The Impact Of The Collective Rating Presence On Consumers’ Perception

Zhuolan Bao
Michael Chau

Follow this and additional works at: https://aisel.aisnet.org/iceb2015

This material is brought to you by the International Conference on Electronic Business (ICEB) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICEB 2015 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
THE IMPACT OF THE COLLECTIVE RATING PRESENCE ON CONSUMERS’ PERCEPTION

Zhuolan Bao, The University of Hong Kong, Hong Kong, mbao@connect.hku.hk
Michael Chau, The University of Hong Kong, Hong Kong, mchau@business.hku.hk

ABSTRACT

In online markets, collective ratings by prior buyers are often displayed in a marked place and influential for later consumers. While the aggregated ratings transfer overall evaluation towards products, they might also bring biases to potential consumers. In this study, we hypothesize that collective rating, as a piece of information, acts as 1) a predisposition which affects people’s perception towards other information; and 2) a risk level of product’s performance which changes the way people perceive consensus or deviant word-of-mouth information from online reviews. Using online reviews of multiple product categories from Amazon.com, our study reveals the impact of collective ratings on consumers’ perception of WOM information and sheds light upon the conflictive results on perception biases of product reviews. Implication for understanding and facilitating consumer perception of online reviews are discussed.

Keywords: Collective rating, online product reviews, predisposition, perceived risk, helpfulness.

INTRODUCTION

Online markets facilitate our life by providing convenient and fast shopping experiences. In the past two decades, many information systems and marketing researchers have been exploring the factors of market success in online context.

To mitigate the information asymmetry due to the Internet nature, many platforms began to use review systems, which encourage buyers to post their experiences and evaluation of products. By reading prior buyers’ rating and text content, potential consumers can optimize their purchase decisions.

However, as the amount of information increases, potential consumers are overwhelmed by a large amount of information with uncertain quality and credibility. In addition, there are also marketers and reviewers who post fake reviews for certain benefits. In order to help consumers identify the product quality with less time and efforts, two approaches are commonly adopted by shopping websites on their review systems. The first approach is the collective rating information. Figure 1 is an example on the shopping website Amazon.com. The collective evaluation approach displays an aggregate evaluation from buyers who post reviews of the product. The evaluation is shown by either a number or a distribution diagram. The aggregated rating information helps identify good products. Potential buyers can quickly obtain knowledge of product evaluation or use it to search for products within their requirement of product quality. Past research has shown that sellers and marketers benefit from the aggregated average rating, since the rating is positively associated with the product price, sales and the trustworthiness of sellers.

The second approach is a voting mechanism which allows customers to vote for reviews that they feel helpful or unhelpful regarding their purchase decision, as shown in Figure 2. The voting information is displayed as, for example, “32 of 40 people found the following review helpful”. With higher helpfulness, reviews are more likely to be read and considered. Therefore, the helpfulness votes are important in identifying good reviews and hence good products.

Customer Reviews

4.5 out of 5 stars

Share your thoughts with other customers

Write a customer review

See all 2,861 customer reviews

Figure 1. Collective rating information of customer reviews on Amazon

The second approach is a voting mechanism which allows customers to vote for reviews that they feel helpful or unhelpful regarding their purchase decision, as shown in Figure 2. The voting information is displayed as, for example, “32 of 40 people found the following review helpful”. With higher helpfulness, reviews are more likely to be read and considered. Therefore, the helpfulness votes are important in identifying good reviews and hence good products.
A number of studies focus on how people perceive review information. Early research explored that review helpfulness could be determined by reviews’ observable features and reviewers’ characteristics, such as review age, review length, and reviewer’s expertise [11][31][34][41]. Also, content analyses have been applied by scholars to investigate review helpfulness, suggesting that more readable, subjective, and emotional reviews tend to be perceived more helpful [7][29][45][46].

Although past work has studied collective ratings functions and helpfulness voting behaviors intensively, there is limited study of their connections when the two approaches both facilitate consumers in identifying good products. A motivating scenario is: when a customer considers buying a particular product, she might first refer to the product average rating for an impression of its quality, and then she would read and vote for reviews posted by prior buyers. An intuitive following question would be: will the impression of collective rating influence the perception of review content?

