

THE EFFECT OF QUANTITATIVE ELECTRONIC WORD OF MOUTH ON CONSUMER PERCEIVED PRODUCT QUALITY

Yingtao Shen, Stetson University
Shenyu Li, Shanghai University of Finance and Economics
Michelle DeMoss, Stetson University

ABSTRACT

Online shoppers today can easily share reviews of their product usage experiences on a third-party website. Many retail websites allow individuals to submit qualitative and/or quantitative reviews, using a typical five-point scale. With the increase in online shopping behaviors, the impact of these reviews, or electronic word of mouth, on consumers' decisions is drawing more attention. Yet how do these on-line consumer reviews impact consumer behavior? This study empirically tests how quantitative product reviews affect online shoppers' perception of product quality. We find that: (1) extreme ratings have a greater impact on prospective customers' perceived quality than do median ratings while negative reviews have greater impact than positive reviews; (2) none of the volume, valence and consensus of electronic word of mouth can predict perceived quality reliably; (3) there is no difference between the perceived product quality formed by Chinese and American subjects when judging product quality, based on the same quantitative online product reviews.

JEL: M31

KEYWORDS: Electronic Word of Mouth, Online Review, Perceived Quality of Products

INTRODUCTION

Buying online is inherently riskier to shoppers than buying at a traditional brick-and-mortar retail store. First, online shoppers cannot physically examine or try the products or consult salespeople before a purchase because shoppers and online vendors are typically removed from each other geographically (Alba et al., 1997). Compared with offline shoppers, online shoppers may feel more uncertain about the quality of products. Hence, they have a higher chance or risk of selecting brands with lower-than-expected quality. Second, it is typically more inconvenient to return unwanted products to online vendors than it is to conventional stores, although many online vendors currently do offer warranties and return-and-refund policies. However, most online vendors do not cover return shipping costs, which can be very high for bulky or heavy products. Third, it may take longer than the duration of the warranties to find flaws in certain products like digital cameras. When those products malfunction, it is harder for customers to dispute the issue with online vendors and have their products replaced or repaired.

A complete consumer decision-making process includes several steps: Need recognition, information search, pre-purchase evaluation of alternatives, actual purchase, consumption, post-consumption evaluation, and divestment (Blackwell et al., 2006). Once consumers recognize their needs, they begin searching for information about alternative products that may satisfy those unmet needs. In many situations, information search is the key step in determining a purchase decision. Online shoppers, like their offline counterparts, need to collect information about alternative products. Due to the earlier-mentioned higher risk and the lack of direct access both to the products and sellers, online shoppers tend to base their purchase decisions more often than not on information provided online unlike traditional shoppers who can visit a store and see a product directly (Dellarocas, 2003).

There are three different, but major, types of product information available to prospective online buyers of a product, i.e., product and transaction clause information provided by the sellers, reviews created by professional reviewers (e.g. *Consumer Reports*), and reviews posted by previous buyers. Since the last type of information is provided by other customers rather than by the sellers, it is called electronic word of mouth (eWOM).

eWOM is somewhat similar to the more traditional word of month concept, a typical example of which is to listen to one’s neighbors or friends about a product or service, such as a new car or an unknown restaurant. However, since eWOM is distributed online, it is very different from traditional WOM in several ways. First, eWOM offers a larger number of reviews than does traditional WOM (e.g. verbal face-to-face conversation among several friends or family members). Often, a product is rated by hundreds or even thousands of customers since it has been sold online. Second, eWOM lasts longer and may reach a much larger audience than traditional WOM, since it can remain posted on online marketplaces for a longer period of time and can be read by all online shoppers interested in the product. For example, on Amazon.com, it is not unusual to find reviews that have been posted for more than 5 years. Therefore, eWOM may have a much greater long-term impact on prospective customers’ choices. Third, eWOM is typically anonymous. For example, customer reviewers on Amazon often identify themselves with nicknames or initials from their names, plus the names of their presumable locations, such as “Nightwing, St. Louis, Mo., USA”. Fourth, eWOM, particularly narrative reviews, can be highly detailed and comprehensive. Many reviewers go to great lengths to report all the pros and cons of a product and will carefully document everything they have experienced with the product. Many reviews on Amazon are longer than 500 words, and some even run to 1000 words. Finally and most importantly, eWOM is more quantitative in most cases than is traditional WOM.

