Are Neutral Sentiments Worth Considering When Investigating Online Consumer Reviews? Their Relationship with Review Ratings

Jaebong Son
California State University, Chico
json@csuchico.edu

Hyung-koo Lee
HEC Montreal
Hyung-koo.lee@hec.ca

Hyoungyong Choi
Hankuk University of Foreign Studies
hchoi@hufs.ac.kr

On-Ook Oh
University of Colorado at Denver
onook.oh@ucdenver.edu

Abstract

Online consumer reviews (OCRs) play a significant role for firms to understand consumer satisfaction. Prior research on OCRs has used sentiment analysis to identify and quantify consumers' subjective impressions in terms of positive and negative sentiments. However, OCRs also contain objective facts about the product or service, which are represented by neutral sentiments. In this study, we argue that it is important to distinguish neutral sentiments from those of positivity and negativity when investigating consumer satisfaction. Through a lens of expectation-confirmation theory, we delineate the roles of subjective information in relation to consumer satisfaction, in the sense that consumer satisfaction is mainly formed by one's subjective expectations and evaluations, not by objective facts of the product or service. The empirical results obtained from OCRs about hotels demonstrate that consumer satisfaction is significantly higher in positive OCRs than neutral ones, and significant lower in negative OCRs than neutral ones. Furthermore, neutral sentiments drastically improve the explanatory power of empirical models, thereby enhancing our understanding of consumer satisfaction. Academically, this study sheds light on the importance of neutral sentiments. Practically, neutral sentiments, when being separated from the other two sentiment categories, contribute to more accurately reflecting consumer satisfaction.

1. Introduction

Consumer satisfaction (or CS) is a crucial factor for the success of a firm [e.g., 1, 2]. CS is known to affect not only marketing performance [3], but also customer loyalty [4], which eventually result in firm performance and reputation [5, 6]. According to Rust and Zahorik [7], the benefits of CS also include minimizing failure costs and maximizing financial profitability. As businesses become more dynamic and consumer needs are ever-changing [8, 9], CS has been highlighted as a key performance metric that every firm should continually monitor and evaluate.

With the advancement of the Internet and information technology, firms can keep track of CS and gauge whether or not consumer needs are fulfilled [10]. In particular, online consumer reviews or OCRs are considered an important source of information, as consumers express their emotions, feelings, and experiences of the product or service by leveraging diverse information formats of textual contents, numerical review star ratings (also called as review ratings), photos and videos. [11, 12]. Hotel management and marketing studies have investigated the textual contents of OCRs by extracting factors that reflect CS that is represented by review ratings ranging from 1 (very unsatisfied) to 5 (very satisfied) [13-15]. *Sentiment analysis* is one of the popular methodologies to study how consumers' opinions, thoughts, and emotional valence are associated with CS [16, 17]. For example, Geetha, et al. [18] revealed that the positive sentiments of hotel guest reviews were positively associated with review ratings, meaning that positive hotel experiences increased CS. Similarly, Hu, et al. [19] found that the sentiments of book reviews were a strong predictor of book sales—the more positive sentiments, the higher book sales.

In fact, the purpose of sentiment analysis fits well with CS, as CS is formed by an individual's subjective experience and cognitive evaluation manifested after consuming the product or service [20, 21]. However, the sentiments of OCRs are not only positive or negative, but also neutral, because OCRs can convey factual information about the product or service [22, 23]. We found from relevant research that while both positive and negative sentiments are indicative of CS [18, 19], neutral sentiments are more associated with

URI: https://hdl.handle.net/10125/79902
978-0-9981331-5-7

the objectivity of the product or service [e.g., 24]. However, OCR research leveraging sentiment analysis has grouped neutral sentiments with either positive or negative sentiments [e.g., 18, 19], possibly ignoring the role of objective facts in regard to CS and thus weakening empirical evidence. This study attempts to fill the gap by answering the following research question:

RQ. Do neutral sentiments improve our understanding of consumer satisfaction over and above positive and negative sentiments?

In what follows, we review the literature on OCRs and CS, and then develop hypotheses to answer the research question. After presenting the research methodology and empirical results, we conclude by discussing the findings, limitations, and implications for future research.

2. Literature Reviews and Hypothesis Development

2.1. Review ratings and firm performance

OCRs, as a peer-generated product information, significantly impact on financial performance [6, 25]. Anderson [26] reported that review ratings positively affect the profitability of hotels, in that while maintaining the same occupancy, hotels can increase room price by 11.2 percent per 1-point increase in review ratings. The literature on online auctions revealed that as review ratings increase, price premium increases as well—a 1% decrease in review ratings reduces price premium by 0.11% [27]. Cui, et al. [28] examined the effect of OCRs on sales of new products. They collected OCRs available on Amazon.com about search products (e.g., consumer electronics) and experience products (e.g., video games). Their empirical results indicated that review ratings have a positive effect on the sales of products, and such effect is stronger for search goods than experience goods. In different e-commerce contexts of books, television shows and movies, studies consistently reported that review ratings are a significant factor forecasting revenue [e.g., 29, 30, 31].