Our study focuses on the interaction effect of collective ratings and helpfulness voting mechanism. We ask the following research questions: Will the presence of the product collective rating introduce biases or change the way people perceive word-of-mouth information? We address this question by presenting a framework to examine the multiple roles of collective ratings, and then applying this framework to the perception of review helpfulness.

Our work adds to an increasing number of studies exploring the helpfulness perceptions of online reviews. Instead of studying how people vote for reviews with different review features, we expand our focus to a broader perspective of information flow in the online review settings. We argue that the collective ratings, directly and indirectly, affect the way people perceive review helpfulness. Our findings also give explanations to the inconsistent findings of helpfulness perception biases.

**THEORY BACKGROUND AND HYPOTHESES**

According to Mudambi et al. [34], online review is defined as peer-generated product evaluation posted on company or third party websites and helpful online reviews are those facilitate consumers purchase decisions. Buyers and sellers can both benefit from helpful online reviews. For buyers, obviously, review helpfulness helps reduce their search costs. For marketers, they could obtain a strategic advantage in consumer attention by identifying and displaying helpful reviews [11][46].

In studying the consumers’ perception towards word-of-mouth information, we make use of review helpfulness for the following reasons. First, in online review context, review helpfulness is consistently defined as a reflection of the diagnosticity value of a review [25][34]. Diagnosticity perception displays a diverse and integrated consumer perception. Hoch et al. [23] define perceived diagnosticity as the extent to which it helps the consumer assign a product to one (and only one) cognitive category. Jiang et al. [26] bring the definition into online context and used it to represent consumers’ cognitive belief that a website facilitates their product understanding. In this study, we are interested in the overall perception of word-of-mouth information, which might include various dimensions, such as credibility, importance, relevance, informativeness, accuracy and so on. Review helpfulness, as an integrated reflection, fits our purpose of the study.

Second, in the domain of information processing, review information is processed in multiple stages. Kuan et al. [29] posit that message evaluation is based on not only the effect of comprehension, but also the impact of attention. While vivid reviews tend to attract readers’ attention and are more likely to be voted [35], they are not necessarily more diagnostic than pallid information [22]. To avoid misinterpreting readers’ evaluation of review information, they examine the effects of review votes as vividness perception, and the effects of review helpfulness as diagnostic perception. In our study, we share the same emphasis on the comprehension outcome of the review information. Therefore, we follow the extant studies and adopt the overall helpfulness perception as a proxy of review diagnosticity.

Figure 2. Voting mechanism of online reviews
One common finding of helpfulness perception is negativity bias, that reviews with negative ratings are more likely to be helpful [7][29][46]. The arguments are based on the consistent evidence of generalized negativity bias in multiple disciplines, for that bad things are rare and revealing [20], and they receive more attention and more thorough processing than good things [5]. Meanwhile, Pan et al. [36] propose an opposite view. They conclude that positive reviews are more helpful than negative ones, because positively-rated reviews are more congruent with consumers’ predispositions and more likely to be perceived helpful. However, both of the two biases neglect the role of actual collective product ratings that the readers are disposed, as the collective information can introduce prior beliefs to readers about the products.

**Prior Belief**

Prior belief, or predisposition towards a product, has the potential to affect a person’s judgment of WOM information in assessment process [1][12][44]. From the perspective of covariation assessment, Alloy et al. [1] proposed that when doing evaluation, prior expectation and currently available information contribute and interact in the assessment process. They suggest that the stronger the individuals’ prior beliefs, the more the feelings will dominate the interpretation and use of information. Wilson et al. [44] found that no matter the predisposition was newly established or well-founded, the results are the same.

