On the Usefulness of Online Review Valance

Mehmet Akgul
McMaster University
akgulm@mcmaster.ca

Ali Reza Montazemi
McMaster University
montazem@mcmaster.ca

Hamed Qahri-Saremi
Colorado State University
hamed.qahri-saremi@colostate.edu

Abstract

Online reviews are posted by consumers to inform others about their post-usage attitude towards the focal product/service. Often two-sided reviews that provide both pros/cons of a product/service are considered more informative than one-sided reviews. While research has looked into the usefulness of positive versus the negative aspects of the two-sided reviews, the findings are inconclusive. Some studies find negative aspects of the two-sided reviews to be more useful than positive aspects, some find the reverse to be true, and yet there are research findings that show both positive and negative aspects are equally useful. As a result, online review platforms are at loss to deal with the effects of positive/negative aspects of the reviews. Drawing on the Evaluative Space Model, our empirical study of 4705 restaurant reviews from TripAdvisor show that usefulness of reviews depends on how the attitudes of the receivers of the reviews are tilted towards positive or negative aspects of the focal product/service. And that the relationship between review usefulness and reviewers' attitude is nonlinear.

Keywords: Online Review Helpfulness, Evaluative Space Model, Negativity Bias, Positivity Offset, Text Analytics

1. Introduction

Online review platforms attract millions of consumers. For example, Yelp reached 265 million reviews for a variety of businesses (Yelp, 2023), and Facebook has around 2.96 billion monthly active users as of December 31, 2022 (Meta, 2023). Consumers increasingly rely on informal information sources for making decisions about products/services: 85% of consumers usually read up to ten reviews before making their purchase decision (Picher Vera et al., 2016). Online reviews (also called electronic word of mouth (eWoM)) refer to "any positive or negative statement made by potential, actual, or former customers about a product or company, which is made available to a multitude of people and institutions via the Internet" (Hennig-Thurau et al., 2004, p39). Because of its impact on consumer decision-making, 64% of marketers consider online reviews more effective than traditional marketing tools (Whitler, 2014).

Although consumers can benefit from a vast number of reviews available on consumer review platforms, this also brings additional challenges and costs. It is not only time-consuming to find and read the reviews, but the amount of information also makes it difficult for the consumer to process and judge reviews as a result of the information load. Information load refers to "a complex mixture of the quantity, ambiguity and variety of information that people are forced to process. As load increases, people take increasingly strong steps to manage it" (Weick, 1995, p. 87).

The amount of information makes it difficult for the consumer to process and judge reviews (Malhotra, 1984). Therefore, to mitigate the information load, it is recommended that review platforms provide more useful information to consumers (Wang et al., 2020) by providing more helpful reviews, as review helpfulness is a measure of perceived value for the consumer (Mudambi & Schuff, 2010). To that end, there has been extensive research in identifying factors that make a review helpful (Qahri-Saremi & Montazemi, 2019), and one such factor is the review star rating.

Star rating shows the reviewer's overall attitude of a product/service, where 1 star represents an extremely negative experience, and 5 stars represent an extremely positive experience of the product/service (Mudambi & Schuff, 2010). Extant literature has used star rating to figure out the usefulness of reviews' positive/negative attitude for the consumers (Hong et al., 2017). Notwithstanding its importance, the findings about the relationship between the star rating and review helpfulness are mixed: while some studies found that there is a positive relationship (e.g., Chatterjee, 2020), i.e., positive reviews are found to be more helpful; some found a negative relationship (e.g., Chua & Banerjee, 2015), i.e., negative reviews are found to be more helpful. On the other hand, some other studies found no relationship between review

URI: https://hdl.handle.net/10125/106417
978-0-9981331-7-1

star rating and review helpfulness (e.g., Hong et al., 2017).

Our contention in this paper is that the reason behind such discrepant findings is the bipolar nature of the star rating: The star-rating measurement of consumers' attitudes is deficient because it cannot distinguish between indifferent consumers and those who are ambivalent. Indifference means that the consumer has no particular preference for or against the product. Ambivalence refers to "a state in which individual experiences both high positive and negative reactions to an attitudinal object" (Yang & Unnava, 2016, p. 332). Ambivalent consumers can be enticed towards products/services usage (Hamby & Russel, 2020), while indifferent consumers cannot be enticed (Solomon, 2008). To be able to distinguish between ambivalence and indifference, we need to understand the separate effects of reviewers' positive and negative attitudes on the usefulness of reviews. Given the bipolar nature of star rating, its use in the assessment of review usefulness is questionable.

As such, the usefulness of reviews depends on how the reviewers' attitude is tilted towards positive or negative aspects of the focal product/service. In this paper, we draw on the Evaluative Space Model (ESM) to study the effects of reviewers' positive and negative attitudes, as reflected in their reviews of the focal product/service (i.e., text positivity and text negativity), on the helpfulness of the reviews. Drawing on ESM, we show the differential effects of positive and negative attitudes on review helpfulness and that the relationship between review usefulness and reviewers' positive/negative attitudes is nonlinear. To that end, in this exploratory research, we address the following research question:

RQ: How do review text positivity and review text negativity affect its helpfulness?

2. Theoretical Foundation

2.1. Information Flow and Attitude Measurement

In an online review, the reviewer conveys a message (the review) to the consumers. As the communication between the sender (the reviewer) and the receiver (the consumer) happens, the receiver judges the message and then makes inferences from that message. Tang and Guo (2015, p. 69) contend that "eWoM communication begins when an eWoM sender develops attitudes toward a product/service based on their consumption experience(s)." They then convert (encode) their attitudes into a textual review of the product/service and an assigned star rating posted on consumer review platforms (e.g., TripAdvisor, Yelp). The receiver reads the message and then develops an attitude (i.e., positive, negative, or indifference) towards the product/service and the message itself.

