International Journal of Business and Management; Vol. 13, No. 4; 2018
ISSN 1833-3850 E-ISSN 1833-8119
Published by Canadian Center of Science and Education

More Is Less: Only Moderate Polarized Online Product Reviews can Affect Sales

Li Yang[1]

[1] School of Economics and Management, Tsinghua University; Pan-Asia Business School, Yunnan Normal University, China

Correspondence: Li Yang, School of Economics and Management, Tsinghua University, Beijing; Pan-Asia Business School, Yunnan Normal University, Kunming, China., E-mail: yangli@ynnu.edu.cn

Received: February 6, 2018 Accepted: March 6, 2018 Online Published: March 18, 2018

doi:10.5539/ijbm.v13n4p192 URL: https://doi.org/10.5539/ijbm.v13n4p192

Abstract

It is widely proved that positive online word-of-mouth (WOM) can boost sales and negative online WOM harm sales. Then will more positivity or negativity of messages in online product reviews text have greater impact on product sales? This research attempts to tackle this ignored research question. The answer is counter-intuitive: it depends on how positive or negative they are! The results of a two-way fixed-effects panel data analysis based on the data about tablet market in Amazon and a novel sentiment analysis technique demonstrate that the most and least polarized online product reviews actually have no effect on sales and only moderate positive / negative reviews can affect sales. Such effects can be explained by the optimal arousal theory and attribution theory. Inspired by the findings, three strategies for user-generated content (UGC) management are proposed.

Keywords: word-of-mouth, product reviews, consumer reviews, customer reviews, user-generated content, product sales

1. Introduction

With the growing popularity of Internet applications like social media and electronic commerce, consumers rely more on online word of mouth (WOM) to get rich information about products before making a deal. Online product reviews, as a new type of WOM, help potential customers infer products' quality and sellers' reputation from others' real consumption experience, and finally find the product that best meet their needs. For consumers, online review is a new source of product information and for sellers, it is a new element of marketing communication mix (Chen & Xie, 2008). It has the potential to be a kind of effective and manageable marketing resource for sellers. Therefore, dozens of studies to examine the characteristics and effects of online WOM emerged in recent years, and most of them focus on its volume's and valence's impact on sales (Rosario, Sotgiu, De Valck, & Bijmolt, 2016; Ya, Vadakkepatt, & Joshi, 2015).

Valence of online WOM is recognized as a key feature of WOM to affect sales across different product categories by most of such empirical studies (Floyd, Freling, Alhoqail, Cho, & Freling, 2014). It is found that positive WOM can boost sales and negative WOM will harm sales. The argument behind this proposition is that people are influenced by others and may favor information inferred from others' actions (Godes & Mayzlin, 2004), and positive WOM will enhance customers' quality expectations and positive attitudes towards a product and the effect of negative WOM will go the opposite way (Y. Liu, 2006; Tang, Fang, & Feng, 2014). If that is the case, the extent of positivity or negativity of WOM may also affect sales because more positive WOM may result in higher quality expectations, generate more positive attitude towards the product and thus increase the possibility to buy the product. Following the same logic, more negative WOM will decrease the possibility of purchase. More positive or negative WOM may have greater impact on sales. Such inference also follows our intuitions and most relevant empirical studies find that ratings positively affect sales. Although ratings partly represent positivity / negativity, it is too rough to use rating as the proxy of positivity or negativity for three reasons. Firstly, the text of reviews with the same rating often expresses widely different positivity / negativity. Take two pieces of online product reviews both of which have five-star ratings from the data used in this research as an example:

"This is an amazing product. I am really loving it!!!"

"This is a good tablet PC for the price."