Many major online market places, such as Amazon, eBay, and websites run by major brick-and-mortar retailers such as Wal-Mart and Best Buy, have adopted the same 5-point scale for their product rating system. This rating system allows customers who have bought and used a product to review it quantitatively by selecting a category on a 5-star scale to communicate their personal usage experience. These online reviews such as the example shown in Figure 1 can then be viewed by potential buyers when gathering information about a product or brand.

Figure 1: Sample Quantitative Product Review on a 5-star Scale on Amazon.com

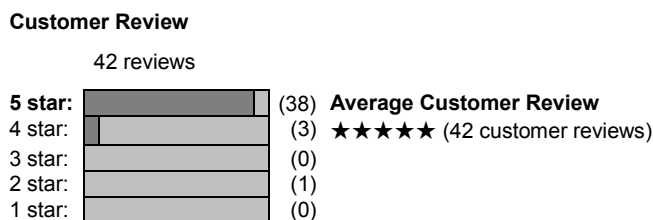


Figure 1 shows an example of these quantitative reviews on a product shown on Amazon.com. This specific product has 42 reviews in its history with an average of 4.9 (when the customer’s mouse points on or clicks on the 5 stars on the right side of the screen, a pop-up window shows the average), with 38 customers selecting 5-star, 3 customers 4-star, and 1 consumer 2-star. When a new review is added, Amazon automatically updates the total number of the reviews and the average rating.

The introduction of the 5-point scale has successfully transformed eWOM from collections of unstructured, qualitative opinions into structured, quantitative surveys with large and continuously increasing sample sizes. As the example in Figure 1 demonstrates, by showing number of reviews, average review and distribution of reviews, this new 5-point rating system clearly delivers the three major characteristics of WOM- volume, valence and consensus- in a highly effective way. According to Khare, Labrecque, and Asare (2011), volume is defined as the number of reviews provided by consumers; valence conveys an aggregate opinion (positiveness vs. negativeness) of the quality of a target to the WOM user; and consensus refers to the level of agreement in reviewers’ ratings for a target product.

When a prospective customer is searching for information about alternative products at these online marketplaces, she can easily read these quantitative reviews on the products. When comparing multiple alternatives, these quantitative peer reviews should influence her choice of brand, because these quantitative ratings provide her with a more seemingly precise measure of the products, and very importantly, facilitate direct and relatively easier comparison among multiple brands. For example, an online shopper who wants to buy a MP4 player for the first time may find one brand has 120 reviews with an average of 4.0 out of 5 with 50 5-star, 40 4-star, 20 3-star, 5 2-star, and 5 1-star reviews; another brand offers 60 reviews with an average of 4.3 out of 5 with 30 5-star, 20 4-star, and 10 3-star reviews. These two products have different eWOM in terms of volume, valence, and consensus. So an interesting question arises: How does she make her choice based on these quantitative reviews? So far, this important question remains largely unanswered.

Despite the increasing popularity and proliferation of this new 5-star product rating system described above, this new element of the promotion mix has not received academic attention. This study's goal is to empirically measure how this rating system affects customer's perceived product quality. Specifically, we begin the paper with a current literature review of WOM and present our research objectives. We then introduce the research methodology, followed by analysis of the results. In the last section, we summarize the findings and analyze the managerial implications and limitations of this study.

LITERATURE REVIEW AND RESEARCH OBJECTIVES

Past research has shown that traditional WOM can significantly influence consumer choice of products (Katz and Lazarsfeld, 1955; Engel et al., 1969; Richins and Root-Shaffer, 1988). Consumers tend to believe that their peers' reviews are more credible and trustworthy than messages from marketers (e.g., Day, 1971) because these reviews are independent of marketer selling intent (Silverman, 1997) and more influential than neutral print sources, such as *Consumer Reports* (Herr et al., 1991).

Recently, online consumers were found to rely on eWOM to make their buying decisions. A survey conducted by Opinion Research Corporation in 2008 suggests that 83% of respondents indicated that online product evaluations and reviews had at least a certain level of influence on their purchasing decisions, and about 70% of respondents had sought information online for a product or service before a purchase. Chevalier and Mayzlin (2006) found that eWOM has had a significant impact on sales of books at Amazon.com and BN.com. Liu (2006) found that eWOM also determines box office movie revenue. Korgaonkar and Karson (2007) found that even people's preference for online retail format is influenced by eWOM.