2.2. Expectation-confirmation theory and review ratings

Expectation-confirmation theory (ECT) provides a solid foothold to consider review ratings as CS (i.e., consumer satisfaction). ECT posits that one's satisfaction is formed in comparison to his or her original expectations with perceived performance [32]. To be specific, when a product's or service's perceived performance outperforms (or underperforms) one's original expectations, his or her satisfaction increases (or decreases).

Review ratings have been used as a measure of the overall quality of the product or service [33], even though there exist criticisms that review ratings are limited in capturing such overall quality [34, 35]. Based on ECT, Engler, et al. [33] systematically analyzed review contents (e.g., "great" as a word for *experience*; "expect" as a word for *expectation*) and provided empirical evidence of the relationship between review ratings and CS. Their findings implied that review ratings are determined depending on both one's pre-purchase expectations and perceived performance after consumption. Therefore, we use review ratings as a proxy for CS—the higher review ratings, the more satisfaction.

2.3. Review sentiments and CS

Recent studies investigating OCRs have provided ample evidence of why the textual contents of OCRs are of importance for research [e.g., 36, 37]. For example, Pavlou and Dimoka [38]'s study on the price premiums of eBay sellers reported that seller-reputation cues obtained from feedback comments contribute to improving the coefficient of determination for price premiums by 20-30% (R^2=0.5) compared to that of prior studies (R^2=0.2-0.3).

Identifying factors contributing to CS through text analytics has also drawn a great attention from academia and industry [e.g., 15, 37, 39]. One promising analytics technique is sentiment analysis, which systematically identify, measure, and categorize emotional information into positivity, negativity, and neutrality [e.g., 40, 41]. There are three broad approaches to sentiment analysis: (1) a lexicon-based approach in which the sentiment of textual information is identified based on a predefined list of positive and negative words; (2) a machine learning approach in which sentiments are classified based on learning algorithms (e.g., decision trees, neutral networks); and (3) a hybrid approach that combines both of the lexicon-based and machine learning approaches [41]. In particular, as sentiment analysis is capable of characterizing information subjectivity as positive, negative, or neutral, it has been widely used to explore consumers' emotions, opinions, evaluations, and attitudes that are assumed to relate closely to CS [e.g., 17, 18, 40, 41].

2.4. Hypothesis development

ECT theorizes that a consumer's satisfaction is formed as a function of pre-purchase expectations in comparison with post-purchase perceived performance [32]. In other words, expectations as one's personal beliefs construct a frame of reference for perceived performance to be evaluated [e.g., 42]. It is important to note that both expectations and perceived performance represent *subjectivity* rather than *objectivity*. Oliver [32] strengthened the information subjectivity of consumers' expectations by asserting that a person's expectations are influenced by his or her prior experiences. Consequently, these arguments point out that a consumer's satisfaction does not indicate objective facts of the product or service, but reflect his or her own subjective beliefs, evaluations and opinions. Subjectivity refers to one's opinions about reality [43], so the subjectivity of OCRs mainly discusses consumers' emotions, perceptions and experiences of the product or service. Viewed in this light, subjectivity is indicative of *individuality* [44]. Objectivity, on the other hand, is a factual statement about reality [24], so the objectivity of OCRs mostly describes objective portrayal of the product or service. Thus, objectivity suggests *commonality* [e.g., 44].

Based on the above discussions, we summarize positive, neutral, and negative sentiments as follows. First, positive sentiments disclose consumers' individuality concerning how much consumers are *satisfied* with the product or service. Second, neutral sentiments indicate commonality (e.g., factual information) of the product or service rather than consumers' individuality. Last, negative sentiments reveal consumers' individuality by the extent to which consumers are *dissatisfied* with the product or service.

Table 1. Variable description

<table>
<tr><th>Variables</th><th>Explanation</th><th>Mean</th><th>S.D.</th><th>Range</th></tr>
<tr><td colspan="5">Dependent</td></tr>
<tr><td>Review Rating_i</td><td>Review i's star rating (or satisfaction)</td><td>3.86</td><td>1.26</td><td>1-5</td></tr>
<tr><td colspan="5">Main</td></tr>
<tr><td>Pos_Neu</td><td colspan="4" rowspan="2">Pos_Neu to compare reviews whose dominant sentiments are positive with those with the neutral sentiments; Neg_Neu to compare reviews whose dominant sentiments are negative with those with the neutral sentiments.
<table>
<tr><th></th><th colspan="3">Dominant sentiments of reviews</th></tr>
<tr><th>Dummy codes</th><th>Positive</th><th>Negative</th><th>Neutral</th></tr>
<tr><td>Pos_Neu</td><td>1</td><td>0</td><td>0</td></tr>
<tr><td>Neg_Neu</td><td>0</td><td>1</td><td>0</td></tr>
</table></td></tr>
<tr><td>Neg_Neu</td></tr>
<tr><td>PosNeu_Neg</td><td colspan="4">By considering neutral sentiments as positive, PosNeu_Neg to compare reviews whose dominant sentiments are positive (i.e., positive=positive with neutral) those with negative sentiments: 1 for reviews with positive sentiments; -1 for OCRs with negative sentiments.</td></tr>
<tr><td>Pos_NegNeu</td><td colspan="4">By considering neutral sentiments as negative, Pos_NegNeu to compare reviews whose dominant sentiments are positive those with negative sentiments (i.e., negative=negative with neutral): 1 for reviews with positive sentiments; -1 for reviews with negative sentiments.</td></tr>
<tr><td colspan="5">Control</td></tr>
<tr><td>Photos_i</td><td>The number of photos in review i</td><td>0.16</td><td>1.14</td><td>0-50</td></tr>
<tr><td>Length_i</td><td>The number of words in review i</td><td>128.2</td><td>131.6</td><td>12-2512</td></tr>
<tr><td>Five_FourThree</td><td colspan="4" rowspan="2">Five_FourThree and Four_Three to control hotel ratings' effect on Review Rating.
<table>
<tr><th></th><th colspan="3">Hotel rating</th></tr>
<tr><th>Contrast codes</th><th>5-star</th><th>4-star</th><th>3-star</th></tr>
<tr><td>Five_FourThree</td><td>2</td><td>-1</td><td>-1</td></tr>
<tr><td>Four_Three</td><td>0</td><td>1</td><td>-1</td></tr>
</table></td></tr>
<tr><td>Four_Three</td></tr>
</table>

