Tech Science Press

The Role of Emotions Intensity in Helpfulness of Online Physician Reviews

Adnan Muhammad Shah and KangYoon Lee[*]

[1]Department of Computer Engineering, Gachon University, 13120, Korea
[*]Corresponding Author: KangYoon Lee. Email: keylee@gachon.ac.kr

Abstract: Online physician reviews (OPRs) critically influence the patients' consultation decisions on physician rating websites. The increasing number of OPRs contributes to the challenge of information overload. The worth of development needs to be explored further. Based on the OPRs collected from RateMDs and Healthgrades, and Plutchik's wheel on human emotions framework, the purpose of this study was to examine the impact of emotional intensity (positive and negative) incorporated in OPRs on review helpfulness (RH). The proposed model was empirically tested using data from two physician rating websites and applying a mixed-methods approach (text mining and econometrics). The results suggested that anger, disgust, and fear (negative emotions), while joy, anticipation, and trust (positive emotions) significantly contributed to the perceived RH. Moreover, the patient's disease severity moderated the association between negative emotions (anger and disgust), or positive emotions (joy and trust) and the perceived RH. Anger incorporated in an OPR has a more negative impact on perceived RH for severe diseases than the mild diseases. The findings contributed to the significance of emotions in OPRs from an innovative perspective and provided practical insights for health rating platform owners to help patients in expressing their emotions more precisely.

1 Introduction

The utilization of Web-based social platforms reveals an evolving drift for health consumers who are looking for health information for their clinical decisions. People increasingly use social media to connect with peers having the same health problems and get updates on treatment options [1]. In the current age of digital health, physician rating websites (PRWs) have been important for patients, as more individuals explore online information for their healthcare decisions [2]. PRWs such as RateMDs and Yelp allow online word-of-mouth (WOM) as online physician reviews (OPRs) [3]. OPRs are commonly considered as a bundle of knowledge that represents the segments of health consumers, with the intention of engaging, sharing, and even persuading other health consumers [4]. As a result, research regarding OPRs posted on PRWs has drawn researchers' interest by focusing on how and in what way they affect the behaviors of other patients in relation to the user-generated content [5].

Given the OPRs provides relevant information regarding a physician's healthcare quality [4], a massive quantity of reviews can produce significant information overload that fatigue readers and increases search cost [6]. Several PRWs such as Vital and Yelp have features that allow users to assign helpful votes to a review based on their perceptions. The helpfulness of a review (from there onward RH) has a significant influence on patients' information search process. RH reflects the extent to which a review helps users in their buying decisions by questioning on PRWs "*Was this review helpful*?" Based on these assumptions, the more the reviews get more helpful votes, the more likely these reviews to be read by users [7].

Prior research has been performed to analyze the determinants that affect RH concerning a review's content and context features. Content features are extracted from a review, such as review length [8,9], visibility [10,11], quality [12], linguistic style [13], and review emotions [9,14,15]. The context features include reviewer expertise [16,17] and review extremity [7,18]. Most importantly, an evolving stream of literature on review content as well as emotional words turns out to be relevant in terms of emotional enchantment. The emotional features can be divided into two components: sentiments and discrete emotions. The review sentiment (positive or negative) is significantly associated with the RH [19–21]. In contrast, previous studies also explored the impact of emotional intensity on RH [6,22–27]. However, certain problems in previous research still need to be examined.

(1) A continuation of previous work is required to track the impact of different positive and negative emotional intensity on physician RH. To address this problem, we present the following question.

   How does the intensity of the positive and negative emotions incorporated in reviews affect the perceived RH?

(2) Regarding the influence of different emotions on RH, we examine which specific emotion will positively or negatively affect perceived RH. Therefore, our second research question states that:

   Which emotion leads to more and less RH voting?

(3) Prior work explored the moderating role of patient's disease severity in influencing the patient evaluation of healthcare service quality [28,29]. For example, Yang, Guo [30] indicated that the influence of response speed and interaction frequency on patient satisfaction were higher for high-risk disease than the low-risk disease. In particular, the earlier research still did not address the moderating role of disease severity on the association between specific emotions (positive/negative) and perceived RH. Based on these arguments, we proposed the following:

   How could the disease severity (high-disease severity or low-disease severity) moderate the influence of individuals' emotions (positive or negative) intensity on the perceived RH?

To answer the above research questions, the specific objectives of the study involved: (1) drawing on Plutchik's emotion framework to examine the influence of eight core emotions intensity dimensions (anger, sadness, disgust, fear, joy, anticipation, surprise, and trust) incorporated in a review on RH; (2) to examine which emotion dimension leads to more and less helpful voting; (3) to explore the moderating role of disease severity in the relationships between specific emotional dimensions and RH.

Given the search and experience goods context, previous studies on RH either focus on survey-based data collection [17], experiments [31], interviews [32] and secondary data analysis from online platform [9,13,20]. The exponential growth of big data increases secondary data analytics with several advantages (e.g., time and cost-effective, large volume, real-time and more objective response) [33,34]. So far, only two studies have been published [35,36], which explicitly analyze UGC to predict OPRs helpfulness using traditional qualitative and machine learning methods. However, we propose a model in the current study to analyze patients' discrete emotions implicitly using a mixed-methods approach (i.e., text mining

and econometrics). In particular, we test hypotheses using larger OPRs datasets (nearly 0.1 million reviews) from RateMDs.com and Healthgrades.com. Major contributions of this research are as follows:

(1) This is the first research to look at the impacts of positive or negative emotional intensity incorporated in OPRs on RH prediction.
(2) The methodology applied in this research implicitly mine positive and negative emotional aspects from OPRs based on the text mining algorithms.
(3) The moderating role of a patient's disease severity, which could interpret the inconsistency of RH between high and low-disease severity, is likely to reveal interesting findings to this research.

2 Theoretical Background and Research Hypotheses

Several emotion theories suggest essential human emotions from different aspects [37,38]. For example, Mehrabian and Russell [39] considered three dimensions to quantify the emotional state of the customer, have established a pleasure-arousal-dominance (PAD) model. Within the arousal component, high-arousal emotions activate a physical response and then initiate knowledge sharing. High and low arousal were related to optimism and uncertainty, respectively [40]. Francisco and Gervás [41] also suggested three types of emotions (pleasantness, activation, and dominance) and classified phrases in fairy tales for these types. Ekman [37] proposed six basic dimensions of emotion (i.e., joy, sadness, anger, fear, disgust, and surprise). Plutchik [42] further added the two emotion dimensions (i.e., trust and anticipation). The study categorized four dimensions as positive emotions (joy, surprise, trust, and anticipation) and other four as negative emotions (sadness, fear, disgust, and anger). Prior research showed that discrete emotions (e.g., happiness, frustration, and anger), which described the events that have induced emotions (e.g., valence, control, certainty) can be contrasted in different dimensions [43].