2.2. Attitude Measurement

Attitude is defined as "a psychological tendency that is expressed by evaluating a particular entity with some degree of favour or disfavour" (Eagly & Chaiken 1993, p. 1). In online communication, review writers convey their attitudes toward a product/service within the review text as well as their overall attitude by associating a star rating with their review. On the receiver side, review readers show their attitude towards the review by voting for its helpfulness. For example, readers can press the "Helpful" button under the reviews on Amazon or use the thumbs-up emoji on TripAdvisor to show that they found the review helpful.

2.3. Review Star

Generally, it is assumed that the reviewer rating (e.g., star rating) of the reviews conveys the reviewers' attitude portrayed in the message/review. However, star rating, being a bipolar measure, is unable to correctly capture the overall attitude of the reviewer. Such traditional bipolar measures of attitude fail to distinguish between ambivalence and indifference (Van Harreveld et al., 2015): that is, people with strong opposite (positive and negative) attitudes and people who are indifferent (Schneider & Schwarz, 2017) both select the midpoint of the bipolar scale (Klopfer & Madden, 1980) to show their attitude towards the product/service.

2.4. Evaluative Space Model

A bipolar scale (e.g., star rating) measures the two poles of a bipolar concept, such as "satisfied" and "dissatisfied" (DeCastellarnau, 2018). A unipolar scale, on the other hand, uses only one pole of the concept to measure the extremity, such as "significant" and "not at all" (DeCastellarnau, 2018).

The use of a bipolar scale in the measurement of attitudes has resulted in questions about their sufficiency (Cacioppo et al., 1997), as the lack of distinction between unipolar and bipolar attributes has resulted in "misinterpretations of the empirical findings" (Rossiter 2010, p. 105). For example, the star rating of a review, which is believed to measure a reviewer's attitude, is unable to identify indifference and ambivalence: individuals who are indifferent and those who are ambivalent are both inclined to give a 3-star rating to a product/service (Klopfer & Madden, 1980).

The bipolar conceptualization of attitudes means that the underlying positive and negative attitudes are reciprocal: that is, a certain degree of increase in negativity (positivity) results in an equal degree of decrease in positivity (negativity). However, drawing on the ESM, considering an endorsement of positivity (negativity) equal to a rejection of negativity (positivity) is misleading. Extant studies suggest that a two-dimensional representation of positive and negative attitudes better than the traditional bipolar model (Cacioppo et al., 1997; Walden et al., 2005). A person's overall attitude can vary depending on the balance between their positive and negative attitudes (Audrezet & Parguel, 2018). If someone feels mostly negative and very few positive ones, they will have a very negative attitude. On the other hand, if someone feels mostly positive and very few negative ones, they will have a very positive attitude. It is also possible for someone to have low levels of both positive and negative feelings, resulting in an indifferent attitude. However, if someone experiences high levels of both positive and negative feelings, their attitude is ambivalent.

According to the ESM, an attitude is formulated as (Cacioppo et al., 1997):

$$\text{Overall Attitude} = P*f(i) + c - N*g(j) + I(ij) + e \quad (1)$$

where P and N are weighting coefficients; i represents the level of positivity activated by an attitude object; $f(i) = i^k$ $(k < 1)$ is the activation (transfer) function for positivity; k is the exponent in the positivity activation function; $c > 0$ is a constant representing the positivity offset; j represents the level of negativity activated by an attitude object; $g(j) = j^m$ $(m < 1)$ is the activation (transfer) function for negativity; m is the exponent in the negativity activation function; I(ij) represents non-additive (interaction) effects; and e represents the error term.

Our focus in this paper is on the receiver's attitude toward the usefulness of the review written by the reviewers. To that end, *i* represents the positivity of the reviewer's attitude toward the focal product/service, manifested in the positive sentiment of the review; *j* represents the negativity of the reviewer's attitude toward the focal product/service, manifested in the negative sentiment of the review. Here, consistent with prior studies (Malik & Hussain, 2017), we assume that the receiver's positive and negative attitude of focal product/service is a function of the positive and negative sentiments of the review. As such, based on ESM, $P*f(i) + c$ represents receiver's positive attitude toward the focal product/service and $N*g(j)$ represents receiver's negative attitude toward the focal product/service.

A sample plot representing Equation (1) is depicted in Figure 1 that shows a three-dimensional surface representation of an attitude. The attitude is modelled as Attitude = $0.4\ \text{Pos}^{0.5} + 1 - 0.6\ \text{Neg}^{0.5}$, where $P = 0.4$, $N = 0.6$, $k = m = 0.5$ and $c = 1$. The x-y plane (i.e., the evaluative plane) represents the level of positivity and negativity activated by an attitude object. The surface shows the possible resulting attitudes towards that attitude object.

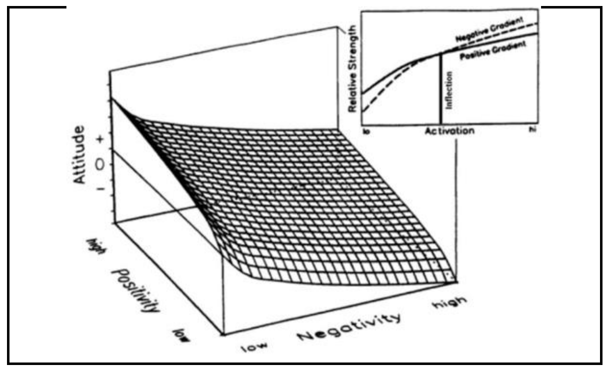

Figure 1. A sample plot of attitude represented by Equation (1) (adopted from Cacioppo et al., 1997)

Although the ESM model provides a formula for measuring attitude as a function of positive and negative attitudes, it doesn't say much about the parameters, P, N, I, k, m and c, of the positive and negative attitudes. Individuals tend to respond differently to positive and negative information, and this asymmetry can be incorporated into the ESM model by using the positivity offset and negativity bias as follows.