Table 2. Reviews with sentiment scores

Example Reviews	Sentiment Proportion		
	Positive	Neutral	Negative
When I checked in they upgraded me to a Suite when was different than any other suite I have been in. The Living room have two chairs pointed toward a 60 inch tv. No walls except for the bathroom which had no door.	0.31	0.38	0.31
Hotel/casino is centralized on the strip. Everything is within walking distance or reachable by the monorail. Casino is just the right size. Food Court is good but a little overpriced. Buffet is good and adequately priced.	0.32	0.37	0.31
Room was clean, shower was great w/extra jets, only thing I noticed was there was coffee for coffee maker first day but none after? stayed there for 3 days. The resort fees and parking fees are not helpful but all the hotels have them.	0.33	0.35	0.32

These distinct characteristics of the three sentiments lead us to postulate that neutral sentiments would be significantly different from the other sentiments in assessing satisfaction. That is to say, consumer satisfaction conveyed in OCRs whose dominant sentiments are positive is higher than that in OCRs whose dominant sentiments are neutral. Likewise, consumer satisfaction held in OCRs whose dominant sentiments are negative is lower than that in OCRs whose dominant sentiments are neutral. We formulate the following two hypotheses concerning the three sentiment categories.

Hypothesis 1 (H1). OCRs with the sentiment category of positive have higher review ratings than OCRs with the sentiment category of neutral.

Hypothesis 2 (H2). OCRs with the sentiment category of negative have lower review ratings than OCRs with the sentiment group of neutral.

3. Research Methodology

To examine the above hypotheses, we collected guest reviews on three hotels whose hotel star ratings (hereafter '*hotel ratings*') range from 3 to 5 shown in Table 1. To minimize any possible geographical effects on guest satisfaction, we chose hotels in the same region of Las Vegas. These reviews were posted during years between 2015 and 2020 on TripAdvisor.com. We performed sentiment analysis on the collected OCRs using the Stanford CoreNLP Toolkit, a Java-based Natural Language Processing library (hereafter '*StanfordCoreNLP*') [45]. Built on the Stanford Sentiment Treebank and a Recursive Neural Tensor Network, the sentiment analyzer of *StanfordCoreNLP* is capable of more accurately categorizing user-generated content (e.g., OCRs) into binary sentiment categories (i.e., positive or negative) or multiple sentiment categories (i.e., positive, neutral, or negative) [22]. Table 2 displays a few actual OCRs' sentiment scores produced by *StanfordCoreNLP*.

Based on the sentiment scores, individual OCRs were grouped into positive-, neutral-, or negative-dominant sentiment categories. We then dummy-coded these categories by designating the neutral sentiment category as a baseline for comparison—*Pos_Neu* and *Neg_Neu*, each of which compares the positive or the negative sentiment category with the baseline category in association with review ratings, respectively. Review ratings as the dependent variable of this study are a 5-likert scaled CS (1=Terrible, 5=Excellent) [e.g., 46]. Along with three sentiment categories, we examined the binary sentiment categories by considering neutral as positive (*PosNeu_Neg*) and as negative (*Pos_NegNeu*) (see Table 1 for more details).