The rationale behind the adoption of Plutchik's framework in this study was as follows. First, Plutchik's wheel on human emotions has been well described in previous psychological studies. Second, Plutchik's model offset positive and negative emotional experiences, in comparison with those in which negative emotions dominate [37]. Third, the emotional aspects of the current research are superset of some other study [37]. Plutchik's framework has been extensively used in various other domains while analyzing textual content posted on the internet, including online reviews [6,22,23,25,44]. Since the context of the study is the online healthcare industry (credence goods) in which patients evaluate the physician's offline service quality through online information channels. Compared with other goods categories, healthcare service is a credence commodity (i.e., an entity whose quality is difficult to quantify for users even after consumption). The online healthcare industry includes more information asymmetry, as these represent both credence goods and online markets. Hence, analyzing OPRs are more likely to elicit distinct emotions of consumers. Satisfied and unhappy customers share their positive and negative experiences with other consumers by posting online. For example, if a patient is unhappy with the service quality, s/he will investigate if the complaint is treated reasonably, generate valence as a failure to assess service, and then express emotions online to others to gain social recognition or encourage empathy. If a customer believes that the service provider's fault or service retrieval failed, s/he will be frustrated, lose faith, and even displease the service provider. In contrast, an individual might feel satisfied or pleased due to the service recovery. As a result, displeasing service practice can contribute to various emotional dimensions. The transition of impacts from multidimensional emotional factors to readers in reviews describes their judgments and interpretation. The conceptual model of our study is shown in Fig. 1.

Although online information is critical for patients' choice of their doctors, there is scant research investigating the impact of discrete emotion on RH. The direction of relationship regarding the impact on positive and negative emotions on RH is still in debate. There are only two studies that investigated the

impact of discrete emotions on RH, but both of them did not discuss the direction of relationship [23,24]. Given that the earlier studies have confirmed the positive effect of either positive or negative emotions on RH, it is considered that all six emotional dimensions (fear, rage, disgust, confidence and excitement, anticipation) have a positive influence on RH. Still, we discussed that the sense of sadness or joy from service experience is addressed particularly as inflated and probably even invented because excessively emotional reviews are regarded as doubtful and pretended. Moreover, in-service lapse research, anger, dissatisfaction, or concern has been used as a discrete emotion rather than sadness; for instance, despite the negative emotions, anger and sadness lead to different kinds of cognitive responses [25].
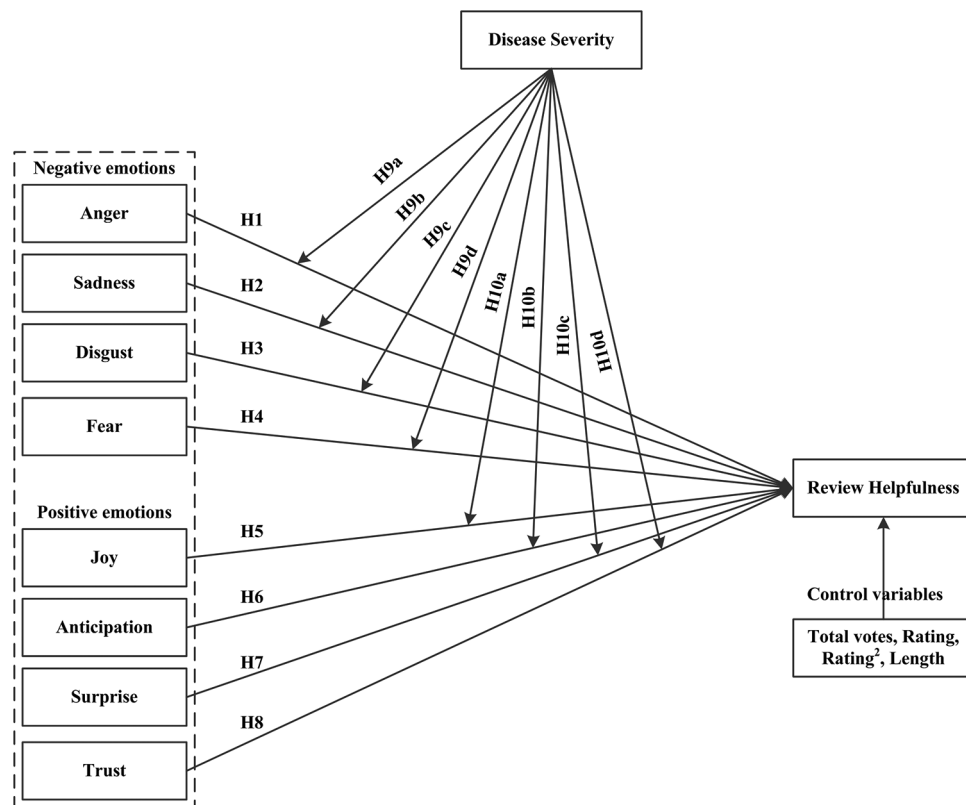
Figure 1: Conceptual framework

From the perspective of online reviews, as angry individual tends to make stereotypical assumptions, racism, and a poor level of argument [45], consumers invest less cognitive energy on anger-incorporated in online feedback [46]. Yin, Bond [46] have shown that anger has a detrimental effect on the RH. Fear in reaction to purchase experience is found to positively impact the attitude, intentions, and behaviors of the customers [47]. Those messages rooted with fear aim to boost anxiety and encourage the recipients to follow the advice by highlighting possible adverse effects. Fearful messages hinder the discovery of new technologies by people that may be regarded as potentially risky or faulty [48]. Sad people often face trouble in their rational decision-making [45]. The sadness-involved reviews lack a detailed assessment of the goods. Previous research has also found the negative impact of sadness on RH [40]. The anger and anxiety as dimensions of negative emotions intensity significantly influenced RH [26]. Thus, the following hypotheses have been proposed to explore the influence of emotions intensity on RH.

H1: Anger-incorporated online reviews negatively affect the perceived RH.

H2: Sadness-incorporated online reviews positively affect the perceived RH.

H3: Disgust-incorporated online reviews positively affect the perceived RH.

H4: Fear-incorporated online reviews positively affect the perceived RH.

H5: Joy-incorporated online reviews negatively affect the perceived RH.

H6: Anticipation-incorporated online reviews negatively affect the perceived RH.

H7: Surprise-incorporated online reviews positively affect the perceived RH.

H8: Trust-incorporated online reviews positively affect the perceived RH.

Individuals' tastes vary with regard to credence services and, therefore, subjective in nature. Credence services are rated by consumers who post either positive or negative reviews [7]. Consumers may use extreme ratings to find similar services and evaluate the information provided by others [49]. For example, health consumers are more likely to consult a doctor who provides higher-quality care. A high-quality healthcare service is more likely to retain customers and gain feedback (i.e., online reviews) detailing their satisfaction with the service. Increased service knowledge among health customers increases the likelihood of receiving high-quality service and helpful reviews. Patients who suffer from serious diseases may expect a better quality of services than patients who suffer from mild diseases [30].