We define consensus information as information consistent to the prior beliefs and deviant information as one that contradicts the prior beliefs. Levin et al. [30] applied intuitive statistics paradigm and concluded that subjects will discount deviant information in making an inference from a sample to a population when they recognize such information to be unrepresentative of the population. Crocker [12] also provided ideas from the covariation processing perspective. Although deviant information can be processed at a deeper level and easier to recall, but if the incongruence can be explained so that it makes sense in the context of the other information, then it is no longer incongruent or the incongruence is qualified and limited [12]. In this way, the deviant information is likely to be recalled but with little influence on assessment process. Hoch et al. [23] explained the impact of predisposition from another angle. They held the opinion that prior impressions are persistent and hard to be changed by other information, even by a contradicted information, because 1) any ambiguous information is interpreted as consistent to expectancies, 2) any consistent information to expectancies increases confidence to expectancies, and 3) any inconsistent information is discounted or ignored [22][23]. In online shopping websites, consumers normally confront various products they may or may not have heard of. Since collective rating information can always provide prior beliefs about the products, consumers’ receptivity to WOM information can be determined by the “fit” with the predispositions. Hence, collective rating will directly influence the helpfulness perception so that consensus information will be more favored than deviant information. Therefore, we hypothesize that,

Hypothesis 1. A review whose rating is closer to the prior collective rating is more likely to be perceived diagnostic.

**Perceived Performance Risk**

Besides a predisposition, the collective rating in online markets provides the satisfaction perception of the product/service item. Since consumer behavior can be viewed as risk taking [4][27], it is essential for online consumers to reduce the risk level by pre-purchase information acquisition [19].

Extant research has defined six components of perceived risk, namely financial, physical, psychological, performance, social, and time-related risk [43]. The collective evaluation of product given by prior product reviews provides a relatively objective evaluation of the product performance. The higher the evaluation is, the more certainty consumers will perceive upon the item and the less the performance risk will be. Since performance risk occurs when the product chosen might not perform as desired and thus not deliver the benefits promised [24], interpreting the collective rating as a measure of performance value is consistent with the notion of the perceived risk in business context.

One might wonder the relationship between performance risk and the product uncertainty concept in Dimoka et al. [14]. Product uncertainty is defined as the buyers’ difficulty in evaluating the product and predicting how it will perform in the future [14]. In our research, performance risk is different from product uncertainty. A high level of product uncertainty indicates a situation where buyers are more difficult to evaluate the product, while a high performance risk suggests that the product is more likely to have a low quality.

WOM is an important risk reliever for consumers at pre-purchase phase [19][40], but the impact of WOM is different as a function of perceived risk. Arndt [2] showed that comparing to low-risk perceivers, the high-risk perceivers tended to make more efforts to seek word-of-mouth information. The high riskers are more active in various WOM sources, such as starting pre-purchase conversation, listening to comments, requesting more information and so on. Online markets have made the approaches of obtaining WOM information easier, so online consumers are more likely to initiating searching behaviors.

Since product rating implies the risk of the purchase, it is inferred that high product rating presents a low-risk purchase environment, and low product rating invokes high-risk perception. Therefore in our context, we posit that consumers are less open, and less willing to accept various information when evaluating products with low risks, than they are when evaluating products with high risks. To summarize, we hypothesize that,
Hypothesis 2: The deviant information is perceived more diagnostic for products with high risk, and consensus information is perceived more diagnostic for products with low risk.

METHODOLOGY

To test these hypotheses, we conduct an empirical study on a real-world setting of online shopping platform.

There are several reasons to choose the Amazon website to test our hypotheses. First, Amazon is one of the biggest online markets all over the world and consistently has the largest number of posted reviews [36]. Many prior studies of online reviews have been conducted on Amazon. Our findings could potentially possess more generalizability as they are produced on the typical and influential online market. Second, previous studies delivered inconsistent results of rating biases by Amazon data. As our research provides alternative views of consumers’ shopping behavior, it is better to test our hypotheses by data from the same source.

The data we use were collected by the Stanford Network Analysis Project (http://snap.stanford.edu/index.html) [32]. Seven categories were chosen in our pilot test, including Electronics, Gourmet & Food, Health, Home & Kitchen, Musical Instrument, Sports & Outdoors, and Tools & Home Improvement. We discarded products that were launched before the helpfulness voting mechanism was added, resulting products whose launch time are more than 2,500 days from now to be deleted. Therefore, our pilot dataset contains a sample of 213,934 reviews on 52,022 products. Following is a description table for the data we collected.

