The following publication Li, H., Ji, H., Liu, H., Cai, D., & Gao, H. (2022). Is a picture worth a thousand words? Understanding the role of review photo sentiment and text-photo sentiment disparity using deep learning algorithms. Tourism Management, 92, 104559 is available at https://doi.org/10.1016/j.tourman.2022.104559.

Is a picture worth a thousand words? The impacts of review photo sentiment using deep learning algorithms

Hengyun Li, Ph.D.
School of Hotel and Tourism Management,
The Hong Kong Polytechnic University,
Hong Kong SAR, China
Email: neilhengyun.li@polyu.edu.hk

Haipeng Ji*
School of Hotel and Tourism Management,
The Hong Kong Polytechnic University,
Hong Kong SAR, China
Email: haipeng.ji@polyu.edu.hk
Telephone: +852 3400-2243; Fax: +852 2362-9362
* Corresponding Author

Hongbo Liu, Ph.D.
School of Hospitality & Tourism Management,
University of Surrey,
Guildford, Surrey, GU2 7XH, UK
E-mail: hongbo.liu@surrey.ac.uk

Danting Cai
School of Hotel and Tourism Management,
The Hong Kong Polytechnic University,
Hong Kong SAR, China
Email: danting.cai@connect.polyu.hk

Huicai Gao
School of Hotel and Tourism Management,
The Hong Kong Polytechnic University,
Hong Kong SAR, China
Email: huicai.gao@connect.polyu.hk

Declaration of competing interest

None

Acknowledgments

- The authors acknowledge the support of research funds from the National Natural Science Foundation of China (71902169) and The Hong Kong Polytechnic University Departmental General Research Fund (Project No. G-UALR).

This is an Accepted Manuscript of an article published by Elsevier in Tourism Management in 2022. Available online: https://doi.org/10.1016/j.tourman.2022.104559

Is a picture worth a thousand words? Understanding the role of review photo sentiment and photo-text sentiment disparity using deep learning algorithms

Abstract: Images have become integral to consumers' sharing of consumption experiences due to their abilities of carrying rich and vivid information. This study investigates the impacts of restaurant review photo sentiment on customers' perceived review usefulness and enjoyment using deep learning and econometric model analysis. The results indicate that (1) reviews with photos are more useful and enjoyable than reviews without photos; (2) a U-shaped relationship exists between review photo sentiment and review usefulness, with the effect of review photo sentiment on review enjoyment being positive and linear. Moreover, the effects can be strengthened by the number of review photos while weakened by the text-photo sentiment disparity. The above findings are reinforced by a sample of restaurant online reviews written by tourists in Las Vegas. This study contributes to the electronic word-of-mouth literature as well as to the application of machine learning technologies in computer vision to tourism and hospitality research.

Keywords: Photo sentiment; text–photo sentiment disparity; photo number; review usefulness; review enjoyment; deep learning

1. Introduction

Studies have demonstrated that user-generated content (UGC) contributes substantially to ongoing growth in tourism consumption, including destinations, hotels, and restaurants. UGC plays a particularly vital role in reducing search efforts and mitigating risk, given the myriad uncertainties of tourism and hospitality products (Bigne et al., 2020; Lu et al., 2020; Ray & Bala, 2021; Zhang et al., 2020). According to Stackla (2021a, 2021b), 92% of consumers trust UGC, 83% of travelers derive travel inspiration from such content, and 79% of people perceive UGC as offering a closer look at products to inform purchase decisions. Among an array of user-generated information, visual content (e.g., photos) can easily capture users' attention thanks to its vividness and visual appeal (Aydin, 2020; Bigne et al., 2020; Geise & Baden, 2015; Guan et al., 2019). For example, Facebook posts containing photos receive 37% more engagement than those featuring only text (Bullas, 2017). McGrath (2017) further contended that the human brain processes visual content more quickly than text. This distinction makes photos even more important in today's age of information overload, when consumers can become overwhelmed by the volume of sensory input available on social media.

In the past few years, advances in technology and mobile devices have boosted the number of user-generated photos (UGPs). Visual content has become an integral part of consumers' experience sharing; more than 576 million travel-related photos have been posted on Instagram to date (Instagram, 2021; Stackla, 2021a). Images provide greater informational, aesthetic, and self-enhancement value than textual content (Li & Xie, 2020). When customers include visual content in their online reviews, the authenticity and trustworthiness of these reviews improve. Customer reviews featuring photos also tend to be considered more helpful than text-only reviews (Ma et al., 2018): nearly two-thirds (65%) of consumers are more likely to trust products whose customer reviews contain photos or videos (PowerReviews, 2017). Given the importance of UGPs, it is necessary to understand the role of visual content included in reviews.

Recent years have witnessed an expanding but limited number of studies on UGPs in marketing, tourism, and hospitality. Most research on online reviews or UGCs has focused on textual content despite the key role of visual content. This disparity may be partly due to the technical and methodological challenges of analyzing scalable visual content (An et al., 2020; Ma et al., 2018). Among the relatively thin body of work on UGPs, Oliveira and Casais (2019) explored the effects of UGPs' content characteristics (i.e., photos showcasing food or a restaurant's physical environment) on consumers' restaurant selection; Li and Xie (2020) explored the effects of mere image presence and image characteristics (i.e., colorfulness, source, image quality, human faces, and facial emotions) on consumer engagement; and Cai and Chi (2020) and Marder et al. (2021) investigated the impacts of posted photos' aesthetics and color on consumers' willingness to dine/buy. However, scholars have largely ignored photo sentiment and its effects, which are essential to understanding consumers' insight, emotions, and expectations (Bigne et al., 2020; Chua & Banerjee, 2016; Geetha et al., 2017; Siering et al., 2018)—especially for experiential products in tourism and hospitality. In addition, although many researchers have analyzed review text sentiment to uncover customers' opinions and emotions (Alaei et al., 2019; Li et al., 2019; Ma et al., 2018), none appear to have investigated review text–photo sentiment disparity (i.e., the interaction or congruence between review text and review photo) and its impacts. Furthermore, most related work has pertained to the utilitarian value of online reviews, often review usefulness as evidenced by "useful" or "helpful" votes (An et al., 2020; Cheng & Ho, 2015; Lee, 2018; Ma et al., 2018). Few studies have assessed the hedonic value of reviews (Ham et al., 2019) as evaluated via "enjoyment" votes (Liu & Park, 2015).

To address the above research gaps, this study aims to investigate the impacts of UGPs in online reviews with the evidence from Las Vegas restaurants. Both the entire restaurant online review sample and the tourist online review sample (i.e., restaurant reviews written by tourists) were used for analysis. First, we examine whether reviews with photo(s) contribute to review usefulness and enjoyment. Second, using a deep learning algorithm, we calculate review photo sentiment and explore its impacts on review usefulness and enjoyment. Lastly, the boundary effects of review photo sentiment are tested by considering the moderating effects of the number of photos in a review (i.e., the richness of reviews' visual content) and review text–photo sentiment disparity. This study contributes to the literature in several ways. First, it represents an initial attempt to reveal the role of review photo sentiment by adopting a deep learning method and econometric modeling with big data. Second, this study enriches research related to UGC by revealing the utilitarian and hedonic value of UGPs. Third, our study expands the literature by uncovering the boundary effects of review photo sentiment with two moderators, namely the number of photos and review text–photo sentiment disparity. Fourth, this study is one of the first to investigate textual and visual content interaction in online reviews.