Because of the poor quality, anger and sadness found in reviews are generally involved in an unfavorable consumption experience. Anger and sadness involved in reviews could lead to extreme scores. Reviews involved extreme ratings were considered more helpful than lower ratings for low-risk disease [7]. In case of high-disease severity, as extreme reviews give more realistic knowledge than the extreme reviews of low-disease severity, we suggest the negative impact of anger and sadness would be attenuated in the review of low-disease severity. Based on the earlier discussion, anger and sadness have detrimental effects on perceived RH. Hence, we claim that the influence of anger and sadness on RH is greater for high-disease severity than for low-disease severity. Therefore, the following hypotheses are proposed:

H9$_a$: Disease severity negatively moderates the relationship between anger and perceived RH.

H9$_b$: Disease severity negatively moderates the relationship between sadness and perceived RH.

We also consider the disease severity as a factor moderates the association between the intensity of the particular emotion and RH. Patients with varying disease conditions may need different levels of quality of health services [28]. Patients suffer from serious disease (high-risk disease) may need a better quality of services than patients with mild disease (low-risk disease) [29]. Hence, we assume that emotions incorporated in OPRs will provide detailed information, likewise service details for high-disease severity than for low-disease severity. In this vein, the sentiment of specific emotions can be anticipated more helpful in the case of serious diseases, as these OPRs are perceived as less displeasing by readers who disagree with the expressed opinion. As a result, this study assumes the following hypotheses:

H9$_c$: Disease severity positively moderates the relationship between disgust and perceived RH.

H9$_d$: Disease severity positively moderates the relationship between fear and perceived RH.

H10$_a$: Disease severity positively moderates the relationship between joy and perceived RH.

H10$_b$: Disease severity positively moderates the relationship between anticipation and perceived RH.

H10$_c$: Disease severity positively moderates the relationship between surprise and perceived RH.

H10$_d$: Disease severity positively moderates the relationship between trust and perceived RH.

3 Methods

This study uses data from RateMDs and Healthgrades since both these PRWs pioneered the RH features, which has been a core feature of these sites until now. The helpfulness of OPRs published on both these sites continues to draw the scholars' attention [7,36,50]. Data were collected in March 2020, using a network spider coded in Python 3.6. The scraped reviews were posted between January 2012 and September 2019, ensuring that all feedback had been made available to users for at least six months.

Regarding the disease severity, we chose 10 different diseases (i.e., Cancer, Heart disease, Accidents, Chronic lower respiratory disease, Stroke, Alzheimer's disease, Diabetes, Suicide, Pneumonia or influenza, and Nephropathies) based on the mortality rate in the United States [51]. Since, due to time, resources, and access constraints, it is nearly impossible to include the entire population of diseases in a single study, a representative sample has therefore been selected. After removing duplicate and non-English OPRs, a total of 94,102 OPRs were obtained. Simultaneously, to evaluate RH, the number of votes "helpful or useful" has been collected for each individual review. After deleting the reviews without helpful or useful votes, we received a total of 89,134 reviews for final data analysis.

Our study's data processing comprises mainly two parts; one focuses on text analytics and the other on empirical inquiry. Text mining is primarily aimed at extracting the number of emotion words from the unstructured text information. The core objective of the empirical analysis was to look at the impact of the discrete emotions on RH. In the text mining process, several pre-processing steps have been performed, including spell correction using an open-source spell checker software *Google Spell Check*, removing special characters, symbols, URL's, punctuations, numbers, words occurrence fewer than 10 times in the corpus, and redundant words in the dataset, etc. Next, Python package NLTK [52] was used for stop words removal, word and sentence tokenization, lemmatization, and part-of-speech tagging. The technical route of the current research is shown in Fig. 2. In addition, the study variables are defined as following.
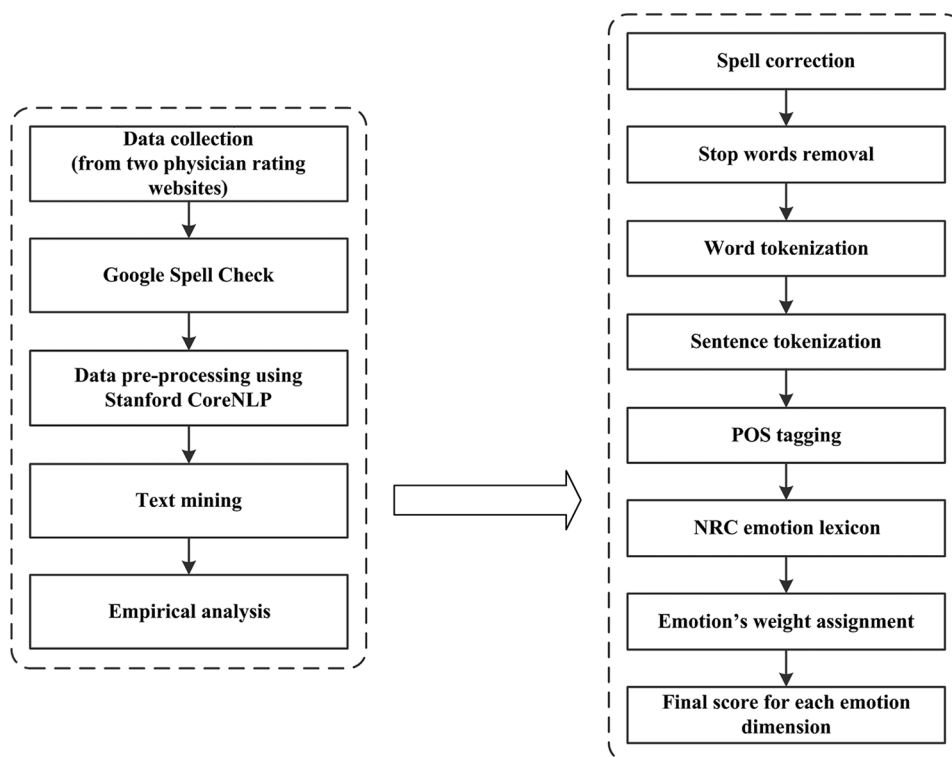
Figure 2: Research technical route

(1) *Dependent Variable*: RH, which worked as the dependent construct in this study, can be described in the following way:

$$Helpfulness_i = \frac{HelpfulVotes_i}{ElapsedDay_i} \tag{1}$$

where the number of votes for review is represented by *HelpfulVotes_i*, and *ElapsedDay_i* indicates the duration (the number of days) between the review posting and crawling date.