Obviously, the former expresses more positivity and satisfaction with the product than the latter, and implies better quality and consumption experience. The former seems to be a stronger stimulate to drive the sales. Secondly, one dimensional solution as ratings to measure positivity / negativity is challenged because many reviews have both positive and negative components simultaneously and independently (Tang et al., 2014). Thirdly, the results of prevalent five-star+ rating systems in most websites are highly unbalanced. Product review ratings in Amazon, the data source of most online WOM studies, are positively biased. Chen, Wang, and Xie (2011) and Chevalier and Mayzlin (2006) report over 50% five-star reviews and over 70% four and five-star reviews. Data in current research include 47% five-star reviews and 21% four-star reviews. The reference point of being neutral is questionable. Therefore, it is hard to tell how many stars represent highly positive and moderately positive. Moreover, to distinguish positive, negative and neutral reviews based on one to five-star ratings is actually a confusing and arbitrary job. Many studies such as Chen et al. (2011) and Chevalier and Mayzlin (2006) regard one-star as negative and five-star as positive. But how about two, three and four stars? What do they mean by positivity and negativity?

Text of product reviews contains richer information about consumers' preference and attitude than ratings do. It seems to be a better solution to extract the information about valence polarity from the review text rather than from the ratings. High cost of manual encoding for mass product reviews deters researchers from doing so. Sentiment analysis (B. Liu & Zhang, 2012; Pang & Lee, 2008) as an innovative tool based on fast evolving technology of natural language processing and text mining, shows the potential to help them out. The new technique to assess the sentiment strength of online communications text (Thelwall, Buckley, & Paltoglou, 2012) gives us a chance to evaluate the effects of varying positivity and negativity inside review text. Does online product review text with more positivity / negativity have greater impact on sales? As far as the literature search of this research goes, no studies ever systematically answered the question. This research adopted the sentiment strength measurement tool created by Thelwall et al. (2012) to tackle the question.

Contrary to the intuitions, the result of current research demonstrates that more positivity or negativity in the text of online product reviews does not always bring about higher sales volume. Actually, more is less: the most and least polarized messages in online product reviews text have no statistically significant effect on sales and only moderate positive / negative messages can affect sales. Such effects can be explained by the optimal arousal theory and attribution theory.

In the online business world, sellers are trying to manage electronic WOM including online product reviews and they can publish glowing reviews as those well-satisfied customers do or promote real reviews full of praise to the top of review lists. However, given such finding, it may not be a smart choice. On the other hand, those reviews full of blames may not be so terrible as many imagined that sellers have to detoxify them as soon as possible. The finding of this research is enlightening for marketing practices.

2. Theoretical Background and Hypotheses

In explaining the mechanism behind the effects of online WOM on sales, researchers borrowed from the advertising literature. The effects are regarded as the results of persuasion and information processing efforts. It says that the effects of online WOM on sales are influenced by the psychological mechanisms of persuasion and information processing. Consumers' reactions to valence of online WOM are also affected by how consumers process the online WOM messages.

One of the theoretical frameworks may shed light on the research question is the optimal arousal theory. According to the theory, the extent of arousal is based on a discrepancy from the "adaptation level". Minor deviations from the adaptation level (either a little more or a little less) can generate the most effective affect, while large discrepancies will weaken its effectiveness (Crowley & Hoyer, 1994). Adaptation level in the theory is equal to a value of stimulus eliciting the neutral response (Helson, 1947). Stimulus above this level is assumed to produce positive responses and that below negative responses. Thus, the adaptation level in the online WOM context should be initialized as being neutral, which means neutral product reviews. Moderate discrepancies from the adaptation level are assumed to increase consumers' motivations to process product messages, while large discrepancies decrease them (Crowley & Hoyer, 1994). Increased motivations to process product review messages will improve consumers' involvement, leading them to absorb more information from the messages (Tang et al., 2014) and thus, such product reviews can make a greater impact on consumers' decisions.