We controlled potentially significant effects on CS for better estimation. First, we included two review characteristics of *Length* and *Photos*. *Length* is a count of words per OCR, and *Photos* is the number of photos per OCR. Zhao, et al. [39] reported an interesting finding on review length and CS, in that hotel guests are inclined to post longer and more detailed reviews, as they are less satisfied. Namely, hotel guests tend to use more words to articulate their negative feelings and emotions (e.g., anger, frustration, displeasure) [47]. Hotel guests embed photos in their OCRs to more vividly express their experiences and evaluations, possibly reflecting the extent of their satisfaction [e.g., 48], as photos convey visual cues that textual review contents alone cannot communicate [e.g., 49]. As a result, *Length* and *Photos* were included in the empirical models of this study as OCR-related control variables. With these review characteristics, we added the following control variable about hotels. *Hotel Ratings* represent the overall quality of hotels—the most basic hotels at 1 star throughout 5 stars for the most luxurious hotels. Studies in tourism revealed that hotel guests have greater expectations for hotels with higher stars than for hotels with lower stars [50]. Furthermore, an empirical study by Rajaguru and Hassanli [51] evidenced the significant moderation effect of hotel ratings on the relationship between hotel financial performance and guest satisfaction. As our dataset includes 3-, 4-, and 5-star rated hotels, we devised the two contrast codes of *Five_FourThree* and *Four_Three*. *Five_FourThree* controls a difference in review ratings between a 5-star hotel and 4- and 3-star hotels. *Four_Three* rules out a difference in review ratings between a 4-star hotel and a 3-star hotel.

Using the aforementioned dependent, independent and control variables, we articulated four empirical models. First, model 1 (M1) is the baseline model that only consists of the control variables, such as *Photo*, *Length*, *Five_FourThree* and *Four_Three*. Second, model 2 (M2) adds M1 *PosNeu_Neg*. Third, model 3 (M3) adds M1 *Pos_NegNeu*. As the most comprehensive model, Model 4 (M4) adds M1 *Pos_Neu* and *Neg_Neu* to compare the three sentiment categories of positive, neutral, and negative. The below equation represents M4:

$$Review\_Rating_i = \beta_0 + \beta_1 Pos\_Neu + \beta_2 Neg\_Neu + \beta_3\ Photos_i + \beta_4 Length_i + \beta_5 Five\_FourThree + \beta_6 Four\_Three + \varepsilon_i$$

Table 3. Results of regression analyses

Models / Variables	DV=*Review Rating*			
	Model 1 (M1)	**Model 2 (M2)**	**Model 3 (M3)**	**Model 4 (M4)**
Main				
PosNeu_Neg	–	0.85689^{***} (0.01193)	–	–
Pos_NegNeu	–	–	0.54601^{***} (0.00509)	–
Pos_Neu	–	–	–	0.96709^{***} (0.02143)
Neg_Neu	–	–	–	-0.59829^{***} (0.02396)
Control				
$Photos_i$	0.04798^{***} (0.00487)	0.04550^{***} (0.00464)	0.04304^{***} (0.00464)	0.02608^{***} (0.00371)
$Length_i$	-0.00230^{***} (0.00006)	-0.00234^{***} (0.00006)	-0.00114^{***} (0.00005)	-0.00105^{***} (0.00005)
Five_FourThree	0.17824^{***} (0.00418)	0.16493^{***} (0.00395)	0.14373^{***} (0.00385)	0.12501^{***} (0.00347)
Four_Three	0.26681^{***} (0.00668)	0.24050^{***} (0.00632)	0.20972^{***} (0.00605)	0.19196^{***} (0.00561)
Constant	3.92920^{***} (0.00576)	3.16153^{***} (0.01208)	4.02888^{***} (0.00478)	3.42831^{***} (0.02090)
Model summary				
R^2	0.13660	0.23433	0.29878	0.41697
Adj. R^2	0.13652	0.23424	0.29870	0.41689
n	42457			

† Unstandardized coefficients and robust standard errors in parentheses are shown ($^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$).

The variance inflation factor (VIF) analysis on M4 indicated that multicollinearity is not a concern (Mean=1.799, Max=3.306) [52]. The studentized Breusch-Pagan (BP) test on M4 demonstrated the existence of heteroskedasticity, so we employed OLS with robust standard errors to estimate the empirical models [53, 54]. The statistical results of the models are shown in Table 3.

4. Results

From M1, M2, and M3, we observed that the binary sentiment categories of positive and negative improved the explanatory powers of M2 and M3 over M1—the R^2 of M1 is 0.136, while those of M2 and M3 are 0.234 and 0.298, respectively. In other words, the explanatory power of M2 is improved by 0.097 over that of M1, and this difference in R^2 was made only by *PosNeu_Neg*, which treated neutral as positive sentiments. Similarly, the different in R^2 between M1 and M3 was 0.162, and such different was achieved by considering neutral as negative sentiments (*Pos_NeuNeg*). Models M2 and M3 clearly demonstrate an importance of leveraging review sentiments in explaining review ratings.

Unlike M2 and M3, M4 includes neutral sentiments as a separate category, resulting in the highest R^2 of 0.416. To be exact, the R^2 of M4 is higher by 0.182 than that of M2 and by 0.118 than that of M. The improved explanatory power of M4 validates the importance of separating neutral sentiments from the other two sentiments. Therefore, we leverage the results of M4 to evaluate the hypotheses.