2. Literature Review

Consumer-generated photos play a prominent role in consumers' judgements and decisions related to online reviews. A large volume of online reviews can lead to information overload and impede information processing. The elaboration likelihood model (Petty & Cacioppo, 1986) asserts that a large amount of information causes consumers to rely on peripheral cues, such as images and affect, to facilitate information processing and decisions (Li et al., 2020). Photos are more vivid, attention-grabbing, and easier to comprehend (Aydin, 2020; Guan et al., 2019). This type of information can potentially serve as peripheral cues to reduce the complexity of information processing. Despite this important role, UGPs have been largely overlooked within the literature on online reviews. The advancement of image analytic techniques such as deep learning has sparked a growing body of research on UGPs in online reviews. Scholars have reported that UGPs generally improve online review helpfulness (An et al., 2020; Cheng & Ho, 2015; Lee, 2018; Ma et al., 2018). This effect is more pronounced for negative online reviews, lower-priced hotels, and when photos convey content relevant to review text (An et al., 2020; Li et al., 2021). However, contradictory findings regarding the roles of review photos on review helpfulness have also been reported. Hlee et al. (2021) documented a negative relationship between the number of review photos and restaurant online review helpfulness, and Lee (2018) failed to find a significant relationship between the number of review photos and review helpfulness in restaurants. More empirical research is needed to disentangle these conflicting results. Other pictorial features (e.g., photo sentiment) and moderators, such as the congruence between review text sentiment and photo sentiment, should be considered to fully understand UGPs' influence on review helpfulness.

Denoted by online reviews' number of "helpful" votes, review helpfulness represents online reviews' overall quality and informativeness (Lee & Choeh, 2020; Ma et al., 2018). It helps build trust towards the online reviews and facilitates information processing, especially when the information load is high (Choi & Leon, 2020). The number of "helpful" votes can help consumers quickly identify the most trustworthy and useful information among the large volume of online reviews, and consumers then rely on such reviews to make purchase decisions. Therefore, review helpfulness promotes adoption of the information and suggests the persuasion effect of online reviews (Li et al., 2020). Many empirical studies provided support for the persuasion effects of review helpfulness. For example, Lopez and Garza (2021) and Filieri et al. (2018) found that review helpfulness positively predicts product purchase intention. Lee and Choeh (2018) and Topaloglu and Dass (2021) revealed that when the online reviews are deemed helpful, online review features such as number of reviews, review length and review content have a stronger influence on box office sales.

Online reviews have both utilitarian (information) value and hedonic value (Ham et al., 2019). Hedonic value focuses on an object's "fun" and "playfulness" (Babin et al., 1994). Along with utilitarian value, it is a fundamental dimension of overall perceived value (Babin et al., 1994; Sánchez-Fernández & Iniesta-Bonillo, 2007). To capture the hedonic value of online reviews, Yelp.com released a feature enabling users to assign reviews "funny" and "cool" votes. These two aspects reflect consumers' perceived enjoyment when reading online reviews (Liu & Park, 2015). Perceived enjoyment plays a core mediating role between stimuli and behavioral intentions (So et al., 2020). Perceived enjoyment of online reviews could enhance the adoption of online reviews, and the role of review enjoyment can be more prominent for hedonic products like restaurants. In information search, consumers not only have functional needs (i.e. seeking helpful information), but also have hedonic needs, like need for enjoyment, entertainment and novelty (Vogt & Fesenmaier, 1998). The hedonic aspects of information hence also play an important role in information adoption, and even a more important role for consumers with hedonic needs (Hlee et al., 2019; Park & Nicolau, 2015). Therefore, review enjoyment is of critical importance for online review adoption and purchase decisions in restaurant context. Additionally, review enjoyment helps increase consumer engagement with online reviews, thereby enhancing online reviews' adoption and persuasiveness (Park & Nicolau, 2015).

Thus far, most studies have focused on the utilitarian value of online reviews (i.e., review helpfulness); only a few have explored the hedonic value of online reviews (i.e., review enjoyment) (Li et al., 2019). Related research has shown that review enjoyment is positively related to review helpfulness (Liu & Park, 2015; Yang et al., 2017) and can be predicted by review length, review rating, the number of photos, reviewer expertise, and the explanatory and sensory cues in review text (Hlee et al., 2019; Li et al., 2019; Park & Nicolau, 2015; Yang et al., 2017). The number of photos has been identified as a significant predictor of review enjoyment (Yang et al., 2017). Yet the effects of UGPs and other pictorial features, such as photo sentiment and the interaction between review text and photos, have not been fully addressed. The present study hence explores the impacts of review photo sentiment on review helpfulness and review enjoyment and how these effects depend on the number of review photos and the congruence between review text sentiment and photo sentiment.

2.1 Presence effects of user-generated photos on review helpfulness and enjoyment

Including consumer-generated photos in online reviews improves review helpfulness in several ways. First, UGPs provide information to supplement review text, thus enhancing online reviews' information richness (Cheng & Ho, 2015). Second, according to the dual coding theory (Paivio, 1990), visual information is processed distinctively from verbal information. The former is more easily encoded and retrieved and is therefore simpler to comprehend than text. As such, images can facilitate the understanding of corresponding textual information (Glenberg & Langston, 1992). Review photos can thus improve one's understanding and interpretation of relevant review text (Li et al., 2021). Third, consumer-generated photos provide visual proof of textual content and promote online reviews' trustworthiness and persuasiveness (Filieri, 2016; Ma et al., 2018). Consumer-generated photos accordingly complement textual content in predicting review helpfulness (Ma et al., 2018). Several studies have provided empirical support for the positive effects of review photos on review helpfulness (An et al., 2020; Bigne et al., 2020; Cheng & Ho, 2015; Li et al., 2021; Ma et al., 2018; Yang et al., 2017).

Consumer-generated photos could also elevate fellow customers' enjoyment when reading online reviews. Images are naturally more attention-grabbing and engaging than text, as pictures are distinct and offer rich information via diverse content and sensory stimuli, including color, shape, and size (Li & Xie, 2020). Including photos in online reviews can make for more interesting reading versus text-only reviews. Additionally, sensory cues are commonly used in marketing to shape consumers' emotions, attitudes, and behavior (Krishna, 2012). Images, as sensory stimuli, can evoke affective responses and aesthetic experiences that are tied to enjoyment when reading online reviews (Hlee et al., 2019; Yang et al., 2017). Review photos also contain various sensory cues about the focal products and environment, such as lighting, color, shape, and texture; these cues reportedly contribute to review enjoyment (Li et al., 2019). The number of photos showcasing the physical environment, food, and beverages in restaurants' online reviews is positively associated with review enjoyment as well (Yang et al., 2017). In line with this discussion, we propose the following hypotheses:

H1a: *Online reviews containing photos have a positive effect on review usefulness.*

H1b: *Online reviews containing photos have a positive effect on review enjoyment.*

2.2 Impact of review photo sentiment on review helpfulness and enjoyment

Sentiment refers to individuals' opinions, attitudes, or judgements that are prompted by feelings (Fang & Zhan, 2015). Sentiment captures the valence of an opinion, falling along a continuum that ranges from negative to neutral to positive (Ma et al., 2018). Sentiment reflects subjective feelings rather than facts through sentimental words such as "great," "love," "bad," and "hate" (Li et al., 2018; Ma et al., 2018). Sentiment has been widely used to understand customers' opinions and emotions expressed in online review text (Bigne et al., 2020; Chua & Banerjee, 2016; Geetha et al., 2017; Kuan & Hui, 2015; Siering et al., 2018). Many studies on customer sentiment have focused on textual data, such as online review text and text-based posts or comments on social media (e.g., Twitter) (Alaei et al., 2019; Ma et al., 2018). However, visual information (e.g., photos and videos) also conveys opinions and sentiment (Campos et al., 2015; Deng & Li, 2018). In addition to text, consumers use photos to demonstrate their consumption experiences—whether positive or negative—in online reviews. Photos can help consumers express emotions and sentiment through visual features (e.g., color temperature and brightness), human actions or expressions (e.g., smiling or crying), and the meanings and features of objects in photos (e.g., flowers or animals) (Wang, 2018). Depending on the valence of photo content and visual features, photo sentiment also falls on a spectrum from negative to positive. Photo sentiment could be an important factor that influences fellow customers' perceptions and judgements about online reviews but has been largely overlooked in the extant literature.