In the WOM literature, a variety of different facets of WOM have been studied. Prior studies have examined non-message characteristics, including provider characteristics (Bone, 1995; Hu, Liu, and Zhang, 2008), recipient characteristics (Bansal and Voyer, 2000), situational and product characteristics (Hogan, Lemon, and Libai, 2004), and message characteristics including volume (Chevalier and Mayzlin, 2006), valence (Liu, 2006; Duan, Gu, and Whinston, 2008), and consensus (West and Broniarczyk, 1998).

Relatively fewer studies have looked into the interaction between these WOM characteristics, particularly the three important message characteristics of volume, valence, and consensus. West and Broniarczyk (1998) conclude that people evaluate a target more favorably when WOM consensus is high and valence is higher than their aspiration level. Khare et al. (2011) show that when WOM consensus is high, high-volume WOM worsens the evaluation of an already negatively-rated WOM target and improves the evaluation of an already positively-rated WOM target, because volume in WOM is a high-scope, decision-making cue.

However, for a brand under review by a 5-star system, when the number of ratings at different categories (stars) changes, its total number of reviews, average of reviews, and distribution of reviews will all change. Thus, the three important characteristics of WOM, namely, volume, valence, and consensus, will also change simultaneously. Research has yet to examine how the interaction of these three major WOM characteristics using the new 5-star rating system affects individual perceptions of product quality.

In this empirical study, the following research questions regarding the 5-star scale were addressed: (1) Are the 5 different categories equally important in determining people’s perceived quality? If not, which categories are more important than others? (2) When the number of reviews at different categories/stars varies, what is the overall effect of all of the quantitative reviews on a product?

RESEARCH METHOD

Conjoint Analysis

We used conjoint analysis as the primary data analysis technique for this study. Conjoint analysis is a widely used marketing research multivariate technique and particularly suitable for investigating how consumer preferences are determined via different categorical independent variables, such as product attributes and price levels. Conjoint analysis assumes that people’s preference or utility of a product can be broken into separate amounts of utility or disutility for each separate attribute of that product, a concept called part-worth. To do conjoint analysis, typically researchers first create a set of hypothetical product profiles with their attributes at different levels; then these product profiles are presented to and evaluated by participants who either rate or rank these profiles according to their preferences. The preferences are then be decomposed into different part-worths to determine how important each product attribute actually is. The greatest advantage of conjoint analysis is that it mimics actual decision-making in the real world where people will select among multiple products, with none being the perfect option, by forcing participants to trade off product attributes in the same way consumers presumably do in the real world.

Conjoint analysis is particularly suitable for examining the effect of quantitative reviews on consumers’ perception of quality. A given quantitative review with different numbers (such as 10, 20, or 50 reviews) at 5 different stars is essentially like a product with 5 different attributes (such as size) at different levels (e.g., small, medium, and large). By manipulating the numbers at different stars, we can create hypothetical quantitative review profiles. Comparing these reviews entails a trade-off among different attributes (categories or stars) which mimic what online consumers do in the real world. Therefore, we can ask people to rank these review profiles according to the perceived quality of them and then use conjoint analysis to analyze the ranking data to determine how important each attribute (star) and each of its different levels is to their perceived quality of the product.

Table 1: Attributes and Levels

Rating	Number of Reviews
5-star	30, 50
4-star	20, 40
3-star	0, 30
2-star	0, 20
1-star	0,10

Table 1 shows the attributes/categories and their corresponding levels taken into consideration for the conjoint analysis. Each attribute/category has two levels. For example, 5-star has two levels, i.e., 30 or 50 reviews.

To create hypothetic review profiles, we select two different selected levels (number of reviews) for each of the 5 different attributes (categories or stars) as shown in Table 1. These 5 attributes and the 2 levels for each will generate $2 \times 2 \times 2 \times 2 \times 2 = 32$ profiles with averages varying between 3.4 and 4.7. A general

search on major online marketplaces, such as Amazon.com, revealed that review averages of most products sold online will vary between 3 and 5 with a significant portion of them scoring higher than 4 (it is intuitive because sellers are more likely to drop the unpopular products that earn low average reviews), so our selection of levels actually reflects this reality.

