Online reviews play an important role in the uncertainty reduction efforts of customers. In this study we investigate which factors influence the customers’ evaluation of online reviews and prove that both product and review characteristics play a role in the perceived helpfulness.

[to cite]:

[url]:
http://www.acrwebsite.org/volumes/16360/volumes/v38/NA-38

[copyright notice]:
This work is copyrighted by The Association for Consumer Research. For permission to copy or use this work in whole or in part, please contact the Copyright Clearance Center at http://www.copyright.com/.
EXTENDED ABSTRACTS

“Effects of the Consumer Review Process on Attitude Formation and Communication”
Stephen He, Georgia Institute of Technology, USA
Samuel Bond, Georgia Institute of Technology, USA

The past decade has observed a dramatic increase in the use of consumer-generated content (CGC). Modern consumers are involved in creating, sharing, and reading content, and these behaviors have received increased attention from marketing practitioners and scholars. In particular, there has been a growing interest in the effects of product reviews on consumers’ purchase behavior (Chevalier and Mayzlin 2006; Weiss, Lurie, and Machinis 2008). However, very little research has focused on the effects of review writing on the writers themselves (c.f. Moore et al 2009). In order to address this need, the current paper considers a number of questions related to review writing and attitudes; in particular, we explore how the format of the review task affects the process of review writing and readers’ comprehension of the review.

We start by assuming that a consumer has recently engaged in a consumption experience and holds some attitude toward the experience, however weakly formed. We consider two different scenarios, in which consumers write a text review either before or after reporting their attitude regarding the consumption experience. Our central argument is that these two scenarios generate fundamentally different motivations, which in turn dictate the reviewing process and downstream consequences. When the review task precedes any rating, reviewers are motivated to address the different views of unknown readers. This motivation leads to a more open-ended, unstructured review task that allows the reviewer to elaborate extensively on different aspects of the experience and his or her reactions towards them (Tetlock, 1989; Schlosser, 2005). Such an ‘explorative unfolding’ of information allows the consumer to bring to mind a wide variety of attitude-relevant information, resulting in attitude change.

A stark contrast exists when the reviewing task follows evaluation. In this case, reviewers who have already reported their attitude will feel committed to their rating and utilize their review to engage in “defensive bolstering” (Tetlock, 1989). However, individuals are generally poor at recognizing the reasons underlying their preferences (Nisbett and Wilson 1977), and this inability should lead reviewers to encounter difficulty generating sufficient support for their stated evaluations. In keeping with the principle of metacognitive inference (Schwarz, 2004), we suggest that this perceived difficulty may result in attitude moderation. Moreover, due to their one-sided, purposeful nature, text reviews written after a rating may not be very effective at conveying the authors’ attitude to readers.

In a series of lab experiments, we examined hypotheses related to the two different cases above. A pretest utilizing various short animated movie clips revealed consistent attitude differences between participants who wrote a text review of a clip and participants who simply evaluated the movie. These pretest results provided initial evidence that the mere act of writing a text review changes reviewers’ attitudes.

Study 1 explored the effect of task order on attitude change and actual review content. In the experiment, undergraduate participants (N=67) watched the target stimulus, a short animated movie clip, and were then given different instructions according to condition. Two factors were manipulated: task order and type of writing. Regarding the first factor, write-then-rate conditions completed the writing task and then rated the movie, while rate-then-write conditions completed the steps in the opposite order. Regarding the second factor, all participants were asked to generate either a test review of the movie or a filler (summarizing events of the preceding day). In a follow-up session occurring three weeks later, participants were reminded of the movie clip and asked to recall their evaluation.

Replicating our findings, analyses revealed that at t1, the mean attitude towards the movie was considerably lower for the review-then-rate condition than the other three conditions (which did not differ). At t2, this attitude remained basically unchanged. For the rate-then-review condition, not only was t2 attitude significantly lower than that of the control groups, but the absolute difference in attitudes between t1 and t2 was significant as well. These findings support our argument that even after initial assessment, attitude was moderated by the process of writing a text review.