Customer sentiment, no matter positive or negative, could positively influence review helpfulness by reducing individuals' perceived uncertainty and ambiguity when evaluating products or services (Bigne et al., 2020; Chua & Banerjee, 2016; Siering et al., 2018). On the contrary, neutral sentiment can adversely affect review helpfulness due to a lack of clear opinions. Negatively valenced online reviews, which either feature low review ratings or negative emotional content, have been deemed more helpful than neutral or positive online reviews due to the negativity bias effect (Rozin & Royzman, 2001). Specifically, negative information is more influential than positive information (Li et al., 2020; Wu, 2013). Positive emotional content expressed in review text has been reported to negatively affect online review helpfulness due to subjectivity or trustworthiness concerns (Li et al., 2020). However, positive photo sentiment may have a different impact on review helpfulness, as photos are more objective than the emotions and sentiment expressed in review text. Photos displaying positive sentiment could provide visual proof of corresponding review text and hence amplify review helpfulness. Accordingly, we posit that a non-linear relationship exists between photo sentiment and review helpfulness such that reviews featuring negative and positive photo sentiment may be more helpful than those with neutral sentiment.

Photos can also elicit affective responses. In psychological research on emotion elicitation, photos classified as having positive or negative valence per the International Affective Picture System (Lang et al., 1999) can successfully elicit corresponding emotional responses (Ito et al., 1998; Uhrig et al., 2016). As such, photos with positive sentiment can elicit positive emotions while those with negative sentiment can elicit negative emotions among viewers. Review photos can also stimulate mental imagery, which involves a mental visualization of the consumption experience depicted in the photos (Yang et al., 2017; Yoo & Kim, 2014). Photos with positive sentiment are therefore likely to stimulate imagery of positive and pleasant consumption experiences, whereas photos with negative sentiment are likely to evoke imagery of negative and unpleasant consumption experiences. We thus suggest that photo sentiment could positively influence review enjoyment as postulated below:

H2a: *A U-shaped relationship exists between review photo sentiment and review usefulness. Specifically, reviews with extremely negative or positive photo sentiment are more likely to be perceived as helpful than reviews with moderate photo sentiment.*

H2b: *Online review photo sentiment has a linear effect on review enjoyment. Specifically, with review photo sentiment becoming more positive, the review is perceived as more enjoyable; while with review photo sentiment becoming more negative, the review is perceived as less enjoyable.*

2.3 Moderating effect of number of review photos

The number of UGPs conveys content richness and can add to online reviews' comprehensiveness and intrigue while presenting sensory and aesthetic cues (Cheng & Ho, 2015; Hlee et al., 2019; Yang et al., 2017). The number of review photos has been found to positively affect review helpfulness and enjoyment (Cheng & Ho, 2015; Hlee et al., 2019; Yang et al., 2017). The number of review photos can also influence review helpfulness and enjoyment by strengthening the impacts of photo sentiment on these two attributes. This moderating effect could be explained by cue summation theory (Severin, 1967), which suggests that an increase in the number of relevant information cues (either from different modalities or different cues from the same modality) can enhance learning and understanding (Dwyer, 1978; Severin, 1967). In online reviews, the number of relevant information cues grows along with the number of review photos, resulting in greater understanding of the core message conveyed through the review photos (Ghasemaghaei et al., 2018).

As the number of review photos increases, a larger amount of information is available for fellow customers to judge the photos' sentiment (negative, neutral, or positive), and this helps enhance their understanding of photo sentiment and reinforces certainty in their judgement of such sentiment. Moreover, a higher number of review photos and relevant information cues could intensify photo sentiment, whether positive or negative. These processes indicate that increasing the number of review photos could alleviate uncertainty when evaluating consumption experiences to foster perceived review helpfulness when the photos convey positive or negative sentiment. In addition, a greater number of review photos can expand the quantity and diversity of imagery-evoking cues (i.e., perceptual and sensory cues in photos), which in turn evoke stronger emotional and aesthetic experiences corresponding to the photo sentiment (negative, neutral, or positive), and allow fellow customers to fully immerse themselves in the mental imagery of consumption experiences pertaining to the photo sentiment. On the basis of the above discussion, we propose the following hypotheses:

H3a: *The number of photos in an online review can strengthen the curvilinear effect of* *review photo sentiment on review usefulness.*

H3b: *The number of photos in an online review can strengthen the linear effect of review photo sentiment on review enjoyment.*

2.4 Moderating effect of review text–photo sentiment disparity

Review text and review photos can both convey sentiment. Text–photo sentiment disparity refers to the incongruence between the sentiment of review text and the sentiment of review photos. The congruence between text and images in content, meaning, or style has been widely studied in the communication and advertising literature. Scholars have reported mixed findings regarding the effects of text–image congruence: some studies have shown that text–image *congruence* improves learning, understanding, and message recall, fosters more favorable product attitudes, and even boosts product sales due to information fluency (Lochbuehler et al., 2018; Van Rompay et al., 2010; van Rompay & Pruyn, 2008; Wang & Song, 2020). Other research has indicated that text–image *incongruence* can attract more attention and improve products' memorability due to the novelty of information (Heckler & Childers, 1992; Houston et al., 1987; Lee & Mason, 1999).

Within the context of online reviews, congruence between review text and review photos in content has been found to be positively associated with review helpfulness (An et al., 2020; Ma et al., 2018). An et al. (2020) explained the positive impact of text–image congruence on review helpfulness from two angles. First, the consistency in information from different sources (e.g., review text and photos) could enhance the fluency of information processing, namely the ease or difficulty with which information is processed and integrated. In other words, relevant review photos facilitate understanding and processing of review text. Second, inconsistent information from review text and photos could decrease the trustworthiness of information and vice versa.

Similarly, the congruence of review text sentiment and photo sentiment could enhance information processing fluency by offering clear and consistent information about the reviewer's sentiment and attitude towards focal products or services. By contrast, text–photo sentiment disparity leads to uncertainty and ambiguity in the reviewer's sentiment and attitude; the effects of photo sentiment on review helpfulness could then be compromised. Uncertainty and ambiguity are often associated with negative affect such as anxiety and fear (Carleton, 2016). The positive impact of photo sentiment on review enjoyment might thus be mitigated. In addition, text–photo sentiment disparity could decrease review photos' trustworthiness due to inconsistent information, thereby weakening the effects of photo sentiment on review helpfulness and review enjoyment. Therefore, we assert that text–photo sentiment disparity plays a moderating role on the impacts of photo sentiment on review helpfulness and enjoyment as follows:

H4a: *Online review text–photo sentiment disparity can weaken the curvilinear effect of review photo sentiment on review usefulness.*

H4b: *Online review text–photo sentiment disparity can weaken the linear effect of review* *photo sentiment on review enjoyment.*

Our research framework is summarized in Figure 1.