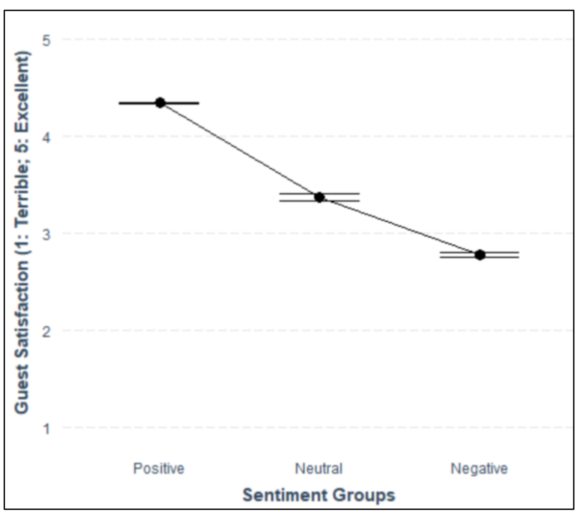

Figure 1. Guest satisfaction by the positive, neutral, and negative sentiment groups

It is notable to discuss the control variables, as they were expected to have significant effects on review ratings [39, 49]. We found that including photos is positively associated with review ratings. An additional photo increases review ratings by 0.026 (β_{Photos}=0.02608***) holding the other variables of the model constant. Similar to what Zhao, et al. [39] reported, review length is found to be negatively associated with review ratings—the longer reviews, the lower review ratings (β_{Length}=-0.00105***). Ten additional words decrease review ratings by 0.0105. We also discovered significant empirical evidence on the relationship between hotel ratings and review ratings. A 5-star hotel has a higher review rating by 0.375 on average than 4- and 3-star hotels ($\beta_{Five\_FourThree}$=0.12501***). A 4-star hotel has a higher review rating by 0.384 than that of a 3-star hotel (β_{Four_Three}=0.19196***). These differences in review ratings are well aligned with what the previous studies on OCRs reported—guests are more satisfied with higher star hotels than lower star hotels.

We now evaluate hypotheses H1 and H2. It turns out that H1 is supported, in the sense that when controlling for the review and hotel characteristics, the positive sentiment category has higher review ratings by 0.967 on average than the neutral sentiment category (β_{Pos_Neu}=0.967***, $F_{1, 42450}$=2036, Positive=4.39 vs. Neutral=3.42). We also found a significant difference in review ratings between the neutral and the negative sentiment categories while holding the other variables of M4 constant (β_{Neg_Neu}=-0.598***, $F_{1, 42450}$=623.25, Neutral=3.42 vs. Negative=2.83). Therefore, H2 is also supported. Figure 1 visually represents each sentiment group's average review rating.

4.1. Post-hoc analysis

The results of M4 showed that neutral sentiments are significantly associated with review ratings (therefore guest satisfaction). We furnish additional evidence by comparing M4 with both M2 and M3 in terms of how close the predicted review ratings would be to the actual review ratings. We use the root-mean-square error or RMSE for such comparison—the lower RMSE, the more accurate prediction. As a goodness-of-fit measure, RMSE is generally used to evaluate statistical and machine learning models [e.g., 55, 56].

To statistically compare M4's RMSE with those of M2 and M3, we leveraged the following steps: (1) randomize OCRs; (2) select randomized OCRs by a random percentage between 10 percentage as the minimum number of OCRs and 90 percentage as the maximum number—i.e., from 4,245 OCRs (10%) to 38,211 (90%); (3) examine M2, M3, and M4 to produce each model's RMSE; (4) repeat (1) to (3) 1000 times. As a result, we obtained each model's 1000 RMSEs calculated from the varying numbers of randomly chosen OCRs. Then, we created two dummy codes to compare M4 with M2 (*M2_M4*: M2=1, M3=0, M4=0) and M4 with M3 (*M3_M4*: M2=0, M3=1, M4=0). The below equation is an empirical model consisting of *RMSE* as the dependent variable, *M2_M4* and *M3_M4* as the main independent variables, and the number of guest reviews (*Obs*) as a control variable.

$$\begin{aligned} RMSE_i = \beta_0 + \beta_1 M2\_M4 + \beta_2 M3\_M4 \\ + \beta_3\, Obs_i \\ + \varepsilon_i \end{aligned}$$

We summarized the empirical result of the above equation in Table 4. It turned out that M4 is more accurate in predicting review ratings than M2 and M3. In other words, while controlling for the number of OCRs, M4's predicted review ratings are significantly closer to the actual review ratings than M2's by 0.14 in RMSE on average (β_{M2_M4}=0.140, M2=1.099 vs. M4=0.959) and than M3's by 0.0927 in RMSE on average (β_{M3_M4}=0.0927, M3=1.052 vs. M4=0.959).

Table 4. Post-hoc analysis result

	DV=*RMSE*
Main	
M2_M4	0.140*** (0.000188)
M3_M4	0.0927*** (0.000188)
Control	
Obs_i	1.13e-08 (8.88e-09)
Constant	0.959*** (0.000133)
Model summary	
R^2	0.9948
n	3000

† Standard errors in parentheses (*p<0.05, **p<0.01, ***p<0.001)

5. Discussion

In this study, we investigated review sentiments in relation to consumer satisfaction or CS by separating neutral sentiments from both positive and negative sentiments. We posited that neutral sentiments' primary information characteristic (i.e., objectivity or commonality) is significantly different from that of positive and negative sentiments (i.e., subjectivity or individuality). The empirical results supported our conjecture, in that when being separated from positive and negative sentiments, neutral sentiments contribute to enhancing our understanding of CS. Furthermore, the post-hoc analysis strengthened the empirical results—the three categories of sentiments (i.e., M4: positive, neutral, and negative) predict CS more accurately than the two categories of sentiments (i.e., M2: positive with neutral, negative; M3: positive, negative with neutral). One plausible explanation would be that the two different information characteristics of subjectivity and objectivity less interfered each other's influence on CS.