Another theory framework aspiring this research is the attribution theory. People tend to assign causes to events to justify others' actions. Three basis criteria that individuals use to process the information about others' actions and make attributions are choice, commonality and desirability (Mizerski, Golden, & Kernan, 1979). Choice means individuals are assumed to have a choice among actions (or inaction). Commonality means only "noncommon effects" are useful for inferring personal causality. The desirability rule says that the more undesirable the action or the effects of the action, the more readily and more confidently causality can be inferred. In the context of online product reviews, consumers always have choices to write positive, negative or neutral reviews and it is up to consumers' wills. So attributions could happen. Positive or negative reviews can be attributed to customer taste dissimilarity or product quality (He & Bond, 2015). Polarized reviews are more uncommon for consumers on average, which may induce personal causality and they can be attributed to customers' personal taste dissimilarity rather than issues of product quality. On the other hand, deceptive reviews take a substantial share of all online product reviews (Anderson & Simester, 2014). Sellers are motivated to boast about their product quality and attack their rivals by means of online WOM and they tend to post polarized reviews. Moderate positivity / negativity is somewhat undesirable for such kind of deceptive reviews and consumers may consider these reviews as real customer reviews more confidently.

Moderate positive / negative reviews can induce optimal level of arousal and are the most effective in motivating review readers to process online product reviews. It may be less possible to attribute moderate positive / negative reviews to personal taste dissimilarity or deception, but more possible to product quality.

Slightly positive / negative reviews can induce arousal below the optimal level. They will bring less novelty to consumers and induce less motivations to process the information inside the review messages. Less positivity / negativity also lower readers' expectations for product quality.

H1: Slightly polarized online product reviews make smaller impact on sales than moderate polarized ones do. Under certain conditions, the effects may disappear.

Extremely positive / negative reviews can induce arousal above the optimal level. They will bring too much novelty to consumers and bring about more doubt on credibility of reviews and induce less motivations to process the information inside the review messages. It may be more possible to attribute extremely positive / negative reviews to personal taste dissimilarity or deception. More positivity / negativity may not bring about greater effects but less.

H2: Extremely polarized online product reviews make smaller impact on sales than moderate polarized ones do. Under certain conditions, the effects may disappear.

3. Data, Variables and Measurement

The online product review data used in this research is from the data sets donated by Wang, Mai, and Chiang (2014) to marketing research society. The dynamic market information including weekly sales rank and product reviews are crawled from the largest online retailer, Amazon's tablets and tablet PCs category. The time window of the empirical research lasts 24 weeks starting from February 1, 2012. The product reviews posted before the starting time point are all assumed to be reviews of the first week. There are 40,741 product reviews about 794 tablet products from 39,278 reviewers in the raw data. After eliminating product items whose sales rank entries are missing, we finally put the data with 36,623 product reviews about 448 tablet products from 35,280 reviewers into this research. About 90% of reviews are retained.

3.1 Slightly, Moderately and Extremely Polarized Product Reviews

This research adopted the sentiment analysis approach (B. Liu & Zhang, 2012; Pang & Lee, 2008) developed from computer science to extract valence information from product reviews. The sentiment analysis can automatically classify unstructured text data into the categories of the positive, negative and neutral based on the semantic information contained in the text. The analytical tool called SentiStrength 2 and developed by Thelwall et al. (2012) is used in the valence classification task. This tool is also adopted in Tang et al. (2014), a study on effects of neutral WOM on sales.

SentiStrength 2 is a lexicon-based sentiment analysis tool. The core of this tool is a dictionary of affective words list including 575 positive words and 1,791 negative words with varying sentiment strengths. Considering the fact that each product review may contain both positive and negative meanings, the tool will report a positive sentiment strength score from 1 to 5 and a negative sentiment score from -1 to -5. Such a lexicon-based approach has some advantages over others and is suitable for the sentiment analysis task on short and informal online English texts in online product reviews. Firstly, SentiStrength 2 is created with the corpus from Internet applications such as YouTube, Twitter, and MySpace and its affective words list is adapted to Internet environment. Secondly, it can use nonlexical information very common in informal online communications scenarios. Meaningful signals such as emoticons (e.g., :-(, :-)), punctuation (!!!!!), all caps (e.g., "BEST") and repeated letters (e.g., "haaaappy," which is more positive than "happy") are considered in the calculation of sentiment strength scores. Therefore, its accuracy outperforms other competitive methods (Thelwall et al., 2012).