2.5. Positivity Offset and Negativity Bias

Positivity Offset refers to the idea that "when people have no, neutral or non-diagnostic information about something (or otherwise lack evidence for positive reactions), their default attitudes tend to be slightly favorable" (Snyder & Tormala, 2017, p. 556). For example, when people have a lack of knowledge about individuals, they perceive them as favorable (Sears, 1983).

Positivity offset is characterized by the constant term c in the ESM model (Cacioppo et al., 1997). In the case of no positive and negative information (i.e., $i = 0$ and $j = 0$), f(i), g(j), and I(ij) are zero, and attitude is equal to $c > 0$. This is depicted in Figure 1 (figure on the top right): at a low activation point, the positive attitude is higher than the negative attitude due to the positivity offset being greater than zero.

Negativity bias refers to "the fact that all else equal, negative information generally carries more psychological weight or has a greater impact on attitudes than positive information (Snyder & Tormala, 2017, p. 556)". This is also supported by the prospect theory, which states that the impact of a loss is more than the impact of a gain with the same amount (Kahneman & Tversky, 1979).

Cacioppo et al. (1997) contend that the negativity bias is characterized in one of two ways in the ESM model: (a) the weight of negativity, N, is greater than the weight of positivity, P (i.e., N > P) or (b) the negativity activation function is steeper than the activation function of positivity (i.e., m > k). This is illustrated in Figure 1 (figure on the top right): the negative attitude is steeper than the positive attitude.

The effect of positive and negative attitudes on the overall attitude is also shown in Figure 1. From Figure 1 (figure on the top right): it is observed that up to the inflection point, the overall attitude is dominated by the positive attitude towards the object. At the inflection point, the effect of positive and negative attitude is the same. However, after the inflection point, the negative attitude dominates the overall attitude.

2.6. Review Helpfulness

Consumers reading an online review show their attitude towards the review by providing a "helpfulness" rating to the associated review. For example, readers can press the "Helpful" button under the reviews on Amazon or use the thumbs-up emoji on TripAdvisor to show that they found the review helpful. Review helpfulness refers to "the degree to which consumers perceive a product review to be helpful in their own purchasing decision-making" (Lopez & Garza, 2022). Thus, an online review's perceived helpfulness can be interpreted as the online review's perceived value for the consumer (Mudambi & Schuff, 2010).

Review star rating shows the reviewer's overall attitude of a product/service, and it is one of the most used variables in the predicting the review helpfulness. In exploring the factors affecting review helpfulness, studies generally used a more quantifiable measure such as star rating (Chen, 2016) due to the cost of assessing the information content in a review text (Godes & Mayzlin, 2004). However, as stated previously, extant studies investigating the relationship between the review star rating and review helpfulness found inconclusive results.

The ESM shows that an overall attitude is a function of positive and negative attitudes. Hence, positive and negative attitudes should be considered separately, and that positivity offset and negativity bias exist in attitude formation.

3. Methodology

We used secondary data to explore the effects of review text positivity and review text negativity on review helpfulness. For the analysis, we focus on online reviews within the context of experience services. In general, services can be classified as "search" or "experience" services (Mitra et al., 1999). While search services can be objectively evaluated without the need for the consumers to experience them, experience services cannot be evaluated objectively; rather, consumers must engage with them and subjectively evaluate them. (Xiao & Benbasat, 2007).

Search services are more standardized and less personalized than experience services. Hence, there is a more perceived risk in the purchase of experience services than in that search services. Thus, for this study, we selected experience services as our context, and we selected restaurant reviews for data collection as restaurants are one of the most common experience services.

3.1. Data Collection

Consistent with the prior studies (e.g., Wang et al., 2019), we downloaded 46,303 publicly available restaurant reviews for four cities in North America from TripAdvisor for the top 20 restaurants in each of the cities (based on TripAdvisor's ranking). The data consists of all the reviews available as of July 12, 2020. We chose reviews from different cities to enhance the generalizability of the results (Wang et al., 2019). For each review, we extracted restaurant, reviewer and review-related data. At the restaurant level, we extracted the name, city, rank, number of reviews, and the restaurant's price. At the reviewer level, we extracted the number of reviews and the number of helpful reviewers' votes. Finally, at the review level, we extracted the review's star rating, review text, review date, and the number of helpful votes the review received.

3.1.1. Dependent Variable. In this study, the dependent variable is review helpfulness, measured as the number of helpfulness votes that the review received from the review readers.

3.1.2. Independent Variables. The following variables are used as independent variables in this study.

Positive Sentiment (PosSent): Positive sentiment of the review text.

Negative Sentiment (NegSent): Negative sentiment of the review text.

The values of PosSent and NegSent are obtained using the VADER sentiment analysis tool (Hutto & Gilbert, 2014). We normalized PosSent and NegSent values to be between 0-1 using MinMaxScaler.

3.1.3. Control Variables: This study uses the following variables as control variables.

Review Length: Review length represents the extent of the information contained in a review. Previous studies show that longer reviews are more helpful as they contain more information about the product/service and its usage (Mudambi & Schuff, 2010; Filieri, 2016). They are also perceived to be more trustworthy than shorter ones (Filieri, 2016). In this study, we use the number of words in the review text as the review length.