The findings reported in this study open opportunities for future research. First, instead of defining each review's overall sentiments as either positive, neutral, or negative, one may perform sentiment analysis centering on the product's or service's aspects (or features) to have a more nuanced comprehension of emotions and feelings. For example, the following sentence shown in Table 2, '… *The Living room have two chairs pointed toward a 60 inch tv.* …,' describes 'living room,' 'chairs,' and '60 inch tv' without positive or negative sentiments. They describe a hotel room's commonality and thus convey objective facts, rarely contributing to gauging guest satisfaction. An aspect- or feature-centering sentiment analysis could better articulate what factors influence CS and what factors do not. Second, review helpfulness as a function of neutral sentiments is of interest, as it indicates how helpful consumers' personal opinions and experiences are for potential customers [36]. In fact, reading OCRs means learning peer-consumers' personal thoughts, experiences and evaluations. However, neutral sentiments are about factual, objective information of the product or service. Therefore, future research may further investigate neutral sentiments in association with review helpfulness. Third, we are aware that many studies use a sentiment score ranging from 0 (negative) to 1 (positive). However, it is uncertain that 0.5 truly means neutral sentiments, even with lower and upper threshold points (e.g., does a sentiment score between 0.4-0.6 mean neutrality?). Instead, we categorized OCRs into positive-, neutral-, or negative-dominant group based on each sentiment score. Of course, we admit that converting continuous variables into categorical variables causes some degrees of information loss. Last but not least, while this study analyzed OCRs on hotels, future studies on OCRs of diverse business contexts (e.g., restaurants, health services, online auctions, etc.) will strengthen the generalizability of the current findings.

6. Conclusions

Online consumer reviews have been deemed an important information source for companies to understand consumer satisfaction [e.g., 57]. Rich evidence demonstrates the positive relationship between consumer satisfaction and firm performance [e.g., 27, 28]. We discussed based on ECT why the information characteristic of consumer satisfaction is closer to subjective than objective and thus why neutral sentiments are different from positive and negative sentiments. Finally, we performed regressions of review ratings as consumer satisfaction on the sentiment categories of positive, neutral, and negative. Founded on the empirical results, we conclude that neutral sentiments are an important sentiment category that must be distinctly included in empirical models to study consumer satisfaction.

This study contributes to academia as well as practitioners. First of all, we expand the applicability of ECT to delineating the relationship between consumer satisfaction and review sentiments. Based on the central tenets of ECT (i.e., expectation, satisfaction), this study empirically showed why neutral sentiments (i.e., objectivity or commonality) are different from positive and negative sentiments, each of which mainly conveys one's subjectivity or individuality. For the existing literature on e-commerce and marketing, we shed light on the importance of contemplating neutral sentiments which are not significantly tackled yet but could bring meaningful implications for consumer satisfaction. In addition, industry practitioners take advantage of this research. The current findings imply that the relationship between review sentiments and consumer satisfaction is distorted, when neutral sentiments are grouped into either the positive or the negative sentiment category. Therefore, hoteliers and hotel operators, for example, may pay more attention to the three sentiments of positive, neutral, and negative, instead of positive and negative, in order to better understand guest satisfaction. Hotel booking agencies can improve their sentiment analysis practices by reflecting the current findings of this study.

7. References

[1] T. Y. Choi and R. Chu, "Determinants of hotel guests' satisfaction and repeat patronage in the Hong Kong hotel industry," International Journal of Hospitality Management, vol. 20, no. 3, pp. 277-297, 2001.

[2] G. K. Kanji and W. Wallace, "Business excellence through customer satisfaction," Total quality management, vol. 11, no. 7, pp. 979-998, 2000.

[3] L. P. Pleshko and J. J. Cronin Jr, "An investigation into the relationship between consumer satisfaction, marketing performance, and firm performance in a services environment," Journal of Customer Service in Marketing & Management, vol. 3, no. 1, pp. 7-17, 1997.

[4] J. L. Heskett, T. O. Jones, G. W. Loveman, W. E. Sasser, and L. A. Schlesinger, "Putting the service-profit chain to work," Harvard business review, vol. 72, no. 2, pp. 164-174, 1994.

[5] S. Helm, I. Garnefeld, and J. Tolsdorf, "Perceived corporate reputation and consumer satisfaction–an experimental exploration of causal relationships," Australasian Marketing Journal (AMJ), vol. 17, no. 2, pp. 69-74, 2009.

[6] W. G. Kim, H. Lim, and R. A. Brymer, "The effectiveness of managing social media on hotel performance," International Journal of Hospitality Management, vol. 44, pp. 165-171, 2015.