The product review is judged as positive when its valence score (positive strength score plus negative strength score) is over zero, as neutral when equal to zero and as negative when below zero. Since the valence score of positive reviews ranges from 1 to 4 and that of negative ones from -1 to -4, reviews with a valence score of 1 / -1 is classified as slightly positive / negative, 2 / -2 as moderate positive / negative, and above 2 / below -2 as extremely positive / negative. Following the measurement strategy adopted by Tang et al. (2014), a weekly count of slightly / moderate / extremely polarized reviews are applied to measure corresponding types of reviews. Average numbers of reviews of each week and each type above are plotted in Figure 1. Basic distributions and ratios of reviews of each type are visualized.

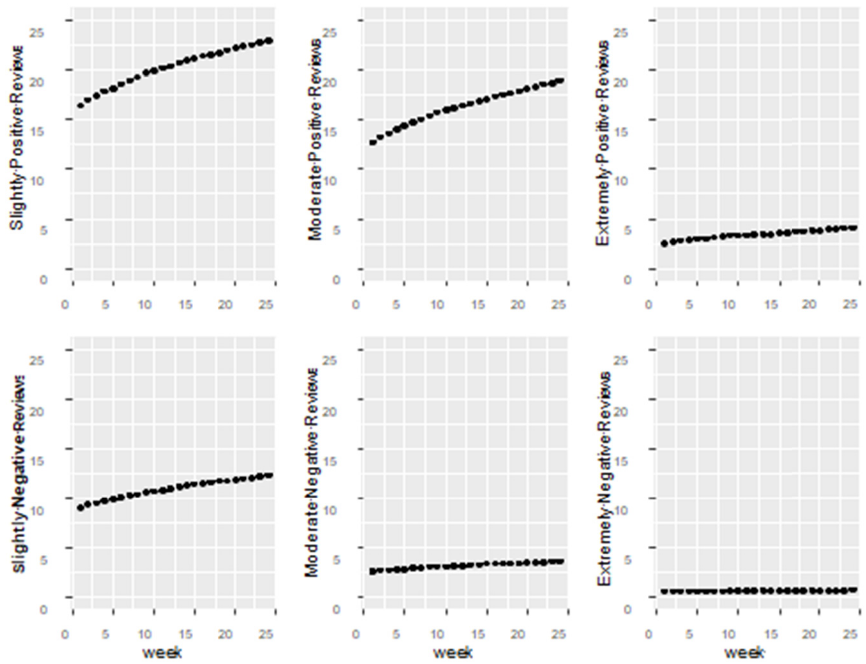

Figure 1. Time-Varying average numbers of reviews of each type

3.2 Control Variables

Positive strength and negative strength of product reviews, positive and negative strengths dispersion, neutral reviews and volumes, ratings and word counts of reviews are controlled in the empirical analysis. These factors are controlled for the possibility that they could affect product sales. The strength of positive and negative reviews, which are measured as the average positive / negative score in each week, represents the intensity of positive and negative reviews, while positive and negative strengths dispersion reflects the variance of opinions about a tablet product, which are measured as variance in positive or negative sentiment scores across all reviews for a product item in each week. The number of words in product reviews of each week is used to control for information volume. The effects of neutral reviews are also controlled. Since Tang et al. (2014) had pointed out that mixed-neutral WOM directly affects sales but indifferent WOM does not, neutral reviews are separated into the mixed-neutral and indifferent neutral and they are measured by weekly count of the two types of neutral reviews. Volumes are measured by the number of weekly reviews and ratings by average ratings.