(2) *Independent Variables*: We measured the independent variables by considering eight emotional dimensions used in the current study. Prior investigation claimed that the number of features involved in a review (e.g., the number of affirmations; the number of concepts) significantly influences the perceived RH [8]. We assume these studies are important to the present work. Using the discrete positive and negative emotions, the proposed model offers improved prediction accuracy for evaluating the RH prediction problem.

(3) *Moderator Variable*: We coded the moderator (Disease severity) with a dummy variable (1, low-disease severity; 0, high-disease severity) from the perspective of mortality rate. Prior studies have also often used patient's disease severity as a moderator variable [7,28,30,33,50].

(4) *Control Variables*: We also included several control variables to make the calculation more convincing. For instance, the total number of votes for each RH, review rating, square of review rating, and review length (word count). Tab. 1 presents an interpretation of the constructs.

Next, we used an emotion lexicon established by the Canadian National Research Council (NRC) [53]. NRC introduced a broad word-emotion lexicon that included over 8265-word types. This NRC classifies the word sentences into 8 emotional dimensions proposed by Plutchik [42].

Table 1: Variables and their operationalization

Variable type	Variable level	Variable	Operationalization
DV	Individual review	Helpfulness	# Helpful votes/# Elapsed day
IV	Individual review	Anger	# anger-related words
	Individual review	Sadness	# sadness-related words
	Individual review	Disgust	# disgust-related words
	Individual review	Fear	# fear-related words
	Individual review	Joy	# joy-related words
	Individual review	Anticipation	# anticipation-related words
	Individual review	Surprise	# surprise-related words
	Individual review	Trust	# trust-related words
CV	Individual review	Total votes	# total votes
	Physician	Rating	# star rating
	Physician	Rating^2	A quadratic term of star rating
	Individual review	Length	# words in a review
M		Disease severity	Dummy variable: 1 for low-disease severity, 0 for high-disease severity

Notes: DV = Dependent variable; IV = Independent variables; CV = Control variables; M = Moderator

After scraping review text from both rating sites, all words were extracted from the review corpus. Every single word is processed, and emotion characteristics are determined using the scoring method shown in Algorithm 1. For each word, the scoring method (lines 7–9) uses the NRC Emotions Lexicon to check for a number of words relevant to each emotion aspect. When a match happens, then the value of the required emotional factor is increased in DE (discrete emotions). This procedure is iterated for all words derived from all phrases in the review text.

The algorithm further calculates the 8 emotions' weights for each review in the corpus (lines 1 to 17). Then for each review, the final scores of all the emotion dimensions are determined (lines 14–16). The configuration of DE's emotional dimensions is described as: DE < Anger, sadness, disgust, fear, joy, anticipation, surprise, trust >. We may assume that DE1 is a tuple containing 8 characteristics (emotional values) of the first review. The final result of a specific emotional dimension relating to the first review is scientifically calculated as:

$$DE_1 (\text{Discrete Emotions}) <trust> = \left(\frac{\text{Number of trust} - \text{related words}}{\text{Number of words in a review}} \right) \times 100 \tag{2}$$

Next, a regression model is presented to analyze data. This allows us to test hypotheses, including the influence of specific emotional intensity on RH. The moderating role of disease severity in the relationships between discrete emotions and RH is also tested. Our empirical model is shown below.

$$\begin{aligned}
\textit{Model}: \text{Helpfulness\%} = {} & \beta_1\,Anger + \beta_2\,Anticipation + \beta_3\,Disgust + \beta_4\,Fear + \\
& \beta_5\,Joy + \beta_6\,Sadness + \beta_7\,Surprise + \beta_8\,Trust + \beta_9\,TotalVotes + \beta_{10}\,Rating + \\
& \beta_{11}\,Rating^2 + \beta_{12}\,Length + \beta_{13}\,Anger \times DiseaseSeverity + \beta_{14}\,Anticipation \times DiseaseSeverity + \\
& \beta_{15}\,Disgust \times DiseaseSeverity + \beta_{16}\,Fear \times DiseaseSeverity + \beta_{17}\,Joy \times DiseaseSeverity + \\
& \beta_{18}\,Sadness \times DiseaseSeverity + \beta_{19}\,Surprise \times DiseaseSeverity + \beta_{20}\,Trust \times DiseaseSeverity + \epsilon
\end{aligned} \tag{3}$$

Algorithm 1: Discrete Emotions Characteristics Scoring Procedure

---

1: for each OPR, u_i in U do

2:         DE_i = [0,0,0,0,0,0,0,0] //DE_i = [Anger, anticipation, disgust, fear, joy, sadness, surprise, trust]

3:             Word-count = 0 // initialize the word count

4:             For each sentence r_j in t_j do

5:                 For each word u_k in r_j

6:                     Increment word-count

7:                     For each emotion dimension, d_m in NRC do

8:                         If (u_k belongs to NRC [d_m])

9:                             Increment the score in DE_i [m]

10:                        End if

11:                    End for

12:                End for

13:         End for

14:         for each element, n in DE_i do

15:            compute DE_i [n]

16:        End for

17:        End for

4 Results

Tab. 2 provides the variable statistics results. In addition, we used Tobit regression to analyze the model. To check the hypothetical relationships, the data were analyzed using STATA 16. Moreover, variables correlations are shown in Tab. 3. A multi-collinearity test was carried out to detect high correlations between the variables: anger, sadness, disgust, fear, Joy, anticipation, surprise, trust, and their interaction. The results show that all the absolute values of correlation coefficients and the variance inflation factor (VIF) statistics for every independent variable are below 0.5 and 10, respectively; hence multicollinearity is not a serious issue in this study, and the results are reliable [54]. The VIF for a regression model variable is the ratio of the overall model variance to the variance of a model that includes only that single independent variable.

Table 2: Variable statistics

Role	Variable	Mean	S.D.	Min	Max
DV	Helpfulness	0.84	0.23	0.05	1.00
IV	Anger	0.93	1.59	0	33.75
	Sadness	3.33	2.68	0	49.23
	Disgust	0.65	1.95	0	36.22
	Fear	1.11	1.89	0	29.00
	Joy	4.53	3.49	0	100.00
	Anticipation	0.95	1.69	0	33.00
	Surprise	1.60	1.85	0	50.00
	Trust	4.11	3.21	0	55.12
CV	Total votes	8.78	49.56	1	1,211
	Rating	4.32	1.36	1	5
	$Rating^2$	17.89	7.81	1	25
	Length	82.41	115.21	1	2,435

Notes: DV=Dependent variable; IV=Independent variables; CV=Control variables.

The results of the Tobit regression analysis are presented in Tab. 4. The findings show that dependent variables are well explained by the independent variables (84.50%), indicates a good fit, with a highly significant likelihood ratio (p < 0.001), and the pseudo $R^2 = 0.0845$ are significant and within an acceptable range [55].