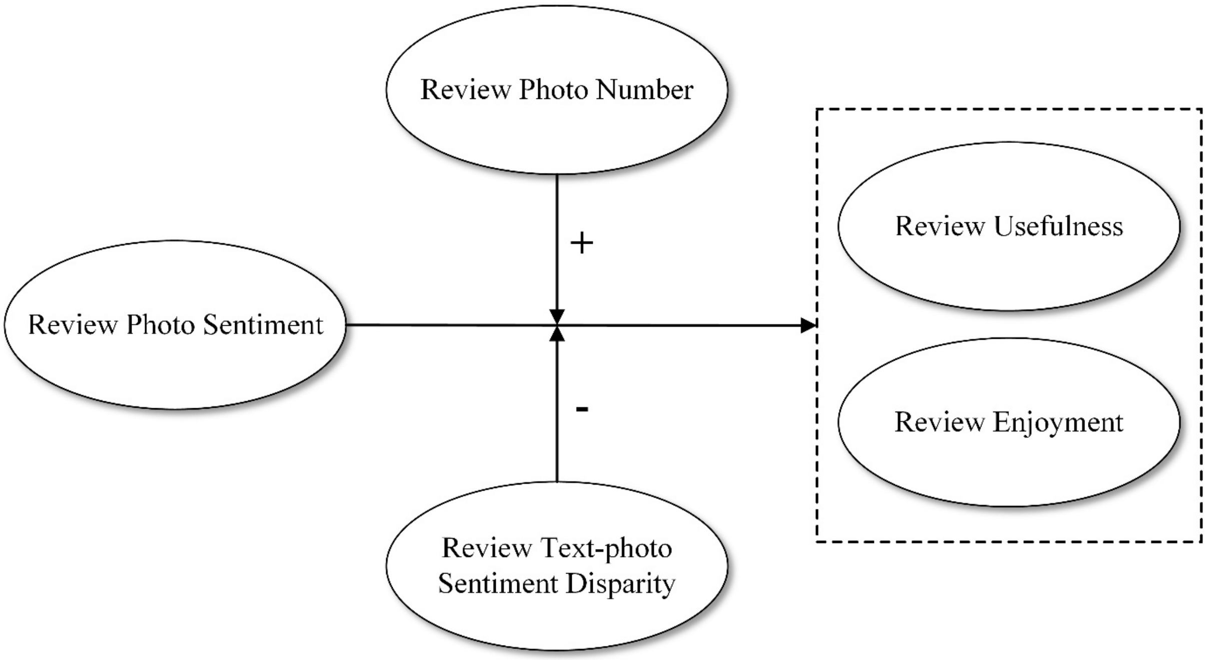

Figure 1. Research Framework

3. Research Methodology

3.1 Online review data

Yelp, a popular online review site in the United States, was used as the main data source in this study. Dining is an experience-oriented good, leading consumers to rely heavily on online reviews when choosing a restaurant (Li et al., 2019; Liu & Park, 2015). Our sample included 300 restaurants in Las Vegas, Nevada, a famous tourism destination. Stratified sampling was adopted to eliminate the potential business heterogeneity effect. First, we divided all restaurants in Las Vegas into groups based on several characteristics, including chain/independent and closed/open. Second, according to the overall proportions of Las Vegas restaurants, we calculated how many restaurants should be sampled per group. Third, to ensure that enough reviews were included in our final sample, we chose a corresponding sample from each subgroup based on the number of restaurant reviews. Online reviews from the most reviewed restaurants in each subgroup were selected as they are more likely to have higher involvement from readers/reviewers (Chen & Lurie, 2013; Li et al., 2019).

The Yelp dataset contained (1) restaurant-level data, such as each restaurant's name, location, and other associated attributes; (2) all reviews per restaurant, including the review date, review rating, review text, review photo(s), and review usefulness and enjoyment votes; and (3) the reviewer's information, such as their "elite" status per Yelp, registration date, and friends and followers. Review data spanned from January 2005 to February 2021, covering restaurants with different price levels, cuisines, business models, and operation statuses. We gathered 401,269 reviews and 336,092 photos in total, respectively; 105,237 reviews contained photo(s) while the remaining 296,032 reviews did not.

3.2 Variable measurement in research framework

Review usefulness was measured by a review's number of "useful" votes (Li et al., 2019; Liu & Park, 2015; Schuckert et al., 2015), otherwise known as review helpfulness (Pan & Zhang, 2011). *Review enjoyment* was evaluated based on the sum of a review's "cool" and "funny" votes (Li et al., 2019; Park & Nicolau, 2015).

With/without photo(s) and *number of review photos*. First, we evaluated whether each review contained photo(s) (*Picture*). If a review contained at least one photo, the value of *Picture* was coded as 1 and 0 otherwise. Second, we calculated a variable counting the number of photos per review (*ReviewPic*).

Review photo sentiment. We calculated review photo sentiment using deep learning method. When a review contains more than one photo, the average of review photo sentiments is used to represent the overall photo sentiment for a particular review. Deep learning, as a branch of machine learning methods, includes multiple processing layers in the architecture of artificial neural network (Schmidhuber, 2015). Compared with traditional machine learning methods, deep learning algorithms can extract intricate structures from a huge number of high-dimensional data (LeCun et al., 2015) and have shown superior performance in visual information processing (Guo et al., 2016; Ma et al., 2018; Zhang et al., 2019). In this study, we involved a supervised learning algorithm, which is the most common form of deep learning and needs a sample of known classifications to optimize the parameters of each layer. Our framework for calculating review photo sentiment appears in Figure 2.

Following Zhang and Luo (2018), a survey excluding review text and rating information was conducted to determine photo sentiment, indicating consumers' attitudes towards the product and desire to purchase it after viewing the review photos. It would be impossible to present all photos in the survey and obtain sentiment scores for each. Instead, we randomly selected 10,000 photos from our dataset and recruited 700 workers from Amazon Mechanical Turk to vote on the sampled photos. After a brief introduction, each worker was shown 300 photos on which they could vote as much as they wished. Votes indicated their attitudes towards the restaurant and willingness to dine there after viewing the review photos. Upon survey completion, each photo had received an average of 21 views and 10.02 votes. Thus, the photo sentiment value is calculated using normalization Eq. (1):

$$S_i = \frac{\frac{x_i^{vote}}{x_i^{view}} - \min\limits_{1 \le j \le 10,000}\left\{\frac{x_j^{vote}}{x_j^{view}}\right\}}{\max\limits_{1 \le j \le 10,000}\left\{\frac{x_j^{vote}}{x_j^{view}}\right\} - \min\limits_{1 \le j \le 10,000}\left\{\frac{x_j^{vote}}{x_j^{view}}\right\}} \tag{1}$$

where $S_i \in [0,1]$ denotes the value of photo sentiment, x_i^{vote} is the number of received votes of photo i, and x_i^{view} is the number of received views of photo i ($i = 1, 2, \ldots, 10,000$).