Table 1. Statistical summary of variables

Variables	Mean	SD	Min	Max
Slightly Positive Reviews	20.32	220.75	0	5329
Slightly Negative Reviews	10.98	100.67	0	2414
Moderate Positive Reviews	16.38	209.21	0	5207
Moderate Negative Reviews	3.21	27.99	0	670
Extremely Positive Reviews	3.43	50.76	0	1261
Extremely Negative Reviews	0.69	6.27	0	151
Mixed Neutral Reviews	14.93	146.20	0	3457
Indifferent Neutral Reviews	1.40	14.02	0	347
Positive Strength	2.57	0.82	0	5
Negative Strength	-2.28	0.81	-5	0
Positive Strengths Dispersion	0.05	0.22	0	2.12
Negative Strengths Dispersion	0.07	0.29	0	2.83
Rating	3.067041	1.28	0	5
Volume	1.066406	14.72	0	659
Word Count	13475.83	126144.20	0	2968171
Log (Sales Rank)	7.75	1.90	0	14.54032

4. Model Specifications

Before obtaining the estimation results, we need to address several concerns about the appropriateness of model specifications and estimation methods to ensure the reliability of the results: potential reverse causality, nonstationarity, endogeneity, nonnormality, and heteroscedasticity.

For the reverse causality concern, we set a time lag between the explanatory variables and sales ranks to eliminate the potential for reverse causality between products reviews and sales rank. Such a time lag is reasonable in tablet product purchase scenarios because the purchase of tablets is unlikely an impulse buying and consumers typically undergo a process of decision making and finally make the order after several days' information searching.

To fit the basic assumptions of panel data analysis, we need to test the cointegration of variables in the model. A Choi's modified P Unit-Root Test (Pm = 168.3, $p < .001$) and Im-Pesaran-Shin Unit-Root Test ($z = -68.113$, $p < .001$) were conducted and both results show that residuals of the regression model are stationary time series and the estimation results can be trusted.

As for the endogeneity issues, a two-way fixed-effects model can reduce product-level unobserved factors such as prices, looks or brand images and time varying unobserved factors such as random marketing campaigns that could affect both reviews and product sales simultaneously, leading to endogeneity. Furthermore, an instrumental variable approach was adopted in this research to control the effects of endogeneity. Lagged dependent variables (sales ranks in the research) are often used as instrumental variables, which is valid for two reasons: they (1) are correlated with the endogenous explanatory variables but (2) not correlated with the error terms. As mentioned above, explanatory variables lagged by one period (one week) are used as instruments in the model, so the effects of endogeneity are further reduced.

To ensure normality of the dependent variable, we chose log (sales rank) as the explained variable in the model. Skewness (-0.4316) and kurtosis (1.1686) of log (sales rank) are calculated and they are within the range between the thresholds of ±2. The residuals of the model estimation were plotted and they are randomly scattered and evenly distributed around 0 and so heteroscedasticity is not an issue.

Besides, lagged log (sales rank) at period t – 1 was added to control for state dependence. To compare the results of a fixed- or random-effect model specification, we used a Hausman test ($\chi^2 = 3382.3, \mathrm{p} < .001$), which suggests fixed effect model is more appropriate.

Given the above, the following model is established to examine the effects of online product review with varying valence on sales:

$$\log(Sales\ Rank_{i,t+1}) = \beta_1 Slightly\ Positive\ Review_{i,t} + \beta_2 Moderate\ Positive\ Review_{i,t}$$
$$+\beta_3 Extremely\ Positive\ Review_{i,t} + \beta_4 Slightly\ Negative\ Review_{i,t}$$
$$+\beta_5 Moderate\ Negative\ Review_{i,t} + \beta_6 Extremely\ Negative\ Review_{i,t}$$
$$+\beta_7 Mixed\ Neutral\ Negative\ Review_{i,t} + \beta_8 Indifferent\ Neutral\ Review_{i,t}$$
$$+\beta_9\ Positive\ Strength_{i,t} + \beta_{10}\ Negative\ Strength_{i,t} + \beta_{11}\ Positive\ Dispersion_{i,t}$$
$$+\beta_{12}\ Volume_{i,t} + \beta_{13}\ Rating_{i,t} + \beta_{14}\ Word\ Count_{i,t} + \sum Week_t + \sum Product_i + \epsilon_{i,t}$$
$$+\log(Sales\ Rank_{i,t})$$

Where i represents a product and t denotes the time period (week in this research). Week and Product refer to two sets of time dummies and product dummies, which means a two-way fixed-effect model was estimated in this research.