Table 2 shows the hypothetical brand review profiles whose design closely resembles that of the reviews posted on most online marketplaces. We decided to select 11 of the 32 possible brand profiles with averages of reviews varying between 4.0 and 4.4, for three reasons. First, ranking all 32 profiles is simply too burdensome for most average participants. Second, we tried to avoid those profiles with review averages that were too low or too high to better “force” participants to trade off among different categories that generate relatively similar review averages. Third, West and Broniarczyk (1998) show that when the valence of WOM is higher than people’s expectation level, the same WOM influences their preference for the target in a different way than when the valence of WOM is lower than people’s expectation level. By making the averages of all 11 profiles equal to or greater than 4, we expect participants to employ the same heuristic when they compare these review profiles.

Table 2: Hypothetical Brand Review Profiles

Brand	Number of Reviews					Volume: Total Number of Reviews	Valence: Average	Consensus: Standard Deviation
	1-star	2-star	3-star	4-star	5-star			
A	10	0	0	20	30	60	4.0	1.4
B	0	20	0	20	50	90	4.1	1.2
C	0	0	30	40	50	120	4.2	0.8
D	10	0	0	20	50	80	4.3	1.3
E	0	20	0	40	50	110	4.1	1.1
F	10	0	0	40	30	80	4.0	1.2
G	0	0	30	20	30	80	4.0	0.9
H	10	0	0	40	50	100	4.2	1.2
I	0	0	30	20	50	100	4.2	0.9
J	0	0	30	40	30	100	4.0	0.8
K	0	0	0	40	30	70	4.4	0.5

Table 2 shows the 11 hypothetical brands compared in the conjoint analysis. For each brand profile, we can see the number of reviews at each of the five rating categories, the total number of reviews of all of the five categories, the average of all of its reviews, and the standard deviation of its reviews. For example, Brand A has 30 5-star, 20 4-star, 0 3-star, 0 2-star, and 10 1-star reviews; its total number of reviews is 60 ($=30+20+0+0+10$); the average and standard deviation of its reviews are 4.0 and 1.4 respectively.

Subject and Data Collection

In total, 125 undergraduate students participated in this study, including 68 students from a private university in the U.S. and 57 from a public university in China. Additionally, 100% of the participants reported that they had shopped online at least once in the past 6 months.

The surveys were administered on paper-based questionnaires during regular class time. The participants were asked to imagine that after they had decided to buy a hypothetical product online, they were browsing an online marketplace where they found 11 different alternative brands of the same product. These brands had similar features and prices. Each of these brands had been rated by previous buyers on a 5-point scale with 5 categories from 1-star to 5-star. The eleven brand profiles shown in Table 2 were presented to the participants on a single page of the questionnaire. For each brand profile, the participant could see the total number of reviews, the number of reviews for each of the 5 categories (stars), and the average of the reviews. Then the participants were instructed to rank the 11 brands in order according to

their perceived quality of each brand based on these given customer reviews. In total, 123 of the 125 participants finished the task of ranking the 11 brands.

RESULTS AND DISCUSSION

The data were cleaned for outliers and missing data and analyzed using SPSS 16.0. The estimated relative importance of the 5 categories and the part-worths of different levels of the categories are shown in Table 3.

Table 3: Estimated Part-worths for Different Rating Categories

Rating	Levels	Part-Worths	Importance (%)
5-star	30	.000 (.000)	26.714
	50	4.467 (.192)	
4-star	20	.000 (.000)	10.486
	40	1.328 (.172)	
3-star	0	.000 (.000)	17.385
	30	-1.407 (.329)	
2-star	0	.000 (.000)	23.931
	20	-4.230 (.393)	
1-star	0	.000 (.000)	21.483
	10	-3.757 (.329)	
Pearson's R	.997***		
Kendall's tau	.991***		

Table 3 presents the results of the conjoint analysis on the ranking data from all of the subjects. For each attribute/category, it shows how much utility the subjects attached to each of its two levels and how important this attribute is in determining the subjects' preference of the 11 brands. For example, 50 5-star reviews on average have a utility of 4.467 while 40 4-star reviews have a utility of 1.328. Overall 5-star reviews' importance is 26.714% compared to the 10.486% of the 4-star reviews. ***indicates that the coefficient is statistically significant at 99% level of confidence, suggesting high goodness of fit of the estimated conjoint model. Standard errors are in parentheses.