Tab. 4 presents the Tobit regression results for our model for the full sample set. The control variables included in the study analysis demonstrated significant impacts on RH. With positive coefficients, rating (β = 0.145, p < 0.001) and length (β = 0.001, p < 0.001), and had positive influence on RH. In contrast, negative effects were found between both total votes (β = –0.003, p < 0.001), $rating^2$ (β = –0.011, p < 0.05), and RH. These findings are in line with previous research [6,44,46].

Table 3: Variable correlations

	1	2	3	4	5	6	7	8	9	10	11	12	13
RH	1.00												
TV	−0.04*	1.00											
Rating	0.05*	0.03*	1.00										
Rating^2	0.23*	−0.16*	−0.04*	1.00									
Length	0.22*	0.03*	0.05*	0.19*	1.00								
Anger	0.04*	0.00	0.00	−0.03*	−0.02*	1.00							
Sadness	−0.06*	0.04*	−0.07*	−0.19*	−0.05	−0.03*	1.00						
Disgust	0.04*	−0.17*	0.03*	0.02*	−0.04*	−0.03*	0.02*	1.00					
Fear	0.02*	−0.06*	0.03*	−0.05*	−0.03*	−0.04*	0.02*	0.49*	1.00				
Joy	−0.07*	0.16*	0.14*	−0.34*	−0.07*	0.04*	−0.0*	−0.05*	0.51*	1.00			
Anc.	0.03*	−0.15*	0.02*	−0.03*	−0.04*	−0.03*	0.02*	0.48*	−0.02*	0.48*	1.00		
Surprise	−0.02*	0.00	0.02*	−0.16*	−0.04*	0.00	−0.02*	0.10*	0.55*	0.03*	0.08*	1.00	
Trust	−0.06*	0.08*	0.13*	−0.30*	−0.07*	0.04*	−0.07*	0.00	0.47*	0.00	−0.02*	0.71*	1.00

Notes: 1. RH = Review helpfulness; TV = Total votes; Anc. = Anticipation 2. n = 89,134. "*" represent the significance level at $p < 0.01$ level.

Table 4: Result of the tobit regression for full sample ($N = 89,134$)

Constant	Cofficient	Std.Error	t-value	Significance
Total votes	−0.003	0.001	−6.134	0.000***
Rating	0.145	0.034	4.142	0.000***
Rating^2	−0.011	0.005	−3.478	0.011*
Length	0.001	0.000	38.45	0.000***
Anger	−0.024	0.0041	−10.65	0.031*
Sadness	−0.015	0.004	−1.457	0.021*
Disgust	0.0271	0.0047	3.41	0.022*
Fear	0.0250	0.007	5.107	0.000***
Joy	−0.040	0.0031	−17.06	0.000***
Anticipation	0.0033	0.0033	2.301	0.0718
Surprise	0.0070	0.0039	3.713	0.064
Trust	−0.012	0.0029	−4.12	0.000***
Log likelihood	−8,764.1			
Pseudo R^2	0.0845			

Note: * p < 0.05, ** p < 0.01, *** p < 0.001

Regarding the direct effects of discrete emotions intensity on RH, six out of eight emotional dimensions were found to significantly influence the RH. Emotions including disgust (H3: β = 0.0271, p < 0.05) and fear (H4: β = 0.0250, p < 0.001) positively influenced RH, while anger (H1: β = −0.024, p < 0.05), joy (H5: β = −0.040, p < 0.001), sadness (H2: β = −0.015, p < 0.05), and trust (H8: β = −0.012, p < 0.001) had negative impact on RH. These results support all hypotheses H1, H3, H4, H5, and H2, except H8.

Finally, the two emotional dimensions, anticipation (H6: $\beta = 0.0033$, $p > 0.05$) and surprise (H7: $\beta = 0.0070$, $p > 0.05$) did not show significant effects on RH at $p < 0.01$, and thus H2 and H7 are not supported.

Simultaneously, the current study also carried out the robustness test using linear regression. Due to the fact that helpfulness was not useful at normal distribution and the lowest value was "0" in a modern, linear regression variable during the robustness control, the current study took the log value (helpfulness + 1). The findings were in accordance with the results of the Tobit regression analysis.

With regard to the moderating role of disease severity, the influence of discrete emotions is measured by separating the data into two sub-samples, high and low-disease severity. As reported in Tab. 5, path coefficients, standard error, and t-statistics were calculated across high and low-disease severity samples. As observed from the analysis, the influence of $Anger_{Hi, Lo}$ ($\beta = -0.027$, $t = -4.103$, $p < 0.01$; $\beta = -0.018$, $t = -2.223$, $p < 0.05$), upon RH has a greater negative effects for high-disease severity than low-disease severity. In contrast, the influence of $Disgust_{Hi, Lo}$ ($\beta = 0.029$, $t = 4.041$, $p < 0.01$; $\beta = 0.026$, $t = 4.340$, $p < 0.01$), $Joy_{Hi, Lo}$ ($\beta = 0.071$, $t = 15.65$, $p < 0.01$; $\beta = 0.062$, $t = 17.87$, $p < 0.01$), and $Trust_{Hi, Lo}$ ($\beta = 0.032$, $t = 3.473$, $p < 0.01$; $\beta = 0.023$, $t = 3.765$, $p < 0.01$) on RH has a greater positive effects for high-disease severity than low-disease severity. Finally, the influence of $Sadness_{Hi, Lo}$ ($\beta = -0.012$, $t = -2.914$, $p > 0.05$; $\beta = -0.013$, $t = -2.105$, $p > 0.05$), $Anticipation_{Hi, Lo}$ ($\beta = 0.004$, $t = 2.453$, $p > 0.05$; $\beta = 0.005$, $t = 2.117$, $p > 0.05$), $Fear_{Hi, Lo}$ ($\beta = 0.024$, $t = 4.003$, $p > 0.05$; $\beta = 0.026$, $t = 4.439$, $p > 0.05$), and $Surprise_{Hi, Lo}$ ($\beta = 0.008$, $t = 4.254$, $p > 0.05$; $\beta = 0.006$, $t = 4.576$, $p > 0.05$) on RH do not differ across high-disease and low-disease severity samples. These results provide clear evidence to support H9$_a$, H9$_c$, H10$_a$, and H10$_d$. For the control variables, the coefficients of total votes, rating, rating^2, and length are significant for both high and low-disease severity samples. Comparing the current study findings with the results of Mudambi and Schuff [18], the extreme reviews proved less helpful than mild reviews (i.e., β value of rating^2 is –0.023). Tab. 5 presents the research model analysis summary for the moderating effects.