5. Results

Table 2 demonstrates the results of the model estimation. As it shows, slightly positive / negative online product reviews do not have statistically significant effects on product sales ($p > 0.05$). Surprisingly, extremely positive / negative reviews do not affect sales either. The two hypotheses are proved.

Table 2. Two-Way fixed effects model estimation results

Variables	Coefficients	Standard Errors	t	P > t
Intercept	3.6642***	0.0866	42.30	0.000
Variables of Interest				
Slightly Positive Reviews	0.0052	0.0055	0.94	0.348
Slightly Negative Reviews	-0.0157	0.0096	-1.63	0.104
Moderate Positive Reviews	-0.0188***	0.0048	-3.91	0.000
Moderate Negative Reviews	0.0534***	0.0120	4.46	0.000
Extremely Positive Reviews	0.0002	0.0134	0.02	0.987
Extremely Negative Reviews	-0.0653	0.0344	-1.90	0.058
Control Variables				
Mixed Neutral Reviews	0.0182*	0.0080	2.27	0.023
Indifferent Neutral Reviews	-0.0215	0.0252	-0.85	0.393
Positive Strength	0.1310***	0.0304	4.30	0.000
Negative Strength	0.0362	0.0214	1.69	0.090
Positive Strengths Dispersion	-0.0720*	0.0320	-2.25	0.025
Negative Strengths Dispersion	-0.0633*	0.0257	-2.47	0.014
Rating	-0.0633***	0.0167	-3.79	0.000
Volume	0.0020	0.0024	0.84	0.400
Word Count	0.0000	0.0000	1.55	0.122
Log (Sales Rank)$_{t-1}$	0.4639***	0.0088	52.72	0.000
Model Fit Statistics				
F Test	204.724***			0.000
Adjusted R2	0.2131			

Notes. *$p < .05$ **$p < .01$ ***$p < .001$.

Moderate positive online product reviews ($\beta = -0.0188$, $p < .001$) positively (log (sales rank) is negatively related to sales volume) affects product sales and moderate negative reviews ($\beta = 0.0534$, $p < .001$) negatively affect sales, and the absolute value of the latter is 183% bigger than that of the former, which shows that negative reviews produce greater impacts than positive ones and the results are consistent with the findings of previous studies (Rosario et al., 2016; Ya et al., 2015). Slightly positive / negative reviews' impacts on product sales are statistically insignificant, possibly because less positivity just induces a little bit of expectations about high / low quality of the products and such effects may be completely offset by reduced motivation of processing review messages due to less arousal deviated from the optimal level. For extremely positive / negative reviews, induced concerns about credibility or personal taste biases may totally destroy the foundation of common mechanisms behind increased / decreased quality expectations due to positive / negative WOM. Consumers may attribute the extremely polarized online product reviews to deceptive reviews or personal biases rather than product quality issues. Too good to be true, and something looks too bad maybe is not so bad.

As mentioned above, this research dropped product items with missing entries of sales rank data. Missing data entries may be due to the fact that some product items had been unlisted or unsold before the research time window began, but they are listed and sold during the time window. So the deletion of such observations might result in sample selection biases. To check robustness of the estimation, we included all data entries and re-estimated the model with unbalanced panels. The results remained consistent, with no significant deviations.

6. Managerial Implications

Corresponding to firm-generated content (FGC) such as advertising, user generated content (UGC), including online WOM, refers to the content created and shared by Internet users. For consumers, UGC seems to be more credible, abundant and customer-oriented, and so it is more influential in driving sales than FGC (Goh, Heng, & Lin, 2013). Marketing managers are skilled in taking advantage of FGC to promote brand images and sales performance, but tend to lose control over surging UGC due to the lack of theories and best practices of UGC management. The findings of this research may give some hints towards better controls over UGC management from sellers' perspectives.