According to Table 3, different rating categories (attributes) do have a significantly different impact on people's perceived product quality. Evidently, 5-star is the most important attribute/category (26.714%), followed by 2-star (23.931%), 1-star (21.483%), 3-star (17.385%), and 4-star (10.486%). However, taking into account the range of reviews (high number of reviews minus low number of reviews) of these attributes/categories, we reach somewhat different conclusions. Most noticeable is that for only half of the range of the number of reviews (0 to 10), 1-star is 80.4% (21.483% vs. 26.714%) and 89.8% (21.483% vs. 23.931%) as important as 5- and 2- star reviews, respectively. We can make a reasonable conjecture, therefore, that if we increase the range of 1- star reviews from 10 to 20, it is very likely that 1-star reviews will have a greater impact on product quality than 2- and 5- star reviews. With the same range of 20, 5-star (changes from 30 to 50 reviews) is 11.6% (26.714% vs. 23.931%) more important than 2-star (changes from 0 to 20). With a much greater range (from 0 to 30 reviews), 3-star is significantly less important than 1-, 2- and 5- star (have ranges of 10, 20 and 20 respectively). Four-star (changes from 20 to 40) is the least important factor with an importance of only 10.486%.

This evidence indicates that when forming their perceptions of these brands, people weighted extreme reviews (1- and 5-star) greater than median ones and weighted negative reviews more than they did positive ones (1- vs. 5-star and 2- vs. 4-star). In the marketing literature, this phenomenon is called extremity bias and negativity bias. Skowronski and Carlston (1989) and Lalwani (2006) summarize the four major explanations in the marketing and psychology literature. Expectancy-contrast theories argue that people's perceptual anchor is typically between a neutral point and the positive end of their judgment scale, i.e., people tend to think somewhat positively about most targets. Therefore, extreme or negative reviews will stand out because they deviate from this anchor. Frequency-weight theories contend that extreme or negative stimuli receive more attention because they are perceived to be more novel, different, or unique than the other stimuli. Category-diagnostics theories suggest that people view extreme or negative information as being more helpful and diagnostic than other information. Range theories argue

that different information cues will have different levels of ambiguity; extreme cues are perceived as less ambiguous than other cues and negative cues are less ambiguous than positive cues.

Assuming that customers will buy only one of these 11 brands, it is interesting to compare the “market share” of each brand (defined as the percentage of respondents who ranked that brand first) as shown in Table 4. Brand C, having an average review of only 4.2, is clearly the market leader, selected by 60.2% of the subjects. Brands K and D rank second and third, respectively, with much smaller shares (17.1% and 7.3%). A comparison of Table 4 with Table 2 shows that neither volume, valence, nor consensus can reliably predict people’s preferences. With higher volume (120 vs.70), but lower valence (4.2 vs. 4.4) and lower consensus (0.8 vs. 0.5), Brand C is much more favored than is Brand K. However, with lower volume (70 vs. 80), but higher valence (4.4 vs. 4.3) and consensus (0.5 vs. 1.3), Brand K is more favored than Brand D.

Table 4: Percentage of Buyers Indicating They Would Select Different Products as Their First Choice

Brand	A	B	C	D	E	F	G	H	I	J	K
Percentage (%)	0.0	2.4	60.2	7.3	4.1	0.0	0.0	4.9	4.1	0.0	17.1

This table shows the “market shares” of the 11 brands among the subjects, assuming that the subjects will buy their first choice only. For example, 60.2% of the subjects should select Brand C over other brands.

Of our total number of 123 valid observations, 57 (46.3%) were from China and 66 (53.7%) from the United States. Although these two countries are both large ecommerce markets, they have very different economic, business, political, and natural environments. According to a report from BCG in 2011, there are also a number of differences between the online business environments of the two countries in terms of the number and growth rate of online shoppers, shopper familiarity with e-shopping, online shopping expenditures, consumer trust, delivery and logistics, etc. Thus, we felt it was important to analyze the two datasets from these two highly important and yet different markets independently and compare the results to uncover any differences.