Table 5: Result of the tobit regression for high and low-disease severity

Variables	High-disease severity			Low-disease severity		
	Coficient	Std. error	t-value	Coficient	Std. error	t-value
Constant	0.457***	0.029	19.451	0.543**	0.039	16.021
Total votes	0.001*	0.002	–6.134	–0.004*	0.001	–9.145
Rating	0.178**	0.036	4.233	0.081**	0.029	2.149
Rating^2	–0.023*	0.004	–3.467	0.002*	0.006	0.023
Length	–0.005**	0.001	35.42	0.001**	0.000	–3.576
Anger	–0.027**	0.008	–4.103	–0.018**	0.005	–2.223
Sadness	–0.012	0.005	–2.914	–0.013	0.008	–2.105
Disgust	0.029**	0.0041	4.041	0.026**	0.0049	4.340
Fear	0.024	0.004	4.003	0.026	0.005	4.439
Joy	0.071**	0.0029	15.65	0.062**	0.0034	17.87
Anticipation	0.004	0.002	2.453	0.005	0.001	2.117
Surprise	0.008	0.0043	4.254	0.006	0.0055	4.576
Trust	0.032**	0.0024	3.473	0.023**	0.0025	3.765
Log likelihood			–1756.855	Log likelihood		–4126.812
Pseudo R^2			0.091	Pseudo R^2		0.087

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

5 Discussion

This paper creatively examined how the eight dimensions of emotional intensity influence physician RH. Following the findings of Shah, Yan [7], this study also investigated the moderating role of disease severity in the relationship between discrete emotions and physician RH. Drawing on Plutchik's wheel on human emotions framework, research hypotheses were tested based on the data sets included 89,134 OPRs from two U.S-based PRWs. Anger, joy, and sadness incorporated in reviews are found to have a negative impact on RH. According to a couple of recent research studies [6,25], anger, joy, and sadness-related emotional content in online reviews revealed greater negative effects on RH. In contrast, disgust, fear, and trust incorporated in reviews positively influence RH. This result confirms the negative bias effect found in RH studies [6,24,25,27]. However, there was no relationship established between anticipation or surprise incorporated in reviews and RH [25]. The reason behind this result was because anticipation may merely express the personal wish or will of a reviewer who does not provide useful information for readers. In the same way, surprise defines the psychological state of the sudden event. Surprise may not deliver consistent emotions to the readers. Hence, surprise may not reveal a reader's exact attitude, which does not provide extensive support for the judgment of reviewers [25].

The current research also produces some important results regarding the patients' disease severity. Disease severity moderated the influence of anger, disgust, joy, surprise, and trust on perceived RH. Anger has been found to be more negatively impacted on perceived RH for high-disease severity than for low-disease severity. In contrast, greater positive associations are found between disgust, joy, and trust and perceived RH for high-disease severity than the low-disease severity. The significant moderating effect of patients' disease severity on emotional content showed that emotional information could provide useful information under a high information overload and making a greater effect on RH for different environments [6,26,27]. These findings were not only interesting, but also intuitively credible.

5.1 Theoretical Implications

This study contributes to examine the importance of discrete emotions intensity for their effects on RH, lead to some valuable implications as follows.

First, this study is one of the first to explore the role of emotional content on RH in a healthcare context. Due to the credence nature of the healthcare industry, emotional content is particularly vital for patients. However, extent studies on the emotional content embedded in online reviews have been performed regarding tangible goods. To the best of the authors' information, only two studies have been conducted on RH in the healthcare context [35,36]. However, these studies did not consider the discrete emotional content in predicting RH. The current research contributes to the literature by identifying the distinct roles of positive and negative emotional content on RH in OPRs context.

Second, extant studies regarding online RH primarily focused on quantitative reviews (e.g., review depth, word count) [18] or review features (e.g., reviewer expertise and reputation) [56]. These studies rarely explored the textual content contained in a review. The current study helps to fill this knowledge gap by mining review text that utilizes the cognitive characteristics of patients (emotional content) such as discrete positive and negative emotions in OPRs and explores their effect on the perceived RH. In addition, the prediction of the helpfulness of OPRs is typically a dynamic topic. Therefore, it is the first study that focus on innovative approaches such as the mixed-methods approach (text mining technology and econometric analysis) and real data from different investigation sources (i.e., RateMDs, Healthgrades, and state medical boards data) instead of considering only a single method and source because users search information from different platforms before making their health consultation decisions [57]. Keeping in view of social media, analyzing data from a single silver-bullet metric is not worthy. This research follows Ambler and Roberts [58] and employs more than one social media metric such as

structured and unstructured data (e.g., ratings, online reviews, and helpful votes) and links these social media components with the prediction of helpfulness of OPRs.

Last but not least, to the best of authors' knowledge this study is first to examine the impact of patient's disease severity in predicting the helpfulness of OPRs. Early investigations have looked from multiple viewpoints at the moderating impact of patient's disease severity [28,30,33]. However, the moderating role of patient's disease severity on the association between the discrete emotions and OPRs helpfulness remains to be established. Our research offers the theoretical insights that disease severity moderates the association between anger, distrust, joy, and trust and OPRs helpfulness, such that the influence is stronger for high-disease severity than the low-disease severity.

5.2 Practical Implications

There are some practical implications for the findings of this study. Platform administrators could integrate the results of the current study into the review implementation guidelines for future reviewers. For instance, reviewers may be encouraged to articulate the exact affective component of the healthcare service provider and his/her quality of services instead common positive/negative attitudes. Although reviewers remain impartial and competent, the quality of communication in reviews may be enhanced by using such writing techniques. For example, if general feelings (e.g., bad or poor) toward service deficiencies were not lightly articulated, reviewers would be encouraged to portray realistic encounters with particular dimensions of emotion (e.g., fear or disgust).

Health service providers can also learn from the findings of this research in the sense that, particular emotion dimensions on review platforms (e.g., RateMDs and Healthgrades) may enable health care service providers to recognize immediately "whether the customers are pleased, unhappy, untrustful and/ or angry with their services or specific features" [53]. Patients who are unhappy with their treatment outcome tend to deal with the stimulus, while patients who feel fear during the treatment process; they are more likely to be passive and retrogressive. As a result, the current research of emotional information from reviews on both of these platforms would provide hospitals with an "emotion recognition system," which could define affective factors, investigate the causes, and perform service recovery actions.

From a business-operation context, given that the patient's disease severity moderates the effect of discrete positive and negative emotions incorporated in the reviews on the OPRs helpfulness, it is urged to use both emotion-related words (disgust, surprise and trust) for high-disease severity than for low-disease severity when writing a review. However, it is recommended to use less anger-related emotional words for high-disease severity than for low-disease severity while posting a review. Health care providers are advised to implement specific procedures to manage reviews of varying emotional trends to restore patients' confidence in the health care services. The framework of the current paper is applicable to various online health rating platforms (e.g., Vitals, Yelp, and Iwantgreatcare).