In order to examine review content itself, we applied the Linguistic Inquiry and Word Count tool (Pennebaker and Francis ME 2001) to analyze reviews written by the two test review conditions. Among other findings, analyses revealed that the review-first group wrote significantly more words than the rate-first group, indicating more elaboration of the movie experience. In addition, the ratio of positive-to-negative word use was closer to one for the review-first group than the rate-first group, indicating that the writing of the former was more balanced. Finally, the review-first group made greater use of articles, suggesting a more objective writing style.

Study 2 examined how differences in review content influence readers’ ability to infer the attitude of the reviewer. Among the reviews written by participants in Study 1, six each were selected from the review-first group and the rate-first group. Participants (N=68) were asked to read the reviews, estimate what ratings the authors assigned the movie clip, and report their confidence in these estimates.

Rating discrepancies were calculated by taking the absolute difference between a reader’s estimated rating and the author’s actual rating from Study 1. In support of our prediction, analyses indicated that rating discrepancies were lower for reviews written by the review-first group that for the rate-first group. Notably, participants were actually more confident in their estimates after reading reviews from the rate-first group, despite also being more inaccurate. Given the text analysis results reported above, it is likely that there balanced reviews of the review-first group seemed more ambiguous to the readers but actually conveyed the writers’ opinion more effectively.

Overall, our results provide initial evidence that reviews written before vs. after global evaluation affect reviewers’ attitudes in systematic ways, differ in terms of structure and content, and create different interpretations among downstream readers. These findings bear important implications for those researching the review-writing process, and also for marketing practitioners interested in utilizing consumer-created content.
“The Ratings Paradox: Why We Prefer Reading Negative Reviews, But Then Subsequently Rate These Reviews as Less Useful”

Kurt Carlson, Georgetown University, USA
Abhijit Guha, Wayne State University, USA

User reviews, on sites like REI and Amazon, have emerged as an important part of the consumer buying process. In fact, in a recent survey of undergraduate students, 78% of respondents reported that they read one or more user reviews before making an online product purchase. This percentage exceeded all other behaviors, which included talking to friends to get advice (58.5%), reading one or more expert reviews (35.4%), and visiting the store to see the product (20.7%).

User reviews typically include a rating (on a 1-5 scale), coupled with an explanation of the rating (the review) and a report on how useful the review was for previous readers of the review. The usefulness data is generated from those who answer yes or no to the question, “Was this review useful?” Prospective consumers can sort the reviews by valence (e.g., one-star vs. five-star) and by usefulness score (the average usefulness of the review, as reported by those who have previously read the review and answered the usefulness question).

If there is a relatively stable relationship between the number of people who read each review and the number who answer the usefulness question, then the total number of responses to the usefulness question can be used to assess which reviews are read more often. Those reviews with the greatest number of people who answered the usefulness question are those that were most read. When examining the useful ratings, we find the following pattern. Negative (one-star) user reviews are read more often than positive (five-star) reviews, but positive reviews are rated as more useful than negative reviews. We dub this the ratings paradox.

We propose that the ratings paradox can be explained as follows. Consumers who read online product reviews generally have a weak preference for one of the products. Empirical evidence supports this claim. In particular, respondents in a recent survey we conducted reported that (on average) 69.5% of the time that they read user reviews they had a preferred product in mind while reading the reviews. We propose that this preference drives consumers to seek negative reviews about the currently preferred product. It also causes consumers to bias their evaluations of negative reviews, making them subjectively neutral. Having distorted away their diagnosticity, negative reviews are rated as less useful than positive reviews.

We rely on a series of lab studies to test our theory. In the first study, participants (N=53) engage in a choice process involving two hotels. After reading some information about the hotels, participants are asked to report which hotel is leading the choice (i.e., which option they currently prefer). After reporting their leader, participants selected one user review they would like to read about the hotels. The reviews consisted of positive and negative user reviews for each hotel. Results reveal that participants exhibited a strong tendency to seek out the negative review about their weakly preferred hotel. Specifically, 70% participants wanted to read the negative review about their preferred hotel (significantly greater than chance, binomial z=7.30, p<0.001). Thus, consumers who have a weak preference sought negative reviews about their preferred product, a result that fits with our observation that negative user reviews are read more often than positive user reviews.