Table 5: Comparison of Part-worths and Importance - U.S. vs. China

Rating	Levels	Part-worths		Importance (%)	
		American Participants	Chinese Participants	American Participants	Chinese Participants
5-star	30	0.000 (0.000)	0.000 (0.000)	26.067	27.399
	50	4.337 (0.219)	4.618 (0.268)		
4-star	20	0.000 (0.000)	0.000 (0.000)	10.657	10.252
	40	1.300 (0.196)	1.361 (0.240)		
3-star	0	0.000 (0.000)	0.000 (0.000)	18.632	16.039
	30	-1.439 (0.375)	-1.370 (0.460)		
2-star	0	0.000 (0.000)	0.000 (0.000)	23.178	24.826
	20	-4.293 (0.448)	-4.157 (0.549)		
1-star	0	0.000 (0.000)	0.000 (0.000)	21.466	21.484
	10	-3.738 (0.375)	-3.778 (0.460)		
	Pearson's R	0.996***	0.994***		
	Kendall's tau	0.964***	0.964***		

*Table 5 presents the results of the conjoint analysis on the ranking data from American and Chinese subjects respectively. For each attribute/category, it shows how much utility each of the two groups attached to each of its two levels and how important this attribute is in determining each group’s preference of the 11 brands. For example, 50 5-star reviews on average have a utility of 4.337 and 4.618, and importance of 26.067% and 27.399% to the American and Chinese subjects respectively. ***indicates that the coefficient is statistically significant at 99% level of confidence, suggesting high goodness of fit of the estimated conjoint model. Standard errors are in parentheses.*

Table 5 compares the results of the conjoint analysis on the two datasets from China and the U.S. The results clearly show that the part-worths, reflecting the importance of the different stars or attributes, are highly similar across these two countries. To both groups, 5-star is the most important attribute/category (26.067% and 27.399%), followed by 2-star (23.178% and 24.826%), 1-star (21.466% and 21.484%), 3-star (18.632% and 16.039%), and 4-star (10.657% and 10.252%).

Table 6: Percentage of First Choices for Different Brands - U.S. vs. China

Brand	A	B	C	D	E	F	G	H	I	J	K
American Participants	0.0	3.0	59.1	9.1	4.5	0.0	0.0	6.1	3.0	0.0	15.2
Chinese Participants	0.0	1.8	61.4	5.3	3.5	0.0	0.0	3.5	5.3	0.0	19.3

This table compares the “market shares” of the 11 brands among the subjects from the U.S. and China, assuming that the subjects will buy their first choice only. For example, 58.2% and 61.4% of American and Chinese subjects should select Brand C over other brands.

Table 6 shows the first choice shares of the 11 brands among American and Chinese participants, respectively. Again the first-choice shares of these brands are highly similar. In both U.S. and China, Brand C is the most popular one with share of approximately 60%. Brand K is the second most popular choice to both groups with share of 15.2% and 19.3% respectively, followed by Brand D (9.1% and 5.3%). None in the two groups selected Brand A, F, G, or J as the first choice.

The results clearly suggest that the American and Chinese participants utilize the 5-star rating system to form their perceived product quality in similar ways. These findings indicate that the data are highly robust, and indeed, the validity of the data does not appear to depend on economic, social, cultural, or other environmental factors. In addition, these findings provide further support that consumers’ preferences cannot be reliably predicted by volume, valence, or consensus.

CONCLUSIONS

eWOM is a relatively new phenomenon in marketing, and therefore, a rich field for consumer behavior research. This study examined how electronic word of mouth impacts individuals’ perceptions of product quality using the new 5-star rating system. Survey data was collected from undergraduate students attending two different universities in the U.S. and China. Participants were asked to rank 11 brands according to their perceived quality of each brand based on given 5-star reviews.

A conjoint analysis of the data reveals different rating categories of the 5-star review do have a significantly different impact on people’s perceived product quality. Specifically the results indicate that when individuals form perceptions of quality based upon the 5-star rating system, they weight extreme reviews (1- and 5-star) greater than median ones and weight negative reviews more than they did positive ones (1- vs. 5-star and 2- vs. 4-star). However, when examining overall ranking of brands, neither volume, valence, nor consensus can reliably predict people’s preference. In addition, when comparing U.S. and Chinese participants, the analysis suggests similarities in how individuals utilize the 5-star rating system to form perceived product quality.