5.3 Limitations and Future Research

There are several limitations of the current study that may support future studies to advance our understanding further. First, this research operated on the basis of distinct emotions identified using a text mining approach, which cannot highlight continuous shifts in particular dimensions of emotional evaluation, such as valence, enthusiasm, and control [46]. Future research may further test the effects of particular appraisal aspects of emotional content on RH. Second, the voting activity on the RH is evidence of patient engagement. Hence, data from helpfulness votes obtained on health ratings sites may not necessarily warrant a review evaluation carried out by patients properly engaged in health care service (e.g., at the initial treatment consultation stage) or with the review (e.g., have mild response toward helpfulness/unhelpfulness). This aspect must be further investigated in a laboratory setting. The

findings of the current study contribute to the roles of specific emotions to cope the information overload challenge in the healthcare field and highlight the future research directions in the given domain.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1]   A. Benetoli, T. F. Chen and P. Aslani, "Consumer perceptions of using social media for health purposes: Benefits and drawbacks," *Health Informatics Journal*, vol. 25, no. 4, pp. 1661–1674, 2019.

[2]   P. J. Schulz and F. Rothenfluh, "Influence of health literacy on effects of patient rating websites: Survey study using a hypothetical situation and fictitious doctors," *Journal of Medical Internet Research*, vol. 22, no. 4, pp. e14134, 2020.

[3]   Y. A. Hong, C. Liang, T. A. Radcliff, L. T. Wigfall and R. L. Street, "What do patients say about doctors online? A systematic review of studies on patient online reviews," *Journal of Medical Internet Research*, vol. 21, no. 4, pp. e12521, 2019.

[4]   Y. Lin, Y. A. Hong, B. S. Henson, R. D. Stevenson, S. Hong *et al.*, "Assessing patient experience and healthcare quality of dental care using patient online reviews in the United States: Mixed methods study," *Journal of Medical Internet Research*, vol. 22, no. 7, pp. e18652, 2020.

[5]   H. Hao, K. Zhang, W. Wang and G. Gao, "A tale of two countries: International comparison of online doctor reviews between China and the United States," *International Journal of Medical Informatics*, vol. 99, pp. 37–44, 2017.

[6]   G. Ren and T. Hong, "Examining the relationship between specific negative emotions and the perceived helpfulness of online reviews," *Information Processing & Management*, vol. 56, no. 4, pp. 1425–1438, 2019.

[7]   A. M. Shah, X. Yan, S. Khan and S. J. Shah, "Exploring the impact of review and service-related signals on online physician review helpfulness: A multi-methods approach," in *Proc. Twenty-Fourth Pacific Asia Conf. on Information Systems*, Dubai, UAE, pp. 1–14, 2020.

[8]   A. Qazi, K. B. Shah Syed, R. G. Raj, E. Cambria, M. Tahir *et al.*, "A concept-level approach to the analysis of online review helpfulness," *Computers in Human Behavior*, vol. 58, pp. 75–81, 2016.

[9]   Y. Zhang and Z. Lin, "Predicting the helpfulness of online product reviews: A multilingual approach," *Electronic Commerce Research and Applications*, vol. 27, pp. 1–10, 2018.

[10] Y. H. Hu and K. Chen, "Predicting hotel review helpfulness: The impact of review visibility, and interaction between hotel stars and review ratings," *International Journal of Information Management*, vol. 36, no. 6, Part A, pp. 929–944, 2016.

[11] B. Fang, Q. Ye, D. Kucukusta and R. Law, "Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics," *Tourism Management*, vol. 52, pp. 498–506, 2016.

[12] A. Y. K. Chua and S. Banerjee, "Helpfulness of user-generated reviews as a function of review sentiment, product type and information quality," *Computers in Human Behavior*, vol. 54, pp. 547–554, 2016.

[13] M. Antioco and K. Coussement, "Misreading of consumer dissatisfaction in online product reviews: Writing style as a cause for bias," *International Journal of Information Management*, vol. 38, no. 1, pp. 301–310, 2018.

[14] Z. Xiang, Q. Du, Y. Ma and W. Fan, "A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism," *Tourism Management*, vol. 58, pp. 51–65, 2017.

[15] C. Zhao, "Are negative sentiments" negative" for review helpfulness?," in *Proc. Twenty-Fourth Pacific Asia Conf. on Information Systems*, Dubai, UAE, pp. 1–14, 2020.

[16] B. Gao, N. Hu and I. Bose, "Follow the herd or be myself? an analysis of consistency in behavior of reviewers and helpfulness of their reviews," *Decision Support Systems*, vol. 95, pp. 1–11, 2017.

[17] R. Filieri, F. McLeay, B. Tsui and Z. Lin, "Consumer perceptions of information helpfulness and determinants of purchase intention in online consumer reviews of services," *Information & Management*, vol. 55, no. 8, pp. 956–970, 2018.

[18] S. M. Mudambi and D. Schuff, "What makes a helpful online review? A study of customer reviews on Amazon. com," *MIS Quarterly*, vol. 34, no. 1, pp. 185–200, 2010.

[19] M. Salehan and D. J. Kim, "Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics," *Decision Support Systems*, vol. 81, pp. 30–40, 2016.

[20] P. J. Lee, Y. H. Hu and K. T. Lu, "Assessing the helpfulness of online hotel reviews: A classification-based approach," *Telematics and Informatics*, vol. 35, no. 2, pp. 436–445, 2018.

[21] M. Lee, M. Jeong and J. Lee, "Roles of negative emotions in customers' perceived helpfulness of hotel reviews on a user-generated review website: A text mining approach," *International Journal of Contemporary Hospitality Management*, vol. 29, no. 2, pp. 762–783, 2017.

[22] G. Craciun, W. Zhou and Z. Shan, "Discrete emotions effects on electronic word-of-mouth helpfulness: The moderating role of reviewer gender and contextual emotional tone," *Decision Support Systems*, vol. 130, pp. 113226, 2020.

[23] A. Felbermayr and A. Nanopoulos, "The role of emotions for the perceived usefulness in online customer reviews," *Journal of Interactive Marketing*, vol. 36, pp. 60–76, 2016.

[24] M. S. I. Malik and A. Hussain, "Helpfulness of product reviews as a function of discrete positive and negative emotions," *Computers in Human Behavior*, vol. 73, pp. 290–302, 2017.

[25] X. Wang, L. Tang and E. Kim, "More than words: Do emotional content and linguistic style matching matter on restaurant review helpfulness?," *International Journal of Hospitality Management*, vol. 77, pp. 438–447, 2019.

[26] H. Li, H. Liu and Z. Zhang, "Online persuasion of review emotional intensity: A text mining analysis of restaurant reviews," *International Journal of Hospitality Management*, vol. 89, pp. 102558, 2020.