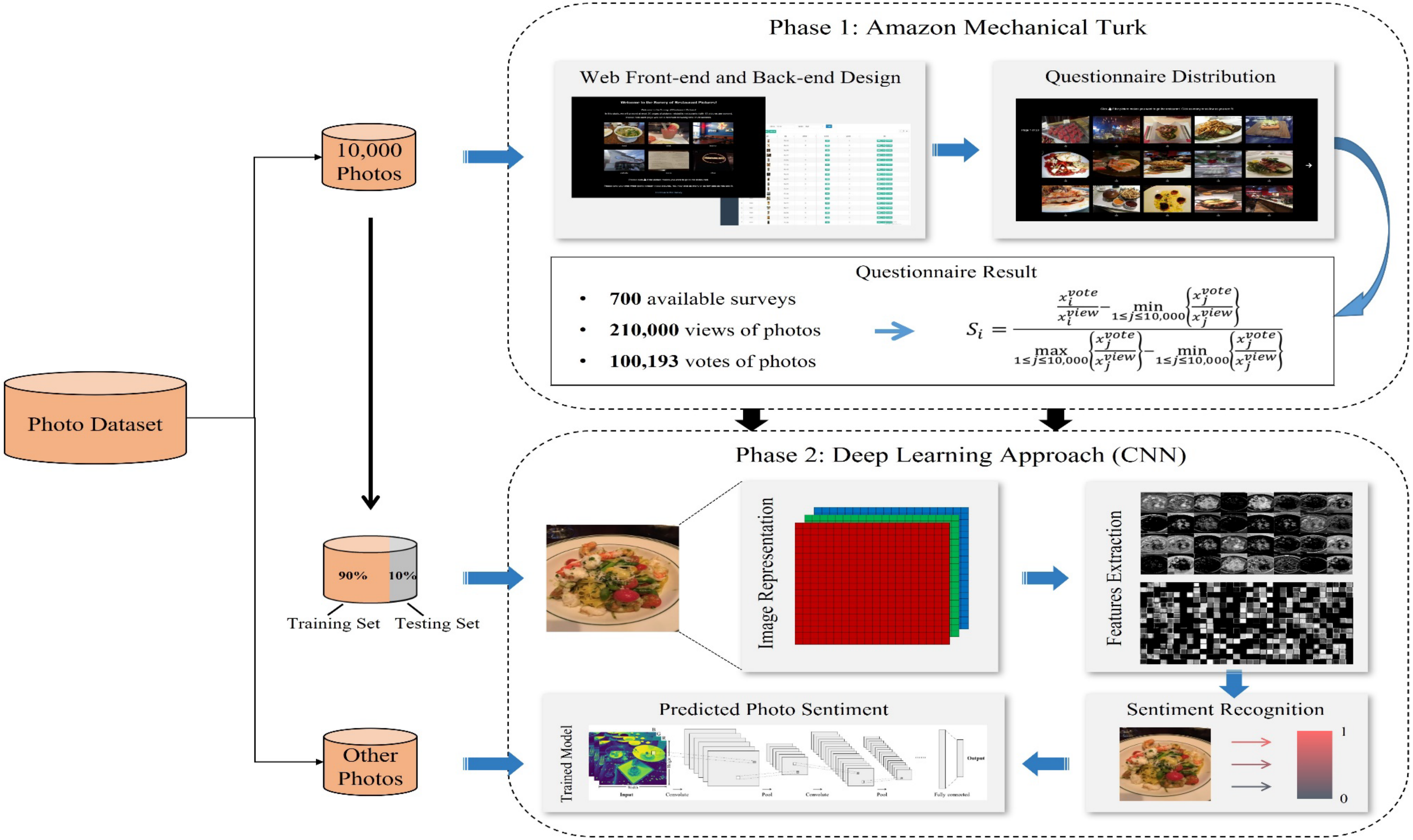

Figure 2. Framework for Measurement of Review Photo Sentiment

To further evaluate review photo sentiment in our dataset, we referred to a popular convolutional neural network (CNN) called GoogLeNet (Szegedy et al., 2015). GoogLeNet has achieved great success in the computer vision field and benefits from the structure of Inception based on the convolution calculation and pooling layer. A picture with 224×224 pixels and three channels (red, green, blue) were taken as input. Output was obtained after proceeding through GoogLeNet, including the convolutional, pooling, fully connected, and deep structure layers. This deep learning network can detect diverse photo characteristics to establish a relationship between them and photo sentiment. In this study, photo sentiment extraction produced a continuous photo sentiment value between 0 and 1 (1 denotes the most positive sentiment; 0 denotes the most negative). The last layer of GoogLeNet is softmax to generate categorical outcomes. We modified this structure so that the output would reflect the ratio of votes received to the number of views. To fine-tune model parameters, we selected a subset of photos from our original dataset with 90% taken as the training set and the remaining 10% as the test set. The model achieved a 0.015 mean square error in predicting the value of photo sentiment after training. In the auto-labeling stage, the neural network conveying relationships among photos' features and sentiment scores was applied to each photo in the full dataset.

Review text–photo sentiment disparity. In this study, disparity refers to deviation between the sentiment of textual content and corresponding review photo sentiment. To analyze review text sentiment, we applied a post-training approach based on the bidirectional encoder representations from transformers (BERT) language model (Xu et al., 2019). This model is increasingly common in natural language processing tasks. As a bidirectional language model, BERT can capture long-term information dependencies; pre-training and fine-tuning greatly improve its efficiency and performance (Devlin et al., 2018). However, limited training samples are insufficient for fine-tuning the pre-trained BERT and cannot achieve optimal performance. Comparatively, Xu et al.'s (2019) post-training approach includes specific domain knowledge and task knowledge for post-training before fine-tuning, demonstrating better performance when tuning examples with limited parameters (Figure 3). We took review text as input whose output was a continuous value between 0 and 1; a higher value indicates more positive expression, similar to the output for review photo sentiment. The absolute value of the difference between review text sentiment and review photo sentiment represents the discrepancy between the two sentiment values (i.e., *Disparity*). Figure 3 displays the process for measuring *Disparity*.

