credibility in Online Reviews," in *Association for Business Communication Annual Convention*, 2007.
- [31] S. Lee, "How do online reviews affect purchasing intention?," *African Journal of Business Management*, vol. 3, pp. 576-581, 2009.
- [32] C. Forman, *et al.*, "Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets," *Information Systems Research*, vol. 19, pp. 291-313, 2008.
- [33] C. Dellarocas, "Analyzing the economic efficiency of eBay-like online reputation

- reporting mechanisms," in *ACM Conference on Electronic Commerce*, Tampa, Florida, 2001.
- [34] N. Hu, *et al.*, "Can online reviews reveal a product's true quality?: Empirical findings and analytical modeling of Online word-of-mouth communication," in *Conference on Electronic Commerce*, Ann Arbor, Michigan, 2006, p. 330.
- [35] S. Standifird, "Reputation and e-commerce: eBay auctions and the asymmetrical impact of positive and negative ratings," *Journal of Management*, vol. 27, pp. 279-295, 2001.
- [36] W. Shen, "Essays on online reviews: The relationships between reviews, reviews, and product sales, and the temporal patterns of online reviews. ," Management, New York University, New York, 2008.
- [37] R. Friedhorsky, *et al.*, "Creating, destroying, and restoring value in wikipedia," in *International ACM conference on Supporting group work table of contents*, Sanibel Island, Florida, USA, 2007, pp. 259-268.
- [38] C. Cheshire and J. Antin, "The Social Psychological Effects of Feedback on the Production of Internet Information Pools," *Journal of Computer-Mediated Communication*, vol. 13, pp. 705-727, 2008.
- [39] M. Olson, *The logic of collective action*. Cambridge: Harvard University Press, 1985.
- [40] Y. Chen, *et al.*, "Social comparisons and contributons to online communities: A field experiment on MovieLens," *Computational Social Systems and the Internet*, 2007.