Reviewer Bias and Credibility in Online Reviews

Jo Mackiewicz

Illinois Institute of Technology

Abstract

This paper examines a sample of 640 online reviews of products and services to see whether reviews exhibited positive bias; indeed, over 48 percent of the 640 reviews that were sampled from two Web sites received the highest rating—5 stars. The paper also investigates the ways that reviews manifested reviewers' concern for establishing credibility. It examines two reviews' varying degrees of careful editing via organized content, correct punctuation, and correct spelling, and it examines reviewers' assertions and displays of expertise. Both of these two reviews—one of a product with low consumer risk and one of a service with high consumer risk—displayed spelling errors, and both reviews signaled reviewers' expertise. The paper concludes with some ideas for developing a richer system for analyzing computer-mediated communication (CMC) genres like online reviews, particularly in relation to reviewers' signals of expertise.

Introduction

In the emerging computer-mediated communication (CMC) genre of online reviews, lay people, as opposed to professional writers, evaluate products and services, and they receive no pay for their time or effort. Glowing reviews like this review of a recipe for Chicken with Grape Tomatoes and Fried Basil, written by GIGI9801, commonly appear on sites like Allrecipes.com:

> DELICIOUS!! Restaurant quality dish! and I couldn't believe how incredibly easy!! The whole family LOVED this!! I used chicken thighs instead of breasts (we prefer dark meat)and followed the recipe exact. The chicken was super tender and full of flavor! As others said, the only problem is this doesn't call for enough rice. Double the rice to 2 cups, serve with some fresh steamed broccoli and you've got yourself an incredible meal!! This is a must try!! Thanks for the wonderful recipe Amanda!! :)

Clearly, this reviewer was pleased with the recipe she/he found, but are most reviews so complimentary? That is, to what extent are online reviews positive, like GIGI9801's review of the chicken recipe? Prior research on consumer reviews of products and services suggests that

reviewers exhibit a positive bias called the positivity effect. This effect stems from consumers' choosing products and services only when they believe, as Chevalier and Mayzlin (2006) say, "there is a high probability of enjoying the experience" (p. 345). This study examines a sample of 320 online reviews of recipes from the site Allrecipes.com to determine whether the online reviews of this popular cooking site display positive bias. This analysis is interesting in that recipes are in general low-risk products: consumers who use these online reviews do not need to trust in them as much as they do reviews of pricier, and therefore, riskier products. Thus, in addition, this paper examines 320 online reviews of riskier products, such as laptop computers and DVD players, from another popular reviews site: Epinions.com.

This paper also examines the extent to which reviewers use language that appears designed to build and maintain readers' trust. It asks the following question: to what extent do reviewers try to construct credibility, the "characteristics that affect the receiver's acceptance of a message" (Ohanian, 1990, p. 41)? After all, reviewers will have little influence on readers if their evaluations are not perceived to be credible. Indeed, studies of third-party testimonials have shown that source credibility is a clear determining factor in the extent to which readers attend to and follow reviewers' advice (Ohanian, 1990). To investigate this question, this paper analyzes the credibility of two online reviews that theoretically would require quite different levels of consumer trust: one review of a recipe, taken from Allrecipes.com, and a review of a tour guide service, taken from Epinions.com. In investigating credibility, this paper analyzes signals of reviewers' careful editing (see Petelin, 2002) and reviewers' assertions and displays of expertise.

Method of analyzing reviews

To obtain the sample of 320 recipe reviews from Allrecipes.com, I gathered 20–22 recipes appearing as "Recipes of the Day" for 15 randomly selected days. (Each day, 20–22 recipes from different content categories, such as "Appetizers and Snacks," "Main Dish," and "Drinks" appeared as recipes of the day.) The reviews were preceded by star rating on a scale of 5 stars (best) or 1 star (worst). Figure 1 shows a 5-star rating.

★★★★★

Figure 1. A 5-star rating from Allrecipes.com.

Similar rating scales appear in most other sites that invite user reviews, such as Amazon.com. The status of these recipes as recipes of the day was not correlated to the recipes' overall star rating (evident in the fact that recipes with an average 1-star rating were listed as recipes of the day).

I also gathered a sample of 320 reviews from Epinions.com. These reviews were about computers and electronics products that ranged in price from US$500 to US$4000. On

Epinions.com and some other sites devoted to consumers' reviews of products and services, reviewers can rate products on several criteria; for example, in Epinions.com, reviewers can rate an Audiovox portable DVD player for sound, ease of use, picture quality, and durability. These criteria ratings are averaged to create one overall rating. I used the products' overall ratings in this study. I tallied frequencies of star ratings, such as 5-star ratings, to gauge possible positive bias in the 640 Allrecipes.com and Epinions.com reviews.