Review Readability: Readability "indicates the extent to which an individual understands and comprehends the product information, which leads to customers accepting information" (Park & Nicolau, 2015). A message is most likely to be read if it is highly readable (Krishnamoorthy, 2015), and it has been shown that there is a positive link between the readability of a review and its helpfulness (Fang et al., 2016; Park & Nicolau, 2015).

To control the review readability, and hence the level of understandability of a review, we used the Coleman–Liau Index (CLI) in the computation of review readability (Coleman & Liau, 1975). CLI shows the average grade level an individual needs to comprehend a text, and the CLI for a given text is calculated using the following formula:

$$CLI = 0.0588L - 0.296S - 15.8$$

where L and S are the average number of letters (characters and numbers) and sentences per 100 words. We utilized spaCy, an open-source Natural Language Processing library, to calculate the CLI index. For English, spaCy provides three models, and we used en_core_web_lg in CLI computation as it is the largest model spaCy provides.

Restaurant Price: Price of a restaurant provides information about the perceived quality of the service or food (Kim et al., 2022), and consumers use price as one of the factors in restaurant assessment (Pantelidis, 2010) and selection (Chow et al., 2007). TripAdvisor provides three price ranges for restaurants, and they are indicated by $ (Cheap Eats), $$-$$$ (Mid-Range) and $$$$ (Fine Dining). After collecting the price range of the restaurants, we coded them as 1 ($), 2 ($$-$$$) and 3 ($$$$).

Restaurant Popularity: A popular restaurant may draw more attention from customers, resulting in more votes for reviews. It has been found that there is a positive relationship between restaurant popularity and review helpfulness (Zhang & Lin, 2018). We use the number of English reviews a restaurant received as the restaurant's popularity.

Reviewer Reputation: "Reviewer reputation refers to the identity-descriptive information displayed on review platforms for users who have contributed reviews" (Chua & Banerjee, 2015, p.355). Chua and Banerjee (2015) show that there is a positive relationship between reviewer reputation and review helpfulness. In this study, we use the number of helpful votes a reviewer has as the reviewer's reputation.

3.2. Procedure

We use sentiment analysis of the reviews to assess the review writers' sentiment. Sentiment analysis is "an active area of study in the field of natural language processing that analyzes people's opinions, sentiments, attitudes, and emotions via the computational treatment of subjectivity in a text" (Hutto & Gilbert, 2014, p.217). For example, Tang and Guo (2015) used text mining to show that the linguistic indicators generated by text analysis are predictive of online review writers' attitudes toward a product/service.

For the text mining tool, we used the VADER (Valence Aware Dictionary and Sentiment Reasoner) library in Python. VADER is a dictionary and rule-based tool developed by Hutto and Gilbert (2014) that enables assessing a sentence's sentiment and intensity. VADER uses a dictionary developed from commonly used word banks such as LIWC (Linguistic Inquiry Word Count), ANEW (Affective Norms for English Words) and GI (General Inquirer), and it extended the dictionary by including emoticons, acronyms and slangs used in online communities (Hutto & Gilbert, 2014). The tool uses internal scores and rules to calculate the sentiments of texts. It was used in previous studies (e.g., Zhang et al., 2019), and according to Hutto and Gilbert (2014), it performs better compared to other sentiment analysis tools.

VADER provides multidimensional (i.e., positive, negative, and neutral) measures of a text. For this study, we only extracted the positive (PosSent) and negative (NegSent) scores for each review text. For text analysis, Chung and Pennebaker (2018) and

Tang and Guo (2015) recommended 100 words per text. Hence, we set the minimum review length to 100 words and removed all reviews shorter than 100 words: this resulted in 12460 reviews out of 46,303 reviews. Moreover, as Chua and Banerjee (2015) contend, not all consumers vote on review helpfulness, and many reviews do not receive helpfulness votes (Fan et al., 2022; Yin et al., 2014). Hence, to reduce noise, similar to Bigne et al., (2021) and Shaft et al. (2020), we removed reviews without helpfulness votes. This resulted in 4705 reviews for our analyses. The descriptive statistics of the variables used in this study are depicted in Table 1.

Table 1 Descriptive statistics for the variables extracted from TripAdvisor.

Variable	Mean	Std.Dev.	Min	Max
Review Helpfulness	1.80	1.79	1	37
Review Star	4.17	1.23	1	5
PosSent	0.43	0.17	0.00	1
NegSent	0.14	0.14	0.00	1
Review Length	190.36	111.95	100	1635
Review Readability	7.63	1.64	1.68	17.07
Restaurant Price	2.46	0.52	1	3
Restaurant Popularity	966.30	699.88	48	2722
Reviewer Reputation	100.23	1810.93	0	122643

3.3. Analysis

Evaluative Space Model suggests that (overall) attitude is the combination of positive and negative attitudes towards stimuli (Cacioppo et al., 1997). Using review helpfulness as the proxy for consumers' (overall) attitude toward textual review, we adopted the following formula to model it:

$$\text{Log(Review Helpfulness)} = P^*(\text{PosSent}^k) + c - N^*(\text{NegSent}^m) + \text{ControlVariables} + e \qquad (2)$$

where PosSent and NegSent are the levels of positivity and negativity activated by the attitude object (i.e., text positivity and text negativity in the restaurant review) $k<1$ and $m<1$ and models the non-linearity of the positive and negative attitudes respectively, P and N are the weighting coefficients, c represents the positivity offset, and e is the error. In Equation 2 $P^*(\text{PosSent}^k) + c$ represents the positive attitude and $N^*(\text{NegSent}^m)$ represents the negative attitude. To find the model parameters, P, k, N, m, and c, we conducted a negative binomial regression analysis similar to Chen et al. (2021) and Yang et al. (2019). Negative binomial regression is suitable for our analysis since review helpfulness is a count variable which cannot be less than one and takes only discrete numeric values. Thus, we are using log of review helpfulness in Equation 2.