Managerial Implications

This study has significant practical implications for online marketing management in terms of customer relationship management and formulating target market strategies. For those marketers who sell their products online or whose customers can post product reviews online, this study illustrates the greater significance of customer satisfaction or dissatisfaction. Particularly, it is not only important to make customers satisfied, but it is even more important to address concerns of customers who are highly dissatisfied. This finding is especially important with eWOM given its prevalence, greater long-term impact, and easy accessibility for consumers.

When a marketer is targeting multiple market segments that are significantly different and cannot be simultaneously satisfied by the same product, it would be wise to design separate products for the separate segments or clearly communicate the value proposition to the intended target market through the product’s marketing mix. A one-size-fits-all product may make some users in each or all of the segments

highly dissatisfied, which may result in negative online reviews. Likewise, a confusing value proposition may lead to a very dissatisfied consumer spreading negative eWOM about a product that does not meet his needs. These negative online reviews may greatly deter prospective buyers and hurt the sales of the product overall.

From the prospective of managers of online marketplaces, this study suggests that the user online product review is highly important in terms of communicating to prospective buyers a perceived product quality. Therefore, customers should be allowed or even encouraged to post reviews of products online. Some online marketplaces only show total numbers and the average of quantitative reviews. Our findings suggest that this strategy is inadequate in aiding buyer purchase decisions. Thus full distribution of all quantitative reviews should be posted. Of even more interest, these findings are consistent across two of the world's major marketplaces, the US and China, suggesting that the study's implications have international applications.

Limitations and Directions for Future Research

Although participants from two different countries participated in this study, the generalizability of the results is somewhat limited due to the use of undergraduate students. However, this study provides a methodology that may be replicated in other populations with greater diversity in relevant demographics such as age and occupation. Future research is also needed to assess how additional online information may affect individual's perceived product quality. Participants in this study were given only the 5-star rating to determine product quality rankings. Specifically, research is needed that measures the potential impact of other informational cues such as price, brand names, product attributes, and seller reputation on online shoppers' purchase decisions. Future research should also measure how qualitative reviews affect perceived quality, when combined with quantitative reviews. It would be of particular interest to measure the relative importance of quantitative and qualitative reviews when these two types of reviews are not perfectly consistent with each other or when they are perceived to be unequally diagnostic about the quality of a product.

REFERENCES

- Alba, J., Lynch, J.W.B., Janiszewski, C., Lutz, R., Sawyer, A., and Wood, S. (1997) "Interactive Home Shopping: Consumer, Retailer, and Manufacturer Incentives to Participate in Electronic Marketplace," *Journal of Marketing*, 61 (3) 38-53.
- Bansal, H.S. and Voyer, P.A. (2000) "World-of-Mouth Processes within Services Purchase Decision Context," *Journal of Service Research*, 3(2), 166-177.
- Blackwell, R., Miniard, P., and Engel, J. (2006) "Consumer Behavior," South-Western.
- Bone, F. (1995) "Word-of-Mouth Effects on Short-Term and Long-Term Product Judgments," *Journal of Business Research*, 32 (3), 213-223.
- Boston Consulting Group (2011) "The World's Next E-Commerce Superpower: Navigating China's Unique Online-Shopping Ecosystem," Retrieved December 20, 2011 from the BCG Web site: <http://www.bcg.com/documents/file91905.pdf>.
- Business Wire (2008) "Online Consumer Reviews Significantly Impact Consumer Purchasing Decisions, New Opinion Research Corporation Survey Finds," Retrieved December 12, 2011 from the Reuters Web site: <http://www.reuters.com/article/2008/06/24/idUS117426+24-Jun-2008+BW20080624>.

Chevalier, J. and Mayzlin, D. (2006) "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research*, 43, 345-354.

Day, G.S. (1971) "Attitude Change, Media and Word of Mouth," *Journal of Advertising Research*, 11, 31-40.

Dellarocas, C. (2003) "The Digitization of Word-of-Mouth: Promise and Challenges of Online Reputation Systems," *Management Science*, 49 (10), 1407-1424.