<table>
<thead>
<tr>
<th>Category</th>
<th># Products</th>
<th># Reviews</th>
<th>Avg. #reviews/product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electronics</td>
<td>7,493</td>
<td>33,668</td>
<td>4.49</td>
</tr>
<tr>
<td>Gourmet &amp; Food</td>
<td>3,251</td>
<td>11,294</td>
<td>3.47</td>
</tr>
<tr>
<td>Health</td>
<td>7,930</td>
<td>33,563</td>
<td>4.23</td>
</tr>
<tr>
<td>Home &amp; Kitchen</td>
<td>9,421</td>
<td>39,188</td>
<td>4.16</td>
</tr>
<tr>
<td>Musical Instrument</td>
<td>2,986</td>
<td>11,218</td>
<td>3.76</td>
</tr>
<tr>
<td>Sports &amp; Outdoor</td>
<td>8,744</td>
<td>36,264</td>
<td>4.15</td>
</tr>
<tr>
<td>Tools &amp; Home Improvement</td>
<td>12,724</td>
<td>48,739</td>
<td>3.83</td>
</tr>
<tr>
<td>IN TOTAL</td>
<td>52,549</td>
<td>213,934</td>
<td>4.07</td>
</tr>
</tbody>
</table>

Measures

We use review helpfulness as our dependent variable (Helpfulness). We measure review helpfulness by the ratio of the helpful votes to the total votes received by a review.

To measure how close the review rating is to the average product rating, we introduce information disparity (InfoDisparity), which is the absolute difference from a review’s rating to the average product rating at that time. To measure it, first, we sort the reviews under each product according to their posting time. Second, we calculate the moving average score of the product when each review was posted. Third, the InfoDisparity for each review is calculated. As we explained above, we measure the perceived shopping risk for each product by the overall average rating score of the product (Avg.ProductScore) that the consumers are reviewing.

At the same time, following prior research, we controlled a series of relevant variables on product level and review level. On product level, we use the launched time of product (LaunchTime), price (Price) and the number of reviews under the product (ReviewNum) as control variables. On review level, we use control the elapsed time of review (ElapsedTime) as a proxy of review age, review’s word count (WordCount), reviewer’s expertise (UserExp), and also some review’s textual features.

Past research has found that many textual features of online review could influence the diagnosticity perception, such as readability, subjectivity, certainty and sentiment. We therefore control them in our research by using various content analysis techniques. First, to control for the reviews’ readability level (Readability), we calculated the Gunning Fog Index. It estimates the years of formal education needed to understand the text on a first reading [18], and had been used in many online review studies of IS discipline [17][28]. Second, to measure the texts’ subjectivity level (Subjectivity), we prepared the subjectivity and objectivity classifiers and calculate the percentage of subjectivity in review content, following the approach of Ghose et al. [16]. Third, we used a dictionary provided by the Linguistic Inquiry and Word Count (LIWC), which was developed by Pennebaker et al. [39] and designed to calculate the degree to which people use different categories of words across a wide array of words. We applied LIWC to calculate the words that appear in categories of certainty (Certainty), positive sentiment (Positive) and negative sentiment (Negative). At last, we used Uniqueness to measure the uniqueness words in each review under a particular product item. It was calculated by the percentage of new words that appear in a review and have not been found in the previous reviews for the certain product.

The descriptive statistics of the variables are listed in Tables 2.
Because there are no observations on the mean and standard deviations of helpfulness unless there is at least one vote, a potential selection bias might exist in our sample [34]. We therefore follow the approach of Kuan et al. [29], using a two-step procedure with a Heckman selection model [21]. Also, it might not be meaningful to calculate the mean and standard deviation of helpfulness percentage when there is only one vote for the review. So we also examine the robustness of results using different minimum numbers of votes to estimate review helpfulness.