Results of the review analysis

The results of this study correspond with Chevalier and Mayzlin's (2006) findings from their study of book reviews on Amazon.com and BarnesandNoble.com. Based on reviewers' star ratings from Allrecipes.com and Epinions.com, the reviews in the present study were overwhelmingly positive. In fact, 507 reviews (79.1%) assigned a 4-star or higher rating, and 309 (48.2%) assigned a 5-star rating, as Table 1 shows.

Table 1. Star ratings of recipes in online reviews.

Star rating	Allrecipes.com		Epinions.com		Overall	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
5	174	54.3	135	42.1	309	48.2
4.5 – 4	99	31.0	99	31.0	198	31.0
3.5 – 3	27	08.4	47	14.7	74	11.5
2.5 – 2	11	3.4	19	6.0	30	4.7
1.5 – 1	9	2.9	20	6.2	29	4.6
Total	320	100.0	320	100.0	640	100.0

Table 1 also shows that in terms of 5-star ratings, reviews from Allrecipes.com were more frequently positive than the Epinions.com reviews of riskier products: 174 of the Allrecipes reviews assigned 5-star ratings (54.3%), whereas 135 (42.1%) of the Epinions.com did the same. This result may stem from Epinion.com reviewers' ability to use a more nuanced rating system. As mentioned before, Epinions.com reviewers could rate products on several scales, and those ratings factored into an overall rating, an overall rating that could manifest a half star (e.g., a 4.5-star rating). Even so, both sites contained mainly positive reviews. In Allrecipes.com, 273 recipes (85.3%) received either a 4- or 5-star rating; in Epinions.com, 234 (73.0%) received a 4-star or higher rating.

In addition, in the sites' combined reviews, only 29 contained 1-star or 1.5-star ratings (4.6%). Epinions.com reviewers, reviewing higher risk products, were more likely to give low ratings, a result that may stem from a stronger sense of responsibility for users' decisions about higher-risk

products. Just 9 (2.9%) of the Allrecipe.com reviews contained a 1-star rating, but 20 (6.2%) of the Epinions.com reviews displayed 1- or 1.5-star ratings.

Clearly there was a bias toward positive reviews in both the Allrecipes.com and the Epinions.com samples. One example from Epinions.com is rt66explorer's 5-star, positive review of an eMachines desktop PC:

> Pros: Speed, price, looks, quiet, has 3 fans, has pc3200 memory not pc2700
> Cons: has cheapish speakers and mouse that work fine
> Not bad looking either. Very striking. We got this unit refurb at Tiger direct for $399.00 and it is great so far. It's my second emachine refurb. Other one is still going strong 2 yrs later. I feel even better about this one. This unit has three fans in it and the advertised memory of 512 megs of pc 2700 is actually 512 of pc 3200. Nice touch! It is also very quiet for having all that air moving. Well laid out too. No floppy drive, but I can't remember the last time I needed one anyway, so I guess it's like having an ashtray if you don't smoke. I too think the speakers and mouse are less than great, although I think the keyboard is fine. At least they throw something in there to get you going. That's the main thing.

In contrast, far fewer reviews were negative, like vivaconcordia's 1-star rating and negative review of the recipe for Flounder Zucchini Bundles:

> The fish was so mushy, I couldn't eat more than 3 bites. I found that it had to be baked longer than 25 minutes. The flavor wasn't too bad but I feel that mushy fish coupled with the mushy zucchini left much to be desired.

In general, reviewers of quite different products—products involving very different levels of consumer risk—displayed positive bias in their reviews.

Creating credibility with careful editing

This section of the paper analyzes two online reviews—each requiring a different level of credibility from the reviewer because of the trust required from the consumer.

A review of a recipe

This analysis of credibility in online reviews begins with a positive recipe review. This 5-star review of Slow Cooker Salisbury Steak displays a relative lack of careful editing:

> Soooo good! I was looking for something quick and easy and this was perfect...it cooked all day and when my sons football practice was done we were ready to eat! It is definately more of a fall winter food, but really good anytime! I loved the gravy and may try to double it next time! I did add mushrooms just for something additional. I served it with redskinned garlic mashed potatoes (which I thought I got off this site, but now I can't find the recipe), my whole family loved it! This will definately be a favorite this winter!

This review by momagain3 appears to have been spontaneously created. That is, it is not organized in any particular order or around a central theme. For example, a sentence about the ease of preparing the recipe is followed by a sentence about the time of year that the recipe is most useful. That sentence is in turn followed by a sentence about a change that the reviewer made in preparing the recipe again. It seems that momagain3 made no attempt to subordinate ideas or to group them.