Since k and m values are not known, we combined optimization with negative binomial regression analysis using the AIC (Akaike Information Criterion). We ran several negative binomial regressions with different values of k and m, and selected the values with the smallest AIC. AIC is one of the most commonly used model selection criteria (Kuha, 2004) and the lower the AIC, the better the model (i.e., the fit between the model and the data is maximum). We used the Python environment for optimization and the *statsmodels* Python module for negative binomial regression.

3.4. Results

A Python program is implemented for the algorithm stated in the previous section. The analysis showed that the lowest AIC is obtained when $k = 0.4$ and $m=0.9$ with AIC= 16997.64. Table 2 presents the parameters of the optimum solution in Equation 2.

Table 2 Optimum Solution for Equation 2

Review Helpfulness	Coef.	St.Err.	p-val
PosSent	-0.636	0.169	<0.01
NegSent	0.698	0.136	<0.01
Review Length	0.001	0.000	<0.01
Review Readability	0.015	0.012	0.195
Restaurant Price - 2	0.236	0.194	0.223
Restaurant Price - 3	0.228	0.195	0.244
Restaurant Popularity	-0.000	0.000	<0.01
Reviewer Reputation	0.000	0.000	0.029
Constant	0.461	0.239	0.054

As shown in Table 2, coefficients of PosSent and NegSent are significant, with an overall Pseudo R-squared of 3.75%. From Table 2, it is observed that $P = -0.636$, $N = 0.698$, and $c= 0.461$. Combining these with k = 0.4 and m = 0.9, Equation 2 could be written as:

$$\text{Log(Review Helpfulness)} = -0.636 * (\text{PosSent}^{0.4}) + 0.461 + 0.698*(\text{NegSent}^{0.9}) + \text{ControlVariables} + e \quad (3)$$

Since we used negative binomial regression, the dependent variable is in terms of log; hence exponentiation needs to be applied to Equation 3, resulting in Review Helpfulness to be formulated as:

$$\text{Review Helpfulness} = \exp(-0.636*(\text{PosSent}^{0.4})) * 1.585 * \exp(0.698*(\text{NegSent}^{0.9}))*\exp(\text{ControlVariables}) \quad (4)$$

The results provide support for:

(1) *the positivity offset* ($c = 0.461>0$). For example, when there is no positive and negative information in a review of a restaurant, that review on average receives ~2 (~exp(0.461)) helpfulness votes from review readers. This is evidence of the positivity offset.

(2) *the negativity bias* ($N>P$ and $m>k$). The weight of negative sentiment (N) is greater than the weight of positive sentiment (P) in the results. Also, the negativity activation function is steeper than the positivity activation function (i.e., $m > k$).

Figure 2 depicts the plot for Review Helpfulness based on Equation 4. Figure 2a shows the 3D plot of Review Helpfulness as a function of PosSent (text positivity) and NegSent (text negativity) and Figure 2b depicts the associated Positive and Negative attitudes as a function of PosSent and NegSent. Figure 2 is drawn for a mid-priced restaurant (i.e., having a price of \$\$-\$\$\$ on TripAdvisor) and for the other control variables, their mean values are used in the calculation of Equation 4. It is noteworthy that the Positive Attitude in Figure 2b also includes the positivity offset.

4. Discussion

The application of ESM reveals various patterns that may not be captured by using bipolar variables such as review star rating. First, consistent with the ESM, our results show positivity offset exists in review helpfulness assessments. As our results indicate, when there is no positive and negative information, i.e., (PosSent=0, NegSent=0), consumers approach the reviews with a positive attitude expecting the review to be helpful (~2 review helpfulness votes as shown on the heat map in Figure 2(a)). The positivity offset is conceptualized as a positive constant term in the ESM model.

Second, ESM analysis clearly depicts the negativity bias in review helpfulness. For example, as shown in Figure 2b, the negative attitude is steeper than the positive attitude. While positive attitude stays almost stable as a function of PosSent (text positivity), negative attitude shows exponential growth as a function of NegSent (text negativity).

Third, figure 2b explains the inconsistent findings about the relationship between review star rating and review helpfulness. As shown in Figure 2b, there is an intersection point between positive and negative attitudes. Before the intersection point, the overall attitude is dominated by the positive attitude; after the intersection point, the overall attitude is dominated by the negative attitude. Hence, we conclude that positive dominant reviews (reviews