First of all, this research casts doubt on effectiveness of glowing content full of applause in the UGC context, which is proved to be valid in the traditional promotions based on FGC. Extremely positive UGC can lead to attributions towards deceptive or personal biased content and raise the issue of credibility. Even though it is supposed to boost sales, the actual results may be counterproductive.

Managerial implications of this research on UGC management involve what to manage and how to manage. As for what to manage, marketers should understand what the major opportunities and threats in the marketing environment with eruptive UGCs from social media, e-commerce websites, email lists and online forums. This research demonstrates that moderate positive UGC is the most effective in driving sales and so sellers need to make good use of opportunities with such kind of emerging UGCs to persuade consumers. Moderate negative UGC may be the most detrimental to product sales. Marketer should pay more attention to such kind of UGCs rather than extremely negative comments. The latter may be not so threatening as we imagined before.

The media containing UGCs is somewhat out of marketers' controls. But they still can take some measures to improve the management over them. Under some circumstances, they can pick some messages to be on the top of the columns, adjust the ranks of content lists or respond to specific comments in public as the official feedbacks. This research also inspires how to exploit such methods to reach the marketing goals. Which kind of content need to be promoted and which kind of content need to be suppressed is the key decisions to make. In the context of this research, it is assumed that most product reviews contain both positive and negative messages and the valence of product reviews is calculated in both the positive dimension and the negative dimension. It implies that moderate positive reviews possibly are those with much strengths and a little weakness. Such a two-sided persuasion strategy with both positive and negative ingredients makes the messages more credible (Crowley & Hoyer, 1994). Moderate positivity does not simply mean presenting a low profile on obvious strengths of products, but implies including a little criticism. Similarly, those UGCs with much drawbacks and a little advantage are the most harmful to the sales because of their high perceived credibility. Besides, different attributions lead to different results, which illustrates the importance of hints for attributions.

Following the discussion above, three strategies for UGC management could be raised. A highlighting strategy could be adopted for those moderate positive UGCs with sufficient praise and a little criticism. Credibility without evident distortions or personal biases should be the criteria in selecting the content deserving more highlights. For those moderate negative UGCs with a little praise and huge criticisms, the best choice may be a dimming strategy, which makes them as invisible as possible and sink to the bottom of the booming UGCs. A guidance strategy is preferred when some extremely positive / negative UGCs appears. Marketers need to induce the quality attributions for positive UGCs and the deception or personal bias attributions for negative UGCs by creating and spreading relevant hints deliberately.

7. Limitations and Future Research

This research focuses on the main effect of varying positivity and negativity of product review text on sales and relies on the data of tablet products in Amazon. Whether this effect is consistent under all circumstances deserves further explorations. As mentioned above, moderate positive / negative reviews have greater impact on sales than extremely positive / negative reviews because varying positivity and negativity relate to different level of efforts in information processing and issues of credibility and attributions. Therefore, more possible moderators of the effect need to be considered in future works.

Efforts of information processing on product review text may be affected by the risk of buying an unsuitable product. If the consequence of making a wrong decision is serious, consumers would be more careful during the purchase process and pay more attention to details of review text. So the values, durability, trialability and observability of the product may be significant moderators on the relationship between varying positivity and negativity of review text on sales.

As for credibility issues, other sources of credibility such as brand images of products or platforms may affect the credibility of product reviews. In an e-commerce platform with great reputations, consumers seldom have credibility concerns about product reviews in an official store of a famous brand also with perfect reputations.

Consumers often attribute extreme polarity of product reviews to personal tastes besides attributions to deceptive reviews. Hence hedonic and utilitarian products may have different stories and so do the products related and unrelated to fashions. These possibilities of research are left up to future research to explore the circumstances under which the findings of current research are valid.