[7] R. T. Rust and A. J. Zahorik, "Customer satisfaction, customer retention, and market share," Journal of retailing, vol. 69, no. 2, pp. 193-215, 1993.

[8] T.-P. Liang and M. Tanniru, "Customer-centric information systems," Journal of Management Information Systems, vol. 23, no. 3, pp. 9-15, 2006.

[9] J. E. Riddle, "Reengineering Ryder to meet rising customer expectations," National Productivity Review, vol. 15, no. 1, pp. 51-62, 1995.

[10] S. M. Mudambi and D. Schuff, "Research note: What makes a helpful online review? A study of customer reviews on Amazon. com," MIS quarterly, pp. 185-200, 2010.

[11] M. Siering and C. Janze, "Information processing on online review platforms," Journal of Management Information Systems, vol. 36, no. 4, pp. 1347-1377, 2019.

[12] J. Son, A. Negahban, Y. Lee, J. Connolly, and D. Chiang, "When More is More and Less is More: Depth and Breadth of Product Reviews and Their Effects on Review Helpfulness," in Proceedings of the 53rd Hawaii International Conference on System Sciences, 2020.

[13] M. Schuckert, X. Liu, and R. Law, "A segmentation of online reviews by language groups: How English and non-English speakers rate hotels differently," International Journal of Hospitality Management, vol. 48, pp. 143-149, 2015.

[14] S. Moon, P. K. Bergey, and D. Iacobucci, "Dynamic effects among movie ratings, movie revenues, and viewer satisfaction," Journal of marketing, vol. 74, no. 1, pp. 108-121, 2010.

[15] Z. Xiang, Z. Schwartz, J. H. Gerdes Jr, and M. Uysal, "What can big data and text analytics tell us about hotel guest experience and satisfaction?," International Journal of Hospitality Management, vol. 44, pp. 120-130, 2015.

[16] A. Giachanou and F. Crestani, "Like it or not: A survey of twitter sentiment analysis methods," ACM Computing Surveys (CSUR), vol. 49, no. 2, pp. 1-41, 2016.

[17] S. Schmunk, W. Höpken, M. Fuchs, and M. Lexhagen, "Sentiment analysis: Extracting decision-relevant