This recipe review displays few of the characteristics of edited language that help generate credibility. For example, a lack of editing is evident in punctuation errors, such as a comma splice (after the parenthetical statement), lack of a possessive apostrophe ("sons football practice"), lack of a forward slash or a hyphen ("fall winter"), as well as in a spelling errors ("definately"). In short, the reviewer has not signaled credibility to establish readers' trust, but this finding is not very surprising given that it is a review of a product that brings little risk to the consumer. In sum, this reviewer's drive to signal credibility via careful editing does not appear to be strong.

A review of a tour guide service

At Epinions.com, reviewers can rate not only electronics and computers, but also nearly any product or service, including service providers like tour guides. The Epinions.com review below, written by the reviewer sewell1591 about a tour service in Beijing, China, is longer than a typical Allrecipes.com review (and hence has been shortened from the original; ellipses mark the places where words have been cut):

> I've made several trips to China, including Beijing, over the past 3 years, including tour (with a different operator) and independent travel. . . . We were met, as promised, just outside immigration at the Beijing airport by a representative of GT, who spoke very good English and had us whisked off to our hotel right away. We stayed at the Prime Hotel, at the northern end of Wangfujing, about a 10 minute walk from the main shopping district and convenient to restaurants My group had scheduled some extra days, and the extensions were at the same rate at the hotel as on the tour and were quite reasonable. Some members took advantage of a $400 add-on and took a day trip to Xian, which included transfers to and from the airports, guide services in Xian, roundtrip airfare from Beijing, and lunch. All reported this to have been a good value and well handled. . . . There was some confusion about coordinating with other members of my party and the exact days we'd be touring. . . . When I reported these problems in a followup after the trip, GT was very apologetic and provided an extra travel credit for another trip. I was satisfied with the professional and speedy way they acknowledged the glitch and feel they have likely taken steps to eliminate the issue in the future. . . . I'm very satisfied with GT and would recommend them to anyone wanting a tour in China! I plan on using them again, myself.

In terms of its editing, this review contains spelling errors ("knowlegable" and "followup"). However, this review displays far more content planning than the Allrecipes.com review analyzed previously; sewell1591 seems to have organized the review by content type. First, the reviewer relates his/her prior experience in China and with other tour services. Then, he/she groups positive evaluations together (e.g., the GT representative "spoke very good English" and

the prices of extensions were "quite reasonable"). Then he/she explains the negative evaluation (i.e., there was "some confusion" about the tour dates). The reviewer concludes with an overall evaluation of the tour: "I'm very satisfied…" In ending with a general, positive evaluation after stating and explaining a negative evaluation, sewell1591 follows a review pattern seen in other discourse genres in which one participant evaluates: face-to-face tutoring interactions (Mackiewicz, 2005) and academic book reviews (Mackiewicz, 2007). The same pattern also seems to appear quite frequently in reviews of electronics and computers, but further research is needed to test this observation. Related to the organization of its content, judging by sheer number of words, this review took more time to write than the shorter recipe review, suggesting the reviewer's thoughtfulness about the topic and concern for the reader and establishing credibility.

To sum up the findings of this section, online reviewers may not be concerned with possible negative effects of spelling errors on their credibility: both reviews examined here display such errors even though they evaluated products/services of quite different consumer risk. However, reviewers of riskier products may be more inclined to organize the content of their reviews to generate credibility.

Creating credibility with signals of expertise

Online reviewers established credibility by asserting their expertise with a product or service. In doing this, they justified their right to make evaluations, whether positive or negative. For example, in a review of Italian Beef Sandwiches, Bonnie R. says that she/he is from a city, Chicago, where such sandwiches are common. Claiming Chicago as a hometown signals Bonnie R.'s familiarity with the subject matter, affording her/him credibility, at least in relation to this recipe. In addition, this reviewer says that she/he has experience in making these sandwiches, which also lends to her/his expertise:

> Thank you! You should rename this one the ONLY REAL Italian Beef Sandwich Recipe! We also moved away from the city of Chicago, and couldn't believe what they were passing off as "Beefs" in our new locale; shredded or loose meat with only a hint of seasoning, and carrots! AAAGH! I came up with this method too after trial and error. I also use flavoring in the water added to the roasting pan, usually a little Worcestshire and a teaspoon of ready-to-use beef bouillon. . . . My recipe calls for an herb/spice rub, with an eight cup each basil, oregano and parsley in dried form. . . . Put on half of this before inserting the garlic and half after. An easy way of doing the garlic is to make shallow scorings lengthwise down the tenderloin and then pat in minced garlic from a jar. . . . Proceed as you do with Dick's recipe, but with a whole tenderloin (sometimes called eye of round) you might need longer for the larger size -- we usually leave it in the oven for 4 hours. The first time we used the recipe, we put it in a roaster for a 70-person grad party, and it was the most requested entree, around 12 pounds completely gone right down to the gravy. Hope the "tweaks" will help..Salute!