[27] M. J. Chen and C. K. Farn, "Examining the influence of emotional expressions in online consumer reviews on perceived helpfulness," *Information Processing & Management*, vol. 57, no. 6, pp. 102266, 2020.

[28] N. Lu and H. Wu, "Exploring the impact of word-of-mouth about physicians' service quality on patient choice based on online health communities," *BMC Medical Informatics and Decision Making*, vol. 16, no. 1, pp. 151, 2016.

[29] X. Cao, Y. Liu, Z. Zhu, J. Hu and X. Chen, "Online selection of a physician by patients: Empirical study from elaboration likelihood perspective," *Computers in Human Behavior*, vol. 73, pp. 403–412, 2017.

[30] H. Yang, X. Guo and T. Wu, "Exploring the influence of the online physician service delivery process on patient satisfaction," *Decision Support Systems*, vol. 78, pp. 113–121, 2015.

[31] H. J. Jeong and D. M. Koo, "Combined effects of valence and attributes of e-wom on consumer judgment for message and product: The moderating effect of brand community type," *Internet Research*, vol. 25, no. 1, pp. 2–29, 2015.

[32] R. Filieri, "What makes an online consumer review trustworthy?," *Annals of Tourism Research*, vol. 58, pp. 46–64, 2016.

[33] A. M. Shah, X. Yan, S. A. A. Shah, S. J. Shah and G. Mamirkulova, "Exploring the impact of online information signals in leveraging the economic returns of physicians," *Journal of Biomedical Informatics*, vol. 98, pp. 103272, 2019.

[34] H. B. Abdalla, A. M. Ahmed and M. A. Al Sibahee, "Optimization driven mapreduce framework for indexing and retrieval of big data," *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 14, no. 5, pp. 1886–1908, 2020.

[35] M. Lockie, M. K. J. Waiguny and S. Grabner-Kräuter, "How style, information depth and textual characteristics influence the usefulness of general practitioners' reviews," *Australasian Marketing Journal (AMJ)*, vol. 23, no. 3, pp. 168–178, 2015.

[36] N. Alodadi and L. Zhou, "Predicting the helpfulness of online physician reviews," in *Proc. IEEE Int. Conf. on Healthcare Informatics*, Chicago, IL, pp. 1–6, 2016.

[37] P. Ekman, "An argument for basic emotions," *Cognition and Emotion*, vol. 6, no. 3–4, pp. 169–200, 1992.

[38] R. Plutchik, "A psychoevolutionary theory of emotions," *Social Science Information*, vol. 21, no. 4–5, pp. 529–553, 1982.

[39] A. Mehrabian and J. A. Russell, "*An Approach to Environmental Psychology,*" Cambridge, MA, US, The MIT Press, pp. xii, 266–xii, 266, 1974.

[40] J. Berger and K. L. Milkman, "What makes online content viral?," *Journal of Marketing Research*, vol. 49, no. 2, pp. 192–205, 2012.

[41] V. Francisco and P. Gervás, "Automated mark up of affective information in English texts," in *Proc. Int. Conf. on Text, Speech and Dialogue*, Brno, Czech Republic, pp. 375–382, 2006.

[42] R. Plutchik, "*The Psychology and Biology of Emotion,*" New York, NY, US, HarperCollins College Publishers, pp. xx, 396–xx, 396, 1994.

[43] J. S. Lerner and D. Keltner, "Fear, anger, and risk," *Journal of Personality and Social Psychology*, vol. 81, no. 1, pp. 146–159, 2001.

[44] S. N. Ahmad and M. Laroche, "How do expressed emotions affect the helpfulness of a product review? Evidence from reviews using latent semantic analysis," *International Journal of Electronic Commerce*, vol. 20, no. 1, pp. 76–111, 2015.

[45] L. Z. Tiedens and S. Linton, "Judgment under emotional certainty and uncertainty: The effects of specific emotions on information processing," *Journal of Personality and Social Psychology*, vol. 81, no. 6, pp. 973–988, 2001.

[46] D. Yin, S. D. Bond and H. Zhang, "Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews," *MIS Quarterly*, vol. 38, no. 2, pp. 539–560, 2014.

[47] M. B. Tannenbaum, J. Hepler, R. S. Zimmerman, L. Saul, S. Jacobs *et al.*, "Appealing to fear: A meta-analysis of fear appeal effectiveness and theories," *Psychological Bulletin*, vol. 141, no. 6, pp. 1178–1204, 2015.

[48] A. Vaish, T. Grossmann and A. Woodward, "Not all emotions are created equal: The negativity bias in social-emotional development," *Psychological Bulletin*, vol. 134, no. 3, pp. 383–403, 2008.

[49] S. Lee and J. Y. Choeh, "The determinants of helpfulness of online reviews," *Behaviour & Information Technology*, vol. 35, no. 10, pp. 853–863, 2016.

[50] A. M. Shah, X. Yan, S. J. Shah and S. Khan, "Predicting the factors that influence helpfulness of online physician reviews," in *Proc. 12th China Summer Workshop on Information Management*, Qingdao, China, pp. 344–349, 2018.

[51] Statista "Top 10 causes of death in the U.S. in 2018," 2018 [cited 2020 August 12], Available from: https://www.statista.com/statistics/235703/major-causes-of-death-in-the-us/.

[52] R. BliegeBird and E. Smith, "Signaling theory, strategic interaction, and symbolic capital," *Current Anthropology*, vol. 46, no. 2, pp. 221–248, 2005.

[53] S. M. Mohammad and P. D. Turney, NRC emotion lexiconx, in *NRC Technical Report*, 2013. S. M. Mohammad and P. D. Turney (2013) NRC emotion lexicon*, in *NRC Technical Report 2*.

[54] J. Hair, W. Black, B. Babin and R. Anderson, "*Multivariate Data Analysis: A Global Perspective,*" Upper Saddle River, NJ, Pearson Prentice Hall, pp. 816, 2010.

[55] M. R. Veall and K. F. Zimmermann, "Pseudo-r2 measures for some common limited dependent variable models," *Journal of Economic Surveys*, vol. 10, no. 3, pp. 241–259, 1996.

[56] M. Siering, J. Muntermann and B. Rajagopalan, "Explaining and predicting online review helpfulness: The role of content and reviewer-related signals," *Decision Support Systems*, vol. 108, pp. 1–12, 2018.

[57] Q. A. Bui, W. B. Lee, J. S. Lee, H. L. Wu and J. Y. Liu, "Biometric-based key management for satisfying patient's control over health information in the hipaa regulations," *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 14, no. 1, pp. 437–454, 2020.

[58] T. Ambler and J. Roberts, "Beware the silver metric: Marketing performance measurement has to be multidimensional," *Marketing Science Institute, Report*, vol. 6, no. 113, pp. 6–113, 2006.