The models that we estimate are as follows.

\[
\text{Voting}_k = \alpha_1 \cdot \text{Avg.ProductScore} + \alpha_2 \cdot \text{Avg.ProductScore} \cdot \text{InfoDisparity} + \alpha_3 \cdot \text{InfoDisparity} + \alpha_4 \cdot \text{Log(UserExp)} + \alpha_5 \cdot \text{Readability} + \alpha_6 \cdot \text{Subjectivity} + \alpha_7 \cdot \text{Certainty} + \alpha_8 \cdot \text{Positive} + \alpha_9 \cdot \text{Negative} + \alpha_{10} \cdot \text{Uniqueness} + \alpha_{11} \cdot \text{Log(ElapsedTime)} + \alpha_{12} \cdot \text{Log(WordCount)} + \alpha_{13} \cdot \text{LaunchTime} + \alpha_{14} \cdot \text{Price} + \alpha_{15} \cdot \text{ReviewNum} + \mu \quad (1)
\]

\[
\text{Helpfulness} | (\text{Voting} \geq k) = \beta_1 \cdot \text{Avg.ProductScore} + \beta_2 \cdot \text{Avg.ProductScore} \cdot \text{InfoDisparity} + \beta_3 \cdot \text{InfoDisparity} + \beta_4 \cdot \text{Log(UserExp)} + \beta_5 \cdot \text{Readability} + \beta_6 \cdot \text{Subjectivity} + \beta_7 \cdot \text{Certainty} + \beta_8 \cdot \text{Positive} + \beta_9 \cdot \text{Negative} + \beta_{10} \cdot \text{Uniqueness} + \beta_{11} \cdot \text{Log(ElapsedTime)} + \beta_{12} \cdot \text{Log(WordCount)} + \beta_{13} \cdot \text{LaunchTime} + \beta_{14} \cdot \text{Price} + \beta_{15} \cdot \text{ReviewNum} + \xi + \lambda(\cdot) \quad (2)
\]

**POTENTIAL CONTRIBUTION & CONCLUSION**

The purpose of this research is to discover whether the presence of the product collective rating introduces biases or change the way people perceive information. We extend our knowledge of collective ratings from new perspectives – forming predisposition and risk perception of each product. We suggest that 1) review’s collective rating has a direct predisposition effect on consumers’ perception towards detailed review information, and 2) the collective information indirectly influences the helpfulness perception behavior of online consumers as its appearance portrays the risk level of product performance.

**Theoretical & Practical Implications**

A main contribution of our study is to extend the current research of reviews helpfulness perception. We start from the perspective of collective rating information. Instead of proving positivity bias [36] and negativity bias [29][42], our results will provide evidence that predisposition influences review feedback perception, resulting that the consensus information is more likely to be favored. Since the two biases have received substantial discussion over the past decade, our work on the collective rating intends to supplement their research findings and help reconcile and explain the inconsistency.

The present research will also contribute to the knowledge of consumer perception towards word-of-mouth information. We suggest that, first, potential consumers tend to follow the collective evaluation before they make purchase decision. Second, our research will extend the role of perceived risk on adoption behavior of information technologies [38]. We propose that under a risky shopping situation, consumers are less willing to take words of consensus information and more acceptable to various types of information.

Additionally, our research will shed light upon online marketing practices. With the direct and indirect effects of collective evaluation, marketers or executives should think about how to apply them on their product pages. As lower ratings’ presence can suffer from both direct and indirect effects of collective rating, sellers should think of ways to minimize the disadvantages. Instead of offering aggregated information of rating, it is worth trying to separate the one rating into several dimensions, such as ratings on product appearance, duration, sellers’ service, package delivery and so on. Moreover, in order to make the most use of positive WOM, marketers or sellers should provide more security or safety cues to reduce the risk perception of
potential consumers.

Limitation & Future Work
The emphasis of the present research is limited to the helpfulness perception of online consumers. However, future work could extend our idea on the adoption behaviors and the economic benefits of consensus or deviant WOM information. Also, in this study, we examine the moderation effect of performance risk on the relationship between WOM information and consumers’ perception. We acknowledge that other risk dimensions are left uninvestigated. Future research may address the problem by other risk facets and explore their impact on the consumers’ perception or behaviors towards information. In order to further generalize our idea, future research could also use multiple methodologies or apply to other contexts to investigate the idea of present study.