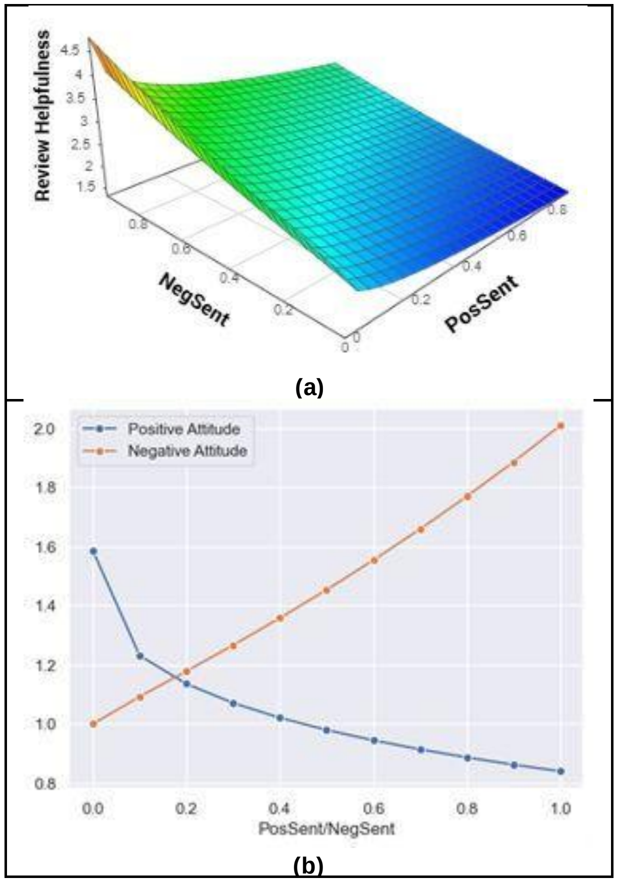

Figure 2. The plot of review helpfulness as provided in Equation 4.

having text positivity and text negativity before the intersection point) and negative dominant reviews (reviews having text negativity after the intersection point) could both be more helpful. On the one hand, if extant literature had assessed review helpfulness using star rating within the zone before this intersection point, positive attitude would have been found to be more useful than negative attitude. On the other hand, if extant literature had assessed review helpfulness using star rating after the intersection point, negative attitude would have been found to be more useful than positive attitude. As such, the effects of positive and negative attitudes are asymmetric, a phenomenon that bipolar scales such as star rating cannot accurately detect.

Fourth, we found that the relationship between the evaluative space (i.e., PosSentand NegSent) and the attitude space (i.e., Review Helpfulness) is non-linear. In particular, different PosSent and NegSent points can represent the same attitude. For example, the review helpfulness of 2(rounded) is represented by both (PosSent=0.4, NegSent=0.3) and (PosSent=0.6, NegSent= 0.4); similarly, the review helpfulness of 4(rounded) is captured by both points (PosSent=0.1, NegSent= 0.9) and (PosSent=0.0, NegSent= 0.7) on the evaluative space. Such a relationship couldn't be captured if the review star rating is used in the assessment of review helpfulness. This is due to the fact that review star rating is a bipolar variable, and "it measures both the direction (side of the scale) and intensity (distance from the center) of the respondent's position on the concept of interest" (Lavrakas, 2008). By using ESM, we modelled the attitude (review helpfulness) using two unipolar scales, namely PosSent and NegSent, enabling us to capture the bidimensional relationship between the underlying positive and negative sentiments and review helpfulness.

5. Concluding Remarks

ESM suggests a nonlinear relationship between positive and negative attitudes – which are functions of PosSent (text positivity) and NegSent (text negativity), respectively – and review helpfulness. Hence, proper identification of positive and negative attitudes become an important task, as review helpfulness heavily depends on the location of evaluative (PosSent, NegSent) points. As mentioned above, many studies used review star rating as one of the explanatory variables for review helpfulness. Yet, the findings related to significance of positive/negative sentiment in terms of helpfulness is mixed. The reason for this is that the bipolar nature of star rating cannot accurately capture the writer’s attitude expressed in a review text. A combination of positive and negative attitudes can result in four possible states, which cannot be accurately represented by a bipolar scale (Audrezet & Parguel, 2018). To that end, the significance of positive/negative attitudes depends on how salient each is in shaping the overall attitude. Thus, they can create four states: Positive to be more influential than negative in determining helpfulness of a review; negative can be more influential than positive, both can be strongly influential, or both can be weakly influential. In addition, the relationship between the evaluative space (PosSent, NegSent) and the attitude space (i.e., review helpfulness) is non-linear. Hence, our contention in this paper is that a bipolar scale, such as review star rating, for capturing the overall attitude toward a product/service would ignore this weight imbalance between negative and positive attitudes and hence the use of review star rating could be the cause of inconsistent findings in the assessment of review helpfulness.

This study has both theoretical and practical contributions. We were able to contribute to the online review literature about review helpfulness by considering the positive and negative attitudes toward a review separately. Our results confirmed the negativity bias, stating that the negative attitudes have more weight than positive attitudes, and the positivity offset, stating that individuals are inclined to approach a stimulus, e.g., review, in a positive way in the case of no information (e.g., expecting a review to be helpful before reading it).

This study also has practical contributions. Service providers can use ESM of reviews for their products/services to enhance their marketing tactics. Case in point, we know that ambivalence elicits physiological arousal. Consumers’ high arousal state reduces their systematic processing of information, instead engaging more heuristic processing which is a more intuitive and less analytical way of thinking (Sanbonmatsu & Kardes, 1988). This arousal caused by ambivalence makes consumers focus on the immediate consequences of their choices. As a result, when the immediate outcomes of consuming a product are positive, ambivalence can increase the likelihood that consumers will approach and choose that product (Hamby & Russel, 2020; Turel & Qahri-Saremi, 2023). For instance, even though a restaurant may have high menu prices that consumers view negatively, they may still be tempted to dine there because of the high quality of the food, services, and ambiance. In this case, service providers should develop marketing strategies that emphasize the immediate positive appeal of their products/services, making it greater than any negative aspects. This can help attract customers and increase sales. In the same vein, future research can investigate the moderating effect of the number of concepts discussed in a review on the relation between its positivity/negativity and its usefulness.