We next investigated why consumers rate negative reviews as less useful than positive reviews. We know that consumers who have a weak preference tend to bias the information they encounter to support this preferred option (Carlson 2006; Russo et al 1998). If consumers bias user reviews in this manner, then negative reviews would be converted into less negative or possibly even neutral reviews. If subjectively seen as “near-neutral” by consumers, negative reviews would rightly be rated as less useful than positive reviews simply because they would be less subjectively diagnostic.

To examine this idea, we conducted two laboratory experiments using undergraduate participants from a US university. In the first experiment (N=60), participants examined one-star reviews and five-star reviews about some products (cameras, TVs, tents etc.) drawn from Amazon and REI. Online, only 18.9% rated the one-star reviews as useful, whereas 87.5% rated the five-star ratings as useful. Laboratory participants, who were not expected to have a preference for either product because they were not shown any information for the products rated the usefulness of the reviews exactly opposite. That is, 86.7% rated the one-star reviews as useful, whereas only 35.0% rated the five-star reviews as useful. These proportions were significantly different from each other (z>2.5, p<0.05).

We next conducted a laboratory experiment (N=44) to examine whether negative user reviews would be distorted to be subjectively perceived as near neutral. Half of the participants (N=21) were shown the hotel stimuli (described earlier) and were allowed to form a preference for one of the two hotels. Participants were then asked to choose between reading negative information about the preferred hotel or negative information about the other hotel. Consistent with Carlson and Guha (2009), 18 out of 21 chose to read negative information about the preferred hotel (significantly greater than chance, binomial z=2.17, p<0.05). Participants were then given the negative information they requested. The 18 participants who read the negative information about their preferred hotel rated this information as 4.55 (i.e., near-neutral) on a 1-9 scale, (where 1 indicated negative information and 9 indicated positive information). In contrast, a separate sample of 23 participants, who examined the negative information without first forming a hotel preference rated the negative information as 2.52 (i.e. closer to the negative end) on the same 1-9 scale (2.52 is less than 4.55; (F(1, 35.0)=9.2, p<0.05), evidencing distortion of negative information. In comparison, the two mean evaluations reveal a strong tendency for those with a preference to convert negative information into near neutral information.

In sum, we find consumers seek negative information about a preferred product and then bias that information making it seem less useful. This pattern of lab results is consistent with our finding that negative user reviews are read more often but rated as less useful, a finding we have dubbed the “ratings paradox.”

“What Determines Customers’ Evaluation of Online Reviews? The Role of Review and Product Characteristics”

Simon Quaschning, Ghent University, Belgium
Mario Pandelaere, Ghent University, Belgium
Iris Vermeir, Ghent University, Belgium

During the last years there has been a huge increase in available choice options for customers (Schwartz et al. 2002), a phenomenon called hyperchoice (Mick, Broniarczyk, and Haidt 2004). But confrontation with extensive options makes decision making more difficult (Schwartz et al. 2002), partly because customers feel more responsible for their choices (Mick et al. 2004). As a result consumers will experience more uncertainty (Anderson 2003). Being confronted with an extensive assortment size may be especially likely in an online context. In addition, products online are intangible which is an additional source of uncertainty (Eggett 2006). According to the uncertainty reduction theory (Berger, and Calabrese 1975), customers will engage in uncertainty reduction efforts to alleviate and eliminate risk caused by uncertainty and to maximize outcome value. Online customer reviews have become
an important source of information and can play a prominent role in this process (Chevalier, and Mayzlin 2006). While previous research has mainly focused on the link between customer reviews and product sales (Chevalier, and Mayzlin 2006; Hu, Liu, and Zhang 2008), this study explores the use of customer reviews in a pre-decision context.

Confronted with an extensive number of alternatives, customers have to make a selection of considered products beforehand for which they subsequently engage in an online search. In an attempt to make an adequate decision, they may consult customer reviews. In the present study we investigate factors that influence customers' evaluation of the helpfulness of reviews. We focus on Amazon.com ratings of helpfulness of reviews. These ratings result from a two-step process. First, customers have to decide whether or not they will rate the helpfulness of a particular review. Second, if they decide to provide a rating, they will have to decide on a positive or negative rating. The current paper investigates whether and how these two decisions are affected by product and review characteristics, like whether the review was positive or negative, the number of words, the position of the review, whether the review is written by credible reviewer, the sales rank of a product, and the type of product.