In claiming Chicago as a hometown ("We also moved away from the city of Chicago") and describing experience in making these sandwiches ("I came up with this method too after trial

and error" and "we usually leave it in the oven for 4 hours"), Bonnie R. indicated that readers could trust her/his evaluation.

It is interesting that recipe reviewers like Bonnie R. note their experience in the same way that reviewers of more costly—and thus more risky—products do. For example, in his/her review of General Tours, sewell1591 began the review by noting that he/she has traveled to China before ("I've made several trips to China, including Beijing, over the past 3 years") and by noting that he/she used a tour on one of those previous trips ("including tour (with a different operator) and independent travel"). This background knowledge establishes sewell1591's qualifications for reviewing. That recipe reviewers point out their experience attests to reviewers' sense of establishing and maintaining an identity, particularly an expert identity, within an online community, regardless of the cost of the products or services under review.

Clearly, just as reviewers create credibility with edited language, so too do they create credibility by displaying expertise. Of course, the extent to which reviewers see themselves as part of the CMC community for which they have written their review will influence the extent to which they establish and maintain credibility. It is likely the case that long-term users of sites like Allrecipes.com and Epinions.com work to establish and maintain credibility online.

Conclusion

In this paper, I analyzed an emerging variety of CMC—laypeople's online reviews of products and services. It examined of 640 reviews to determine whether reviews of risky and not-so-risky products and services manifested the positive bias associated with self-selection of reviewed products and services (Chevalier & Mayzlin, 2006). The reviews strongly displayed a positive bias: over 79% assigned 4-star or higher ratings and over 48% assigned a 5-star rating. Thus, reviews of both low-risk and high-risk products generated the positivity effect.

I also closely examined two online reviews to determine the extent to which they displayed characteristics of careful editing that help create credibility. Analysis of two suggests the reviewers did not consider a few spelling errors to be a great threat to credibility. However, the two reviews differed in the extent to which they displayed organizational planning of content.

This analysis revealed some other language characteristics that might help determine the extent to which reviewers signal or fail to signal credibility, such as using emoticons, which have been discussed previously in CMC literature (Krohn, 2004; Walther & Kuc, 2001). Reviewers also employed spelling (e.g., Soooo) and punctuation (e.g., !!) in unconventional ways to achieve their aims. That is, reviewers used repetition of orthographic symbols for emphasis. Other reviewers achieved the same effect via repetition of words, such as "Yum, yum, yum" in recipe reviews. Identifying such repetitions could facilitate analysis of careful editing and credibility.

In sum, this study suggests that through their reviews, reviewers may indeed manifest a bias toward positivity. To demonstrate credibility, reviewers rely more on carefully edited language and signal their expertise. Further analysis can investigate the extent of positive bias in other review Web sites as well as signals of reviewer credibility. This emerging variety of CMC is worthy of much further analysis as business communicators can use it to analyze how consumers view their organization's products and services.

References

Chevalier, J. A., & Mayzlin, D. (2006, August). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research, 43,* 345-354.

Krohn, F. (2004). A generational approach to using emoticons as nonverbal communication. *Journal of Technical Writing and Communication, 34*(4), 321-328.

Mackiewicz, J. (2005). Hinting at what they mean: Indirect suggestions in writing tutors' interactions with engineering students. *IEEE Transactions on Professional Communication, 48*(4), 365-376.

Mackiewicz, J. (2007). Compliments and criticisms in book reviews about business communication. *Journal of Business and Technical Communication, 21*(2), 188-215.

Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising, 19*(3), 39-52.

Petelin, R. (2002). Managing organizational writing to enhance corporate credibility. *Journal of Communication Management, 7*(2), 172-180.

Walther, J., & Kuc, M. (2001). The impact of emoticons on message interpretation in computer-mediated communication. *Social Science Computer Review, 19*(3), 324-347.

Biography

JO MACKIEWICZ received her PhD in applied linguistics from Georgetown University. She teaches classes in technical editing, document design, and intercultural communication. Her research interests include clarity and credibility of text and graphic elements. She has published in *Journal of Business and Technical Communication, Journal of Technical Writing and Communication, Technical Communication, IEEE Transactions on Professional Communication,* and *Business Communication Quarterly.* As of 2008, she is the Editor in Chief of *IEEE Transactions on Professional Communication.* Contact: mackiewicz@iit.edu.