6. References

Audrezet, A., & Parguel, B. (2018). Using the Evaluative Space Grid to better capture manifest ambivalence in customer satisfaction surveys. Journal of Retailing and Consumer Services, 43, 285–295. https://doi.org/10.1016/j.jretconser.2018.04.008

Bigne, E., Ruiz, C., Cuenca, A., Perez, C., & Garcia, A. (2021). What drives the helpfulness of online reviews? A deep learning study of sentiment analysis, pictorial content and reviewer expertise for mature destinations. Journal of Destination Marketing & Management, 20, 100570.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. Personality and Social Psychology Review, 1(1), 3-25.

Chatterjee, S. (2020). Drivers of helpfulness of online hotel reviews: A sentiment and emotion mining approach. International Journal of Hospitality Management, 85, 102356.

Chen, M.-Y. (2016). Can two-sided messages increase the helpfulness of online reviews? Online Information Review, 40(3), 316–332. https://doi.org/10.1108/OIR-07-2015-0225

Chen, Y., Jin, W., Hu, Y., Zhou, S., & Yang, S. (2022). Does managerial response moderate the relationship between online review characteristics and review helpfulness? Current Issues in Tourism, 25(16), 2679–2694. https://doi.org/10.1080/13683500.2021.1988523

Chow, I. H., Lau, V. P., Wing-chun Lo, T., Sha, Z., & Yun, H. (2007). Service quality in restaurant operations in China: Decision- and experiential-oriented perspectives. International Journal of Hospitality Management, 26(3), 698–710. https://doi.org/10.1016/j.ijhm.2006.07.001

Chua, A. Y., & Banerjee, S. (2015). Understanding review helpfulness as a function of reviewer reputation, review rating, and review depth. Journal of the Association for Information Science and Technology, 66(2), 354-362.

Chung, C. K., & Pennebaker, J. W. (2019). Textual analysis. In H. Blanton, J. M. LaCroix, & G. D. Webster (Eds.), Measurement in social psychology (pp. 153–173). Routledge/Taylor & Francis Group.

Coleman, M., & Liau, T. L. (1975). A computer readability formula designed for machine scoring. Journal of Applied Psychology, 60(2), 283-284.

DeCastellarnau, A. (2018). A classification of response scale characteristics that affect data quality: A literature review. Quality & Quantity, 52(4), 1523–1559.

Eagly, A. H., and Chaiken, S. 1993. The Psychology of Attitudes. Harcourt Brace Jovanovich.

Fan, W., Liu, Y., Li, H., Tuunainen, V. K., & Lin, Y. (2022). Quantifying the effects of online review content structures on hotel review helpfulness. Internet Research, 32(7), 202-227.

Fang, B., Ye, Q., Kucukusta, D., & Law, R. (2016). Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics. Tourism Management, 52, 498-506.

Filieri, R. (2016). What makes an online consumer review trustworthy? Annals of Tourism Research, 58, 46–64. https://doi.org/10.1016/j.annals.2015.12.019

Godes, D., & Mayzlin, D. (2004). Using online conversations to study word-of-mouth communication. Marketing science, 23(4), 545–560.

Hamby, A., & Russell, C. (2022). How does ambivalence affect young consumers' response to risky products? Journal of the Academy of Marketing Science, 50(4), 841–863. https://doi.org/10.1007/s11747-021-00834-7

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? Journal of Interactive Marketing, 18(1), 38–52. https://doi.org/10.1002/dir.10073

Hong, H., Xu, D., Wang, G. A., & Fan, W. (2017). Understanding the determinants of online review helpfulness: A meta-analytic investigation. Decision Support Systems, 102, 1-11.

Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. Proceedings of the International AAAI Conference on Web and Social Media, 8(1), 216–225.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47(2), 363–391.

Kim, J., Lee, M., Kwon, W., Park, H., & Back, K.-J. (2022). Why am I satisfied? See my reviews – Price and location matter in the restaurant industry. International Journal of Hospitality Management, 101, 103111. https://doi.org/10.1016/j.ijhm.2021.103111

Klopfer, F. J., & Madden, T. M. (1980). The Middlemost Choice on Attitude Items: Ambivalence, Neutrality, or Uncertainty?. Personality and Social Psychology Bulletin, 6(1), 97-101.

Krishnamoorthy, S. (2015). Linguistic features for review helpfulness prediction. Expert Systems with Applications, 42(7), 3751–3759. https://doi.org/10.1016/j.eswa.2014.12.044

Kuha, J. (2004). AIC and BIC: Comparisons of assumptions and performance. Sociological Methods & Research, 33(2), 188–229.

Lavrakas, P. (2008). Encyclopedia of Survey Research Methods. Sage Publications, Inc. https://doi.org/10.4135/9781412963947

Lopez, A., & Garza, R. (2022). Do sensory reviews make more sense? The mediation of objective perception in online review helpfulness. Journal of Research in Interactive Marketing, 16(3), 438–456. https://doi.org/10.1108/JRIM-04-2021-0121

Malhotra, N. K. (1984). Reflections on the Information Overload Paradigm in Consumer Decision Making. Journal of Consumer Research, 10(4), 436–440.