Our data consists of both product features and customer review characteristics for 260 different products, of which 116 were books and 144 CD’s, taken from the public website of Amazon.com. We collected information on the price, the position on the Amazon bestseller list (sales rank) and the number of reviews written for the product. We included products with different sales ranks, with groups ranked from 1 to 25, 101 to 125, 501 to 525, 1001 to 1025, 5001 to 5025 and 10001 to 10025. For each product we took a maximum of 20 reviews into our sample, which resulted in a total of 4229 reviews, of which 1664 (39.3%) were book reviews and 2565 were CD reviews (60.7%). For 40 products there were no reviews, so we did not include them in our analysis. For each review, we counted the number of words in the review, the position the review had on the website, whether it was a top 100–reviewer (i.e. credible reviewer), the score (ranging from 1 to 5) given by the reviewer to the product and also the number of customers that voted on the helpfulness and the percentage of voting customers that found it helpful. To explore which characteristics influence revealed customer appreciation we used multilevel regression analysis.

The present findings suggest that whether the review is positive or negative, its length and position, the reviewer credibility, the sales rank and the type of product are important predictors for the review to get a helpfulness response by the reader. Whether this response is positive and thus indicating that the reader values this review or not depends on the sales rank (depending on the type of product), number of words and score given by the reviewer.

Customers are responding more to negative reviews, which presumably render them more uncertain about their inclinations to purchase, than to (moderately) positive reviews. At the same time, however, positive reviews are rated as more helpful than negative ones. While the former can be due to the fact that customers react more extensively to inconsistency, the latter can be explained by peoples’ tendency to systematically prefer information that is consistent rather than inconsistent with their own standpoint (Festinger 1957). A quadratic relationship with review length shows that moderately long reviews are considered more helpful than short or long reviews. This represents a trade-off between a balanced message with different arguments and too much information. Regarding the position of a review, book reviews further down the website, which happen to be the older reviews, attract more responses. Reviews posted long after the product release will get less attention. For CD reviews, this position effect was not obtained. Consumers tend to selectively pay more attention to reviews from top 100-reviewers, but no effect was found on the valence of the vote (helpful or not). Obviously, reviews for bestselling products attract more votes than reviews for products with low sales. However, reviews for products further down the sales ranking are perceived as more helpful. Indeed, they may be more unknown products, for which extra information can be more helpful. This relationship cannot be found for CD’s. Finally, book reviews attract more votes regarding their helpfulness than CD reviews. Possibly, customers rely more on the opinions of others for books than for CDs. The latter are more subject to personal taste.

To conclude, our results show a significant effect of several product and review characteristics on the evaluation of online reviews. For future research it might be interesting to examine the influence of experts on the evaluation of negative reviews. Also an evaluation of the review content, such as objectivity, and its connection to the valuation of the review by the customer would be interesting. Finally, since we believe that the degree of uncertainty is essential to explain potential buyers’ assessment of reviews, a further step would be to test this by manipulating the degree of uncertainty.

“Anonymous and Unanimous: The Impact of Anonymity on Opinion Generalization”

Dilney Goncalves, INSEAD, France
Amitava Chattopadhyay, INSEAD, Singapore

People care about the opinion of others. In particular, the opinion of the majority is often regarded as representing the truth (e.g., Asch, 1951; Festinger, 1954). When consumers search information online, consumers may think that a given review presents a very general opinion about a product or she may think that the reviewer is an outlier and the opinion is very idiosyncratic. Since reviewers are self-selected, the distribution of reviews does not necessarily represent the distribution of opinions in the larger population (XinXin & Hitt 2008). Thus, before making inferences about product quality, consumers are likely to make inferences about how general the opinion expressed in the review is. Weaver et al. (2007) have shown that people might use metacognitive experiences as information when judging the popularity of an opinion. They showed that more familiar opinions were considered more widespread than less-familiar opinions, even when familiarity was increased by incidental factors—like the number of times the same person expressed an opinion.