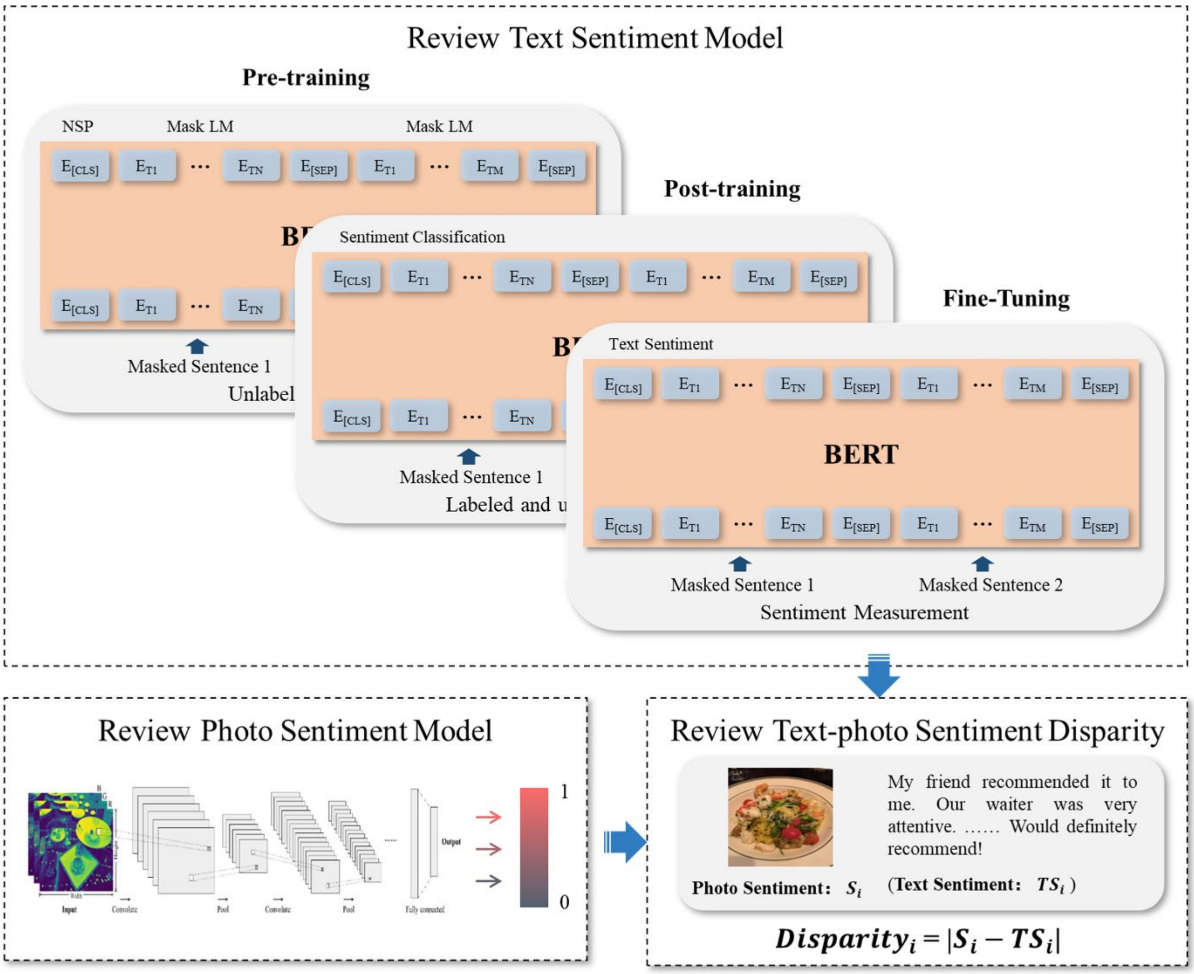

Figure 3. Framework for Measuring Review Text–photo Sentiment Disparity

To exclude the effects of confounding variables, several control variables were examined from three levels. First, at the review level, we controlled for each review's star rating; review text features, including review length and readability as measured with the Gunning-Fog Index (Gunning, 1969); and the number of days since the review was posted. Review length was calculated in Python. Review readability was estimated using TextSTAT, a text statistics calculation library based on Python. Second, at the reviewer level, we controlled for each reviewer's Yelp status (elite vs. non-elite) and social network size (i.e., number of followers and friends). Both factors were expected to positively influence review usefulness and review enjoyment. Third, in line with the literature (Amato & Amato, 2007; Li et al., 2019), we considered business heterogeneity by including restaurant fixed effects to control business-level heterogeneity. Table 1 presents the descriptive analyses of all variables.

Table 1. Variable Descriptive Statistics

Variable	Obs.	Mean	Std. Dev.	Skewness	Min	Max
Full online review sample (with and without photos)						
Usefulness	401,269	1.125	3.890	17.013	0	235
Enjoyment	401,269	1.273	5.843	20.820	0	416
Picture	401,269	0.262	0.440	1.081	0	1
Stars	401,269	4.009	1.245	-1.138	1	5
Length	401,269	114.807	109.259	2.495	1	1009
Readability	401,269	15.025	21.434	5498	0.4	388.71
Date	401,269	1741.546	1063.085	0.679	0	5865
Elite	401,269	0.207	0.405	1.448	0	1
Friends	401,269	167.302	405.390	6.546	0	5000
Followers	401,269	7.186	39.186	26.267	0	5818
Online review sample with photos						
Usefulness	105,237	2.188	6.506	11.336	0	235
Enjoyment	105,237	2.705	10.015	13.019	0	416
SentimentPic	105,237	0.463	0.078	-0.468	0	0.933
ReviewPic	105,237	3.194	2.870	3.067	1	50
Disparity	105,237	0.299	0.165	-0.004	0	0.935
Stars	105,237	4.282	0.976	-1.474	1	5
Length	105,237	141.709	126.211	2.145	1	1008
Readability	105,237	14.869	23.582	6.062	0.4	380.65
Date	105,237	1445.809	892.607	0.793	0	5697
Elite	105,237	0.368	0.482	0.545	0	1
Friends	105,237	271.383	564.432	4.920	0	5000
Followers	105,237	13.435	56.195	21.113	0	5818

3.3 Model specification

We applied a negative binomial 2 regression with robust standard errors to analyze count-dependent variables. All skewed variables (*Length*, *Readability*, *Friends*, *Followers*, and *ReviewPic*) were log-transformed. We first tested the main effects of reviews with/without photo(s) on review usefulness and enjoyment based on all review samples. The specified econometric models are shown in Eqs. (2) and (3), respectively:

$$\begin{aligned}\text{Usefulness}_{ijk} = {} & \beta_{10} + \beta_{11}\text{Picture}_{ijk} + \beta_{12}\text{Stars}_{ijk} + \beta_{13}log\text{Length}_{ijk} \\ & + \beta_{14}log\text{Readability}_{ijk} + \beta_{15}\text{Date}_{ijk} + \beta_{16}\text{Elite}_{ijk} + \beta_{17}log\text{Friends}_{j} \\ & + \beta_{18}log\text{Followers}_{j} + \textstyle\sum_{J} \lambda_j * R_j + \varepsilon_{1ijk}\end{aligned} \tag{2}$$

$$\begin{aligned}\text{Enjoyment}_{ijk} = {} & \beta_{20} + \beta_{21}\text{Picture}_{ijk} + \beta_{22}\text{Stars}_{ijk} + \beta_{23}log\text{Length}_{ijk} \\ & + \beta_{24}log\text{Readability}_{ijk} + \beta_{25}\text{Date}_{ijk} + \beta_{26}\text{Elite}_{ijk} + \beta_{27}log\text{Friends}_{j} \\ & + \beta_{28}log\text{Followers}_{j} + \textstyle\sum_{J} \lambda_j * R_j + \varepsilon_{2ijk}\end{aligned} \tag{3}$$

where *i*, *j*, and *k* respectively denote the review, reviewer, and restaurant.

To further examine the relationships among review photo sentiment and review usefulness and enjoyment, and to test the moderating effects of the number of review photos and review text–photo sentiment disparity, we proposed Eqs. (4) and (5) based on the sample of online reviews with photo(s). The specific econometric models are as follows:

$$\begin{aligned}
\text{Usefulness}_{ijk} = {} & \beta_{30} + \beta_{31} SentimentPic_{ijk}^{2} + \beta_{32} SentimentPic_{ijk} \\
& + \beta_{33} logReviewPic_{ijk} + \beta_{34} Disparity_{ijk} + \beta_{35} Stars_{ijk} \\
& + \beta_{36} logLength_{ijk} + \beta_{37} logReadability_{ijk} + \beta_{38} Date_{ijk} + \beta_{39} Elite_{ijk} \\
& + \beta_{310} logFriends_{j} + \beta_{311} logFollowers_{j} \\
& + \beta_{312} logReviewPic \times SentimentPic_{ijk}^{2} \\
& + \beta_{313} logReviewPic \times SentimentPic_{ijk} \\
& + \beta_{314} Disparity \times SentimentPic_{ijk}^{2} \\
& + \beta_{315} Disparity \times SentimentPic_{ijk} + \textstyle\sum_{J} \lambda_{j} * R_{j} + \varepsilon_{3ijk} \qquad (4)
\end{aligned}$$

$$\begin{aligned}
\text{Enjoyment}_{ijk} = {} & \beta_{40} + \beta_{41} SentimentPic_{ijk} + \beta_{42} logReviewPic_{ijk} + \beta_{43} Disparity_{ijk} \\
& + \beta_{44} Stars_{ijk} + \beta_{45} logLength_{ijk} + \beta_{46} logReadability_{ijk} + \beta_{47} Date_{ijk} \\
& + \beta_{48} Elite_{ijk} + \beta_{49} logFriends_{j} + \beta_{410} logFollowers_{j} \\
& + \beta_{411} logReviewPic \times SentimentPic_{ijk} \\
& + \beta_{412} Disparity \times SentimentPic_{ijk} + \textstyle\sum_{J} \lambda_{j} * R_{j} + \varepsilon_{4ijk} \qquad (5)
\end{aligned}$$