Malik, M. S. I., & Hussain, A. (2017). Helpfulness of product reviews as a function of discrete positive and negative emotions. Computers in Human Behavior, 73, 290–302. https://doi.org/10.1016/j.chb.2017.03.053

Meta. (2023). Meta Reports Fourth Quarter and Full Year 2022 Results. https://investor.fb.com/investornews/press-release-details/2023/Meta-Reports-Fourth-Quarter-and-Full-Year-2022-Results/default.aspx

Mitra, K., Reiss, M. C., & Capella, L. M. (1999). An examination of perceived risk, information search and behavioral intentions in search, experience and credence services. Journal of Services Marketing, 13(3), 208–228. https://doi.org/10.1108/08876049910273763

Mudambi & Schuff. (2010). Research Note: What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com. MIS Quarterly, 34(1), 185-200. https://doi.org/10.2307/20721420

Pantelidis, I. S. (2010). Electronic Meal Experience: A Content Analysis of Online Restaurant Comments. Cornell Hospitality Quarterly, 51(4), 483–491. https://doi.org/10.1177/1938965510378574

Park, S., & Nicolau, J. L. (2015). Asymmetric effects of online consumer reviews. Annals of Tourism Research, 50, 67–83. https://doi.org/10.1016/j.annals.2014.10.007

Picher Vera, D., Martínez María-Dolores, S. M., & Bernal García, J. J. (2016, June). Social networks as a communication, sales and customer service tool. Analysis and effectiveness of eWOM as a marketing strategy. 14th International Conference on Marketing.

Qahri-Saremi, H., & Montazemi, A. (2019). Factors Affecting the Adoption of an Electronic Word of Mouth Message: A Meta-Analysis. Journal of Management Information Systems, 36(3), 969–1001. https://doi.org/10.1080/07421222.2019.1628936

Rossiter, J. R. 2010. Measurement for the Social Sciences: The C-OAR-SE Method and Why It Must Replace Psychometrics, Springer Science & Business Media.

Sanbonmatsu, D. M., & Kardes, F. R. (1988). The Effects of Physiological Arousal on Information Processing and Persuasion. Journal of Consumer Research, 15(3), 379-385. https://doi.org/10.1086/209175

Schneider, I. K., & Schwarz, N. (2017). Mixed feelings: The case of ambivalence. Current Opinion in Behavioral Sciences, 15, 39–45. https://doi.org/10.1016/j.cobeha.2017.05.012

Sears, D. O. (1983). The person-positivity bias. Journal of Personality and Social Psychology, 44(2), 233-250.

Shaft, T., Tian, A. C., & Whang, S. Y. (2020). Anxious and Angry: A Replication Investigating the Effects of Emotions on Perceptions of Online Review Helpfulness. AIS Transactions on Replication Research, 6(1), 17.

Snyder, A. I., & Tormala, Z. L. (2017). Valence asymmetries in attitude ambivalence. Journal of Personality and Social Psychology, 112(4), 555–576. https://doi.org/10.1037/pspa0000075

Solomon, M. (2008). Consumer behavior buying, having, and being (8th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Tang, C., & Guo, L. (2015). Digging for gold with a simple tool: Validating text mining in studying electronic word-of-mouth (eWOM) communication. Marketing Letters, 26(1), 67–80.

Turel, O., & Qahri-Saremi, H. (2023). Responses to ambivalence toward social networking sites: A typological perspective. Information Systems Journal, *33*(2), 385-416. https://doi.org/10.1111/isj.12407

Van Harreveld, F., Nohlen, H. U., & Schneider, I. K. (2015). The ABC of ambivalence: Affective, behavioral, and cognitive consequences of attitudinal conflict. In J. M. Olson & M. P. Zanna (Eds.), Advances in experimental social psychology (vol. 52, pp. 285–324). Academic Press.

Walden, E. A., Browne, G. J., & Larsen, J. T. (2005). Ambivalence and the Bivariate Nature of Attitudes in Information Systems Research. In System Sciences, 2005. HICSS'05. Proceedings of the 38th Annual Hawaii International Conference on (pp. 263). IEEE.

Wang, J.-N., Du, J., & Chiu, Y.-L. (2020). Can online user reviews be more helpful? Evaluating and improving ranking approaches. Information & Management, 57(8), 103281. https://doi.org/10.1016/j.im.2020.103281

Wang, X., Tang, L. R., & Kim, E. (2019). More than words: Do emotional content and linguistic style matching matter on restaurant review helpfulness?. International Journal of Hospitality Management, 77, 438-447.

Weick, K.E. (1995) Sensemaking in organizations. Sage Publications, Inc.

Whitler, K. A. (2014). Why Word Of Mouth Marketing Is The Most Important Social Media. Forbes. https://www.forbes.com/sites/kimberlywhitler/2014/07/17/why-word-of-mouth-marketing-is-the-most-important-social-media/

Xiao, B., & Benbasat, I. (2007). E-Commerce Product Recommendation Agents: Use, Characteristics, and Impact. MIS Quarterly, 31(1), 137-209. https://doi.org/10.2307/25148784

Yang, L., & Unnava, H. (2016). Ambivalence, Selective Exposure, and Negativity Effect. Psychology & Marketing, 33(5), 331–343.

Yang, S., Zhou, Y., Yao, J., Chen, Y., & Wei, J. (2019). Understanding online review helpfulness in omnichannel retailing. Industrial Management & Data Systems, 119(8), 1565–1580. https://doi.org/10.1108/IMDS-10-2018-0450

Yin, D., Bond, S. D., & Zhang, H. (2014). Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews. MIS quarterly, 38(2), 539-560.

Yelp. (2023). Yelp Metrics as of June 30, 2023. (https://www.yelp-press.com/company/fast-facts/).

Zhang, P., Lee, H.-M., Zhao, K., & Shah, V. (2019). An empirical investigation of eWOM and used video game trading: The moderation effects of product features. Decision Support Systems, 123, 113076. https://doi.org/10.1016/j.dss.2019.113076

Zhang, Y., & Lin, Z. (2018). Predicting the helpfulness of online product reviews: A multilingual approach. Electronic Commerce Research and Applications, 27, 1–10.