Duan, W., Gu, B., and Whinston, A. (2008) "The Dynamics of Online Word-of-Mouth and Product Sales—An Empirical Investigation of the Movie Industry," *Journal of Retailing*, 84 (2), 233-42.

Engel, J.F., Kegerris, R.J., and Blackwell, R.D. (1969) "Word of Mouth Communication by the Innovator," *Journal of Marketing*, 33, 15-19.

Herr, P., Kardes, F., and Kim, J. (1991) "Effects of Word-of-Mouth and Product-Attribute Information on Persuasion: An Accessibility–Diagnosticity Perspective," *Journal of Consumer Research*, 17 (4), 454-62.

Hogan, J., Lemon, K., and Libai, B. (2004) "Quantifying the Ripple: Word of Mouth and Advertising Effectiveness," *Journal of Advertising Research*, 44 (3), 271-80.

Hu, N., Liu, L., and Zhang, J. (2008) "Do Online Reviews Affect Product Sales? The Role of Reviewer Characteristics and Temporal Effects," *Information Technology and Management*, 9 (3), 201-14.

Katz, E. and Lazarsfeld, P. F. (1955) "Personal Influence: The Part Played by People in the Flow of Mass Communication," Glencoe, IL: Free Press.

Khare, A., Labrecque, L., and Asare A. (2011) "The Assimilative and Contrastive Effects of Word-of-Mouth Volume: An Experimental Examination of Online Consumer Ratings," *Journal of Retailing*, 87 (1) 111-126.

Korgaonkar, P. and Karson, E. (2007) "The Influence of Perceived Product Risk on Consumers' e-Tailer Shopping Preference," *Journal of Business Psychology*, 22, 55-64.

Lalwani, A. (2006) "Negativity and Positivity Biases in Product Evaluations: The Impact of Consumer Goals and Prior Attitudes," zibs.com, Accessed December 10, 2011 at <http://www.zibs.com/techreports/Negativity%20and%20Positivity%20Biases%20in%20Product%20Evaluations.pdf>.

Liu, Y. (2006) "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue," *Journal of Marketing*, 70, 74-89.

Richins, M. and Root-Shaffer, T. (1988) "The Role of Evolvement and Opinion Leadership in Consumer Word-of-Mouth: an Implicit Model Made Explicit," *Advances in Consumer Research*, 15, 32-36.

Silverman, G. (1997) "How to Harness the Awesome Power of Word of Mouth", *Direct Marketing*, November, 32-37.

Skowronski, J. and Carlston, D. (1989) "Negativity and Extremity Biases in Impression Formation: A Review of Explanations," *Psychological Bulletin*, 105, 131-142.

West, P. and Broniarczyk, S. (1998) "Integrating Multiple Opinions: the Role of Aspiration Level on Consumer Response to Critic Consensus," *Journal of Consumer Research*, 25 (1), 38-51.

ACKNOWLEDGEMENT

This study was supported by the 2012 Summer Research Grant awarded by Stetson University. The authors thank the editor and the two anonymous reviewers for helpful comments and suggestions.

BIOGRAPHY

Yingtao Shen, Ph.D., is an Assistant Professor of Marketing at Stetson University in Deland, Florida. His research appears in journals such as *Advances in Consumer Research*, *Organizational Behavior and Human Decision Processes*, and *Marketing Letters*. He can be reached at Stetson University, 421 North Woodland Blvd., Deland, FL 32723, yshen@stetson.edu.

Shenyu Li, Ph.D., is an Assistant Professor of Marketing at Shanghai University of Finance and Economics in Shanghai, China. His research interests include online marketing, social network theory and modeling, and marketing channel strategy. He can be reached at Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai, China 200433, li.shenyu@shufe.edu.cn.

Michelle DeMoss, Ph.D., is the Chair of the Marketing Department and holds the Dennis C. McNamara, Sr. Endowed Chair of Marketing at Stetson University's School of Business Administration in DeLand, FL. Her research appears in journals such as the *Journal of Consumer Behavior*, *Journal of Business Ethics* and *Journal of Business Education*. She can be reached at Stetson University, 421 North Woodland Blvd., Deland, FL 32723, mdemoss@stetson.edu.