4. Empirical Results

4.1 The effect of review photo(s) presence on review usefulness/enjoyment

We first tested whether a review including photo(s) significantly affected review usefulness and enjoyment. Based on the 401,269 online reviews, including those with and without photos, Eqs. (2) and (3) were estimated for review usefulness and review enjoyment; results appear in Table 2. All models passed the overdispersion test, showing that negative binomial regression 2 models were suitable for our main analysis.

In Table 2, only control variables are included in Models 1.1 and 2.1. Models 1.2 and 2.2 incorporate the variable of interest (i.e., *Picture*). Estimation results demonstrate that compared with reviews without photo(s), reviews with photo(s) were generally perceived as more useful (coefficient = 0.3708260, $p < 0.01$) and more enjoyable (coefficient = 0.4095356, $p < 0.01$). Therefore, Hypothesis 1a ("Online reviews containing photos have a positive effect on review usefulness") and Hypothesis 1b ("Online reviews containing photos have a positive effect on review enjoyment") were each supported.

Table 2. Review Photo(s) Presence Effect on Review Usefulness and Enjoyment

	Review Usefulness		Review Enjoyment	
	Model 1.1	Model 1.2	Model 2.1	Model 2.2
Constant	-4.1013938***	-4.0087692***	-4.5603449***	-4.4680185***
	(-92.15)	(-90.58)	(-84.61)	(-83.33)
Stars	-0.1098423***	-0.1293528***	-0.0443945***	-0.0649988***
	(-49.15)	(-57.54)	(-15.95)	(-23.25)
logLength	0.5980475***	0.5643544***	0.5610036***	0.5238680***
	(157.99)	(148.57)	(123.26)	(114.56)
logReadability	0.0560862***	0.0672519***	0.0548818***	0.0677291***
	(14.99)	(18.05)	(11.80)	(14.62)
Date	0.0000222***	0.0000591***	0.0000690***	0.0001099***
	(7.16)	(18.78)	(18.32)	(28.75)
1.Elite	0.4832915***	0.4452980***	0.5580171***	0.5194708***
	(69.26)	(63.98)	(64.82)	(60.55)
logFriends	0.1923318***	0.1842979***	0.2352433***	0.2262439***
	(124.84)	(119.60)	(126.54)	(121.56)
logFollowers	0.2481943***	0.2337216***	0.3018427***	0.2851399***
	(107.18)	(101.01)	(105.76)	(99.81)
1.Picture		0.3708260***		0.4095356***
		(60.27)		(54.20)
Restaurant Fixed Effects	Yes	Yes	Yes	Yes
Alpha	1.0975767	1.0664454	2.0825563	2.0337277
Likelihood-ratio	2.50719e+05	2.44638e+05	5.04009e+05	4.93870e+05
test of alpha = 0	(P = 0.000)	(P = 0.000)	(P = 0.000)	(P = 0.000)
Log Likelihood	-4.50835e+05	-4.49040e+05	-4.34567e+05	-4.33108e+05
LR Chi2	1.72652e+05	1.76240e+05	1.45368e+05	1.48286e+05
Pseudo R2	0.161	0.164	0.143	0.146

Note: *, **, *** indicates 10%, 5%, and 1% significance level.

4.2 Review photo sentiment and review usefulness

Table 3 displays the estimation results regarding the impact of review photo sentiment on perceived review usefulness. Model 3.1 includes the control variables and independent variables of interest, including review photo sentiment (*SentimentPic-linear*) and its quadratic form (*SentimentPic-quadratic*). The significant and positive parameter of *SentimentPic-quadratic* indicates a U-shaped relationship between review photo sentiment and review usefulness (coefficient = 2.6100819, p < 0.01). As depicted in Figure 4, compared with moderate review photo sentiment, the more positive or negative the sentiment tends to be, the more likely the review is to be perceived as useful. Reviews with extremely negative or positive photo sentiment were thus more likely to be perceived as helpful than reviews with moderate photo sentiment. Accordingly, Hypothesis 2a ("A U-shaped relationship exists between review photo sentiment and review usefulness") was supported. The negative coefficient of *SentimentPic-linear* (β = -2.0929084, $p < 0.01$) revealed that positive review photo sentiment influenced perceived review usefulness more strongly than the impact of negative review sentiment (coefficient of *SentimentPic-linear* = -2.0929084, $p < 0.01$).

Table 3. The Effect of Review Photo Sentiment on Review Usefulness

	Review Usefulness	
	Model 3.1	Model 3.2
Constant	-3.5430521***	-2.9488264***
	(-26.89)	(-13.87)
Stars	-0.0750694***	-0.0864607***
	(-16.32)	(-18.25)
logLength	0.5457993***	0.5182623***
	(82.97)	(76.58)
logReadability	0.0998540***	0.0967729***
	(15.68)	(15.23)
Date	-0.0000375***	-0.0000299***
	(-6.37)	(-5.08)
1.Elite	0.4032596***	0.3950125***
	(37.28)	(36.54)
logFriends	0.2432347***	0.2389319***
	(85.61)	(84.07)
logFollowers	0.2550881***	0.2465644***
	(75.60)	(72.73)
SentimentPic—linear (S)	-2.0929084***	-4.7102544***
	(-4.52)	(-5.31)
SentimentPic—quadratic (S^2)	2.6100819***	5.7128331***
	(5.10)	(5.78)
logReviewPic		0.4920422***
		(3.37)
Disparity		-0.8781259*
		(-1.81)
(S) × logReviewPic		-1.9060813***
		(-2.90)
(S^2) × logReviewPic		2.4049277***
		(3.29)
(S) × Disparity		5.2076227**
		(2.35)
(S^2) × Disparity		-6.7497645***
		(-2.70)
Restaurant Fixed Effects	Yes	Yes
Alpha	0.9883123	0.9767386
Likelihood-ratio test of alpha = 0	1.37850e+05	1.36934e+05
	(P = 0.000)	(P = 0.000)
Log Likelihood	-1.59391e+05	-1.59142e+05
LR Chi2	6.20576e+04	6.25552e+04
Pseudo R2	0.163	0.164