A fundamental difference between online word-of-mouth (e.g.: reviews) and traditional word-of-mouth lies in the fact that consumers receive opinions of strangers. More than unknown to the consumer, reviewers can be completely anonymous. We propose that anonymity can influence how general an opinion feels to the reader. In particular, we hypothesize that anonymous reviews will be judged more general than signed reviews.

We assume that people have experience with a variety of opinions, ranging from very idiosyncratic to very general ones. We further assume that opinions and people are linked in memory in a spreading activation network (Anderson, 1983). Thus, people have learned from experience that when an idiosyncratic opinion is activated in memory, activation spreads to the person linked to the opinion and it feels easy to associate person and opinion. Conversely, when a general opinion is activated, because it is connected to many more people, no particular person is easily associated with the opinion. As a result, it feels difficult to associate an opinion to any particular person when the opinion is widespread. We propose that people make the inverse inference: if it feels easy to associate a person to an opinion, then the opinion is probably idiosyncratic. Conversely, if it is difficult to associate a person to an opinion,
then people judge the opinion to be more general. Although this is usually true, factors unrelated to the generality of an opinion can make it easier or more difficult to think of a person. One such factor is the presence of a name accompanying a review. Next, we describe three studies conducted in this research.

In study 1, we asked participants recruited online (N=251) to read and evaluate a consumer’s opinion about a restaurant. After reading the opinion—which was negative—participants were asked to estimate the percentage of the restaurant’s customers who were dissatisfied. About half the participants saw an anonymous opinion. The other participants saw a signed opinion, which was exactly the same as the anonymous with the addition of a customer’s name in the last line of text. In all studies, two names were used. Since there were no differences between the two name conditions and they were aggregated in all studies. The estimates of opinion generality supported our hypothesis: participants reading an anonymous complaint estimated that 53% of the restaurant’s customers were dissatisfied while those reading a signed complaint estimated that 43% of customers were dissatisfied (p<.05).

To further test our process explanation and to examine whether knowledge about the population would eliminate it, we ran a second study. In this study, participants (N=338) read a restaurant review. Prior to reading it, though, participants were informed of the restaurant’s average rating (high or low). The reviews were either positive or negative. Finally, reviews were either anonymous or signed—all factors manipulated between-subjects. After reading the review, participants were asked to estimate what percentage of consumers were as satisfied (dissatisfied) as the reviewer and to explain how they reached that estimate. A significant average rating x review valence showed that participants used the mean information and estimated greater agreement when the review was in line with the average (p<.001). More importantly to our hypothesis, a significant main effect of anonymity showed that its effect still persisted even when people had knowledge of the restaurant’s average evaluation. Participants were more likely to make attributions to the reviewer when the review was signed (23.5%) than when the review was anonymous (11.5%; χ²=8.26; p<.05). Finally, a mediation test showed that attributions to the reviewer mediated the effect of anonymity on estimates of opinion generality.

In the third study, we examine the effect of reviewer anonymity on choice. Participants (N=183) were recruited online to participate in the study. They were randomly assigned to an anonymous- or signed-review condition. Participants read a scenario in which they had to buy a birthday present in a hurry. They were shown a gift-basket and a positive description of it. This was intended to generate a positive first impression of the gift-basket. Then, participants read a negative review which was either signed or anonymous. Participants decided whether or not they would buy the gift-basket and estimated the percentage of people who were dissatisfied with the gift-basket. After reading the negative review, participants were more likely to buy the gift-basket in the signed condition (18.2%) than in the anonymous condition (4.7%). This effect was mediated by estimates of opinion generality.

Our results provide evidence that anonymous opinions are considered more general than signed opinions. This is because it feels easier to associate an opinion to a particular person when the opinion is signed than when it is anonymous. As a result, people are also more likely to comply with anonymous reviews. Besides demonstrating a novel effect of anonymity in online reviews, this research contributes to and extends the literature on social influence and metacognitive experiences (e.g.: Schwarz, 2004; Weaver et al., 2007) and has direct implications for the design of websites.