References

Anderson, E. T., & Simester, D. I. (2014). Reviews Without a Purchase: Low Ratings, Loyal Customers, and Deception. *Journal of Marketing Research (JMR), 51*(3), 249-269. https://doi.org/10.1509/jmr.13.0209

Chen, Y., & Xie, J. (2008). Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix. *Management Science, 54*(3), 477-491.

Chen, Y., Wang, Q., & Xie, J. (2011). Online Social Interactions: A Natural Experiment on Word of Mouth Versus Observational Learning. *Journal of Marketing Research (JMR), 48*(2), 238-254. https://doi.org/10.1509/jmkr.48.2.238

Chevalier, J. A., & Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research, 43*(3), 345-354.

Crowley, A. E., & Hoyer, W. D. (1994). An Integrative Framework for Understanding Two-Sided Persuasion. *Journal of Consumer Research, 20*(4), 561-574.

Floyd, K., Freling, R., Alhoqail, S., Cho, H. Y., & Freling, T. (2014). How Online Product Reviews Affect Retail Sales: A Meta-analysis. *Journal of Retailing, 90*(2), 217-232. https://doi.org/10.1016/j.jretai.2014.04.004

Godes, D., & Mayzlin, D. (2004). Using Online Conversations to Study Word-of-Mouth Communication. *Marketing Science, 23*(4), 545-560. https://doi.org/10.1287/mksc.1040.0071

Goh, K.-Y., Heng, C. S., & Lin, Z. (2013). Social Media Brand Community and Consumer Behavior: Quantifying the Relative Impact of User- and Marketer-Generated Content. *Information Systems Research, 24*(1), 88-107. https://doi.org/10.1287/isre.1120.0469

He, S. X., & Bond, S. D. (2015). Why Is the Crowd Divided? Attribution for Dispersion in Online Word of Mouth. *Journal of Consumer Research, 41*(6), 1509-1527. https://doi.org/10.1086/680667

Helson, H. (1947). Adaptation-Level as Frame of Reference for Prediction of Psychophysical Data. *The American Journal of Psychology, 60*(1), 1-29. https://doi.org/10.2307/1417326

Liu, B., & Zhang, L. (2012). A Survey of Opinion Mining and Sentiment Analysis. In C. C. Aggarwal & C. Zhai (Eds.), *Mining Text Data* (pp. 415-463). Boston, MA: Springer US.

Liu, Y. (2006). Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue. Journal of Marketing, *70*(3), 74-89.

Mizerski, R. W., Golden, L. L., & Kernan, J. B. (1979). The Attribution Process in Consumer Decision Making. *Journal of Consumer Research*, *6*(2), 123-140.

Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Found. Trends Inf. Retr., 2*(1-2), 1-135. https://doi.org/10.1561/1500000011

Rosario, A. N. A. B., Sotgiu, F., De Valck, K., & Bijmolt, T. H. A. (2016). The Effect of Electronic Word of Mouth on Sales: A Meta-Analytic Review of Platform, Product, and Metric Factors. *Journal of Marketing Research (JMR), 53*(3), 297-229. https://doi.org/10.1509/jmr.14.0380

Tang, T., Fang, E., & Feng, W. (2014). Is Neutral Really Neutral? The Effects of Neutral User-Generated Content on Product Sales. *Journal of Marketing, 78*(4), 41-58.

Thelwall, M., Buckley, K., & Paltoglou, G. (2012). Sentiment strength detection for the social web. *J. Am. Soc. Inf. Sci. Technol., 63*(1), 163-173. https://doi.org/10.1002/asi.21662

Wang, X., Mai, F., & Chiang, R. H. L. (2014). Database Submission—Market Dynamics and User-Generated Content About Tablet Computers. *Marketing Science, 33*(3), 449-458. https://doi.org/10.1287/mksc.2013.0821

Ya, Y., Vadakkepatt, G. G., & Joshi, A. M. (2015). A Meta-Analysis of Electronic Word-of-Mouth Elasticity. *Journal of Marketing, 79*(2), 19-39. https://doi.org/10.1509/jm.14.0169