knowledge from UGC," in Information and Communication Technologies in Tourism 2014: Springer, 2013, pp. 253-265.
[18] M. Geetha, P. Singha, and S. Sinha, "Relationship between customer sentiment and online customer ratings for hotels - An empirical analysis," Tourism Management, vol. 61, pp. 43-54, 2017/08/01/ 2017, doi: https://doi.org/10.1016/j.tourman.2016.12.022.
[19] N. Hu, N. S. Koh, and S. K. Reddy, "Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales," Decision Support Systems, vol. 57, pp. 42-53, 2014/01/01/ 2014, doi: https://doi.org/10.1016/j.dss.2013.07.009.
[20] D. M. Phillips and H. Baumgartner, "The role of consumption emotions in the satisfaction response," Journal of Consumer psychology, vol. 12, no. 3, pp. 243-252, 2002.
[21] D. K. Tse, F. M. Nicosia, and P. C. Wilton, "Consumer satisfaction as a process," Psychology & Marketing, vol. 7, no. 3, pp. 177-193, 1990.
[22] R. Socher et al., "Recursive deep models for semantic compositionality over a sentiment treebank," in Proceedings of the 2013 conference on empirical methods in natural language processing, 2013, pp. 1631-1642.
[23] M. Koppel and J. Schler, "THE IMPORTANCE OF NEUTRAL EXAMPLES FOR LEARNING SENTIMENT," Computational Intelligence, vol. 22, no. 2, pp. 100-109, 2006, doi: https://doi.org/10.1111/j.1467-8640.2006.00276.x.
[24] S. Boudana, "A definition of journalistic objectivity as a performance," Media, Culture & Society, vol. 33, no. 3, pp. 385-398, 2011.
[25] Y. Chen and J. Xie, "Third-party product review and firm marketing strategy," Marketing science, vol. 24, no. 2, pp. 218-240, 2005.
[26] C. Anderson, "The impact of social media on lodging performance," 2012.
[27] D. Lucking-Reiley, D. Bryan, N. Prasad, and D. Reeves, "Pennies from eBay: The determinants of price in online auctions," The journal of industrial economics, vol. 55, no. 2, pp. 223-233, 2007.
[28] G. Cui, H.-K. Lui, and X. Guo, "The effect of online consumer reviews on new product sales," International Journal of Electronic Commerce, vol. 17, no. 1, pp. 39-58, 2012.
[29] C. Dellarocas, X. M. Zhang, and N. F. Awad, "Exploring the value of online product reviews in forecasting sales: The case of motion pictures," Journal of Interactive marketing, vol. 21, no. 4, pp. 23-45, 2007.
[30] D. Godes and D. Mayzlin, "Using online conversations to study word-of-mouth communication," Marketing science, vol. 23, no. 4, pp. 545-560, 2004.
[31] Y.-F. Chen, "Herd behavior in purchasing books online," Computers in Human Behavior, vol. 24, no. 5, pp. 1977-1992, 2008.
[32] R. L. Oliver, "A cognitive model of the antecedents and consequences of satisfaction decisions," Journal of marketing research, vol. 17, no. 4, pp. 460-469, 1980.
[33] T. H. Engler, P. Winter, and M. Schulz, "Understanding online product ratings: A customer satisfaction model," Journal of Retailing and Consumer Services, vol. 27, pp. 113-120, 2015.
[34] N. Hu, L. Liu, and J. J. Zhang, "Do online reviews affect product sales? The role of reviewer characteristics and temporal effects," Information Technology and management, vol. 9, no. 3, pp. 201-214, 2008.
[35] N. S. Koh, N. Hu, and E. K. Clemons, "Do online reviews reflect a product’s true perceived quality? An investigation of online movie reviews across cultures," Electronic Commerce Research and Applications, vol. 9, no. 5, pp. 374-385, 2010.
[36] J. Son, D. Kim, and C. Koh, "Effectiveness of Online Consumer Product Review: The Role of Experiential Information," in Proceedings of the 54th Hawaii International Conference on System Sciences, 2021, p. 4323.
[37] H. Chen, R. H. Chiang, and V. C. Storey, "Business intelligence and analytics: From big data to big impact," MIS quarterly, pp. 1165-1188, 2012.
[38] P. A. Pavlou and A. Dimoka, "The nature and role of feedback text comments in online marketplaces: Implications for trust building, price premiums, and seller differentiation," Information Systems Research, vol. 17, no. 4, pp. 392-414, 2006.
[39] Y. Zhao, X. Xu, and M. Wang, "Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews," International Journal of Hospitality Management, vol. 76, pp. 111-121, 2019.
[40] B. Pang and L. Lee, "Opinion Mining and Sentiment Analysis," Found. Trends Inf. Retr., vol. 2, no. 1–2, pp. 1–135, 2008, doi: 10.1561/1500000011.
[41] B. Liu, "Sentiment analysis and opinion mining," Synthesis lectures on human language technologies, vol. 5, no. 1, pp. 1-167, 2012.
[42] M. Fishbein and I. Ajzen, "Belief, attitude, intention, and behavior: An introduction to theory and research," 1977.
[43] F. González Rey, "The topic of subjectivity in psychology: Contradictions, paths and new alternatives," Journal for the Theory of Social Behaviour, vol. 47, no. 4, pp. 502-521, 2017.
[44] W. Wang, L. He, Y. J. Wu, and M. Goh, "Signaling persuasion in crowdfunding entrepreneurial narratives: the subjectivity vs objectivity debate," Computers in Human Behavior, vol. 114, p. 106576, 2021.
[45] C. D. Manning, M. Surdeanu, J. Bauer, J. R. Finkel, S. Bethard, and D. McClosky, "The Stanford CoreNLP natural language processing toolkit," in Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations, 2014, pp. 55-60.
[46] X. Wang, J. Sun, and H. Wen, "Tourism seasonality, online user rating and hotel price: A quantitative approach based on the hedonic price model," International Journal of Hospitality Management, vol. 79, pp. 140-147, 2019.
[47] T. Verhagen, A. Nauta, and F. Feldberg, "Negative online word-of-mouth: Behavioral indicator or

emotional release?," Computers in Human Behavior, vol. 29, no. 4, pp. 1430-1440, 2013.

[48] I. S. Lo, B. McKercher, A. Lo, C. Cheung, and R. Law, "Tourism and online photography," Tourism management, vol. 32, no. 4, pp. 725-731, 2011.

[49] Y. Ma, Z. Xiang, Q. Du, and W. Fan, "Effects of user-provided photos on hotel review helpfulness: An analytical approach with deep leaning," International Journal of Hospitality Management, vol. 71, pp. 120-131, 2018.

[50] T. Radojevic, N. Stanisic, and N. Stanic, "Ensuring positive feedback: Factors that influence customer satisfaction in the contemporary hospitality industry," Tourism Management, vol. 51, pp. 13-21, 2015.

[51] R. Rajaguru and N. Hassanli, "The role of trip purpose and hotel star rating on guests' satisfaction and WOM," International Journal of Contemporary Hospitality Management, 2018.

[52] L. S. Aiken, S. G. West, and R. R. Reno, Multiple regression: Testing and interpreting interactions. sage, 1991.

[53] H. White, "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity," Econometrica: journal of the Econometric Society, pp. 817-838, 1980.

[54] T. S. Breusch and A. R. Pagan, "A simple test for heteroscedasticity and random coefficient variation," Econometrica: Journal of the econometric society, pp. 1287-1294, 1979.

[55] J. Han, M. Kamber, and J. Pei, "Data mining concepts and techniques third edition," The Morgan Kaufmann Series in Data Management Systems, vol. 5, no. 4, pp. 83-124, 2011.

[56] S. Heravi, D. R. Osborn, and C. Birchenhall, "Linear versus neural network forecasts for European industrial production series," International Journal of Forecasting, vol. 20, no. 3, pp. 435-446, 2004.

[57] D. Kaemingk. "Online reviews statistics to know in 2021." Qualtrics. https://www.qualtrics.com/blog/online-review-stats/ (accessed June 11, 2021).