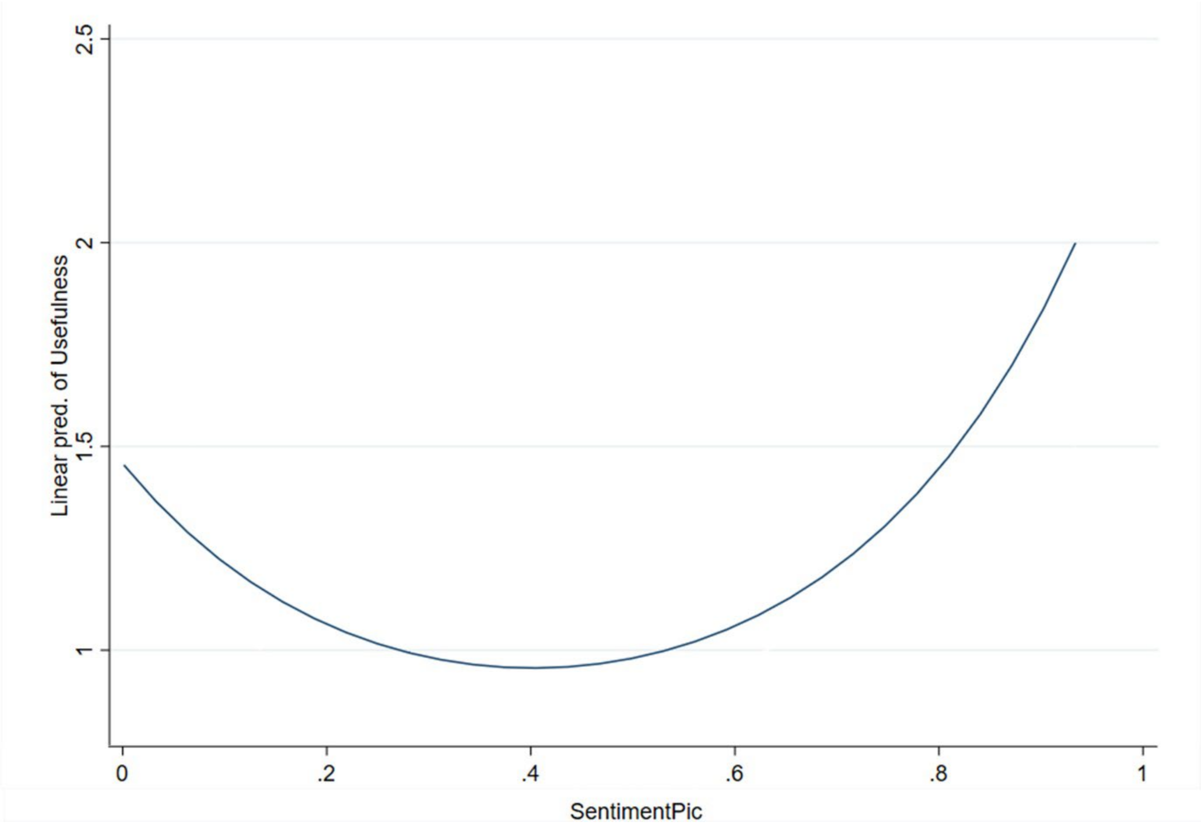

Figure 4. Impact of photo sentiment on review usefulness

Estimation results in Table 3 specify the factors moderating the impact of review photo sentiment on review usefulness by allowing review photo sentiment to interact with two variables of interest: (1) the number of photos in a review (*ReviewPic*) and (2) review text–photo sentiment disparity (*Disparity*). The estimated coefficients of linear interaction terms, including (S) × *logReviewPic* and (S) × *Disparity*, indicate how each moderating variable changes the turning point of the U-shaped relationship between review sentiment and perceived review usefulness (Figure 4). The estimated coefficients of nonlinear interaction terms, namely (S^2) × *logReviewPic* and (S^2) × *Disparity*, reflect how each moderating variable affects the convexity of the U-shaped relationship in Figure 4.

The estimation results of Model 3.2 in Table 3 suggest that the number of review photos accentuates the impact of review photo sentiment on review usefulness (coefficient of the interaction term $[S^2] \times ReviewPic$ = 2.4049277, p < 0.01). Therefore, Hypothesis 3a ("The number of photos in an online review can strengthen the effect of review photo sentiment on review usefulness") was supported. To better understand the moderating effect of the number of review photos, the interaction effect is illustrated in Figure 5; for the group of reviews including more photos, the impact of review photo sentiment on review usefulness was more intense with a steeper slope.

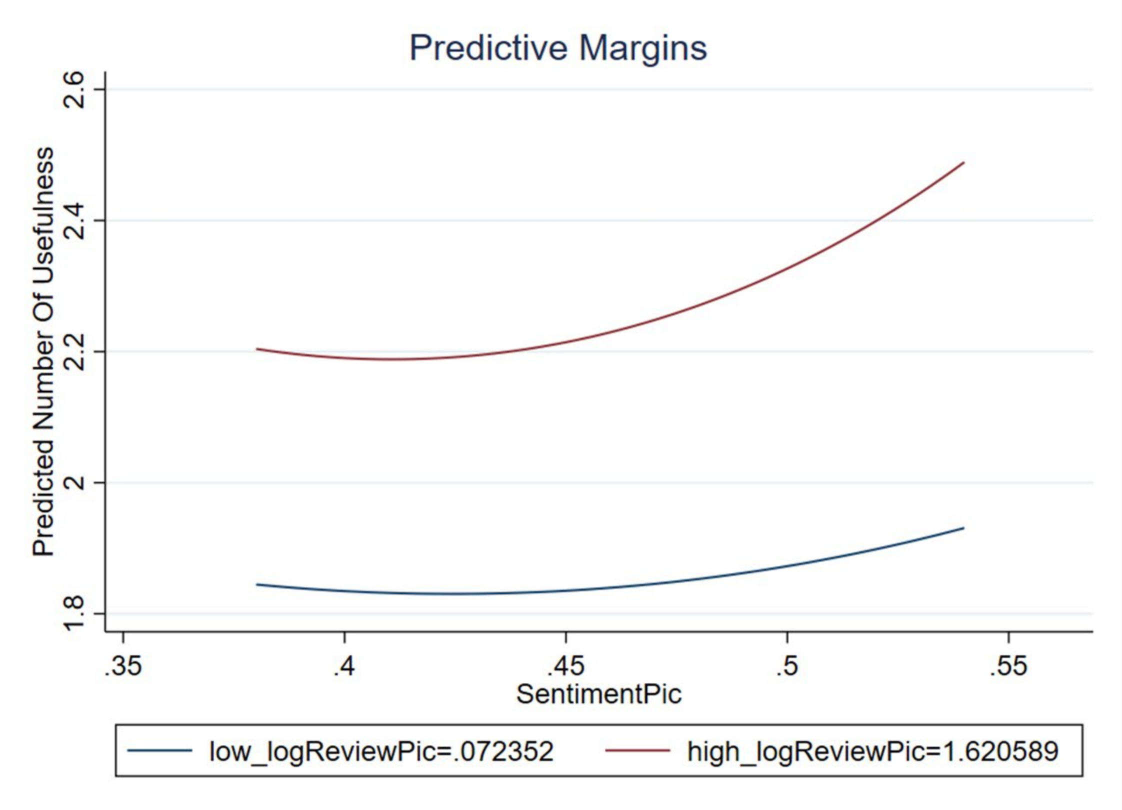

Figure 5. Moderating effect of number of photos in a review on review usefulness

In addition, we tested the moderating effect of review text–photo sentiment disparity (*Disparity*) to evaluate the impact of review photo sentiment on review usefulness. The estimation result of the interaction term *(S^2)* × *Disparity* was negative and significant (coefficient = -6.7497645, $p < 0.01$); that is, review text–photo sentiment disparity can weaken the nonlinear effect of photo sentiment on review usefulness. Hypothesis 4a was therefore supported. To clarify the moderating effect of review text–photo sentiment disparity, the interaction effect is displayed in Figure 6. People's trust in reviews presumably increases with gradual convergence between review text and review photo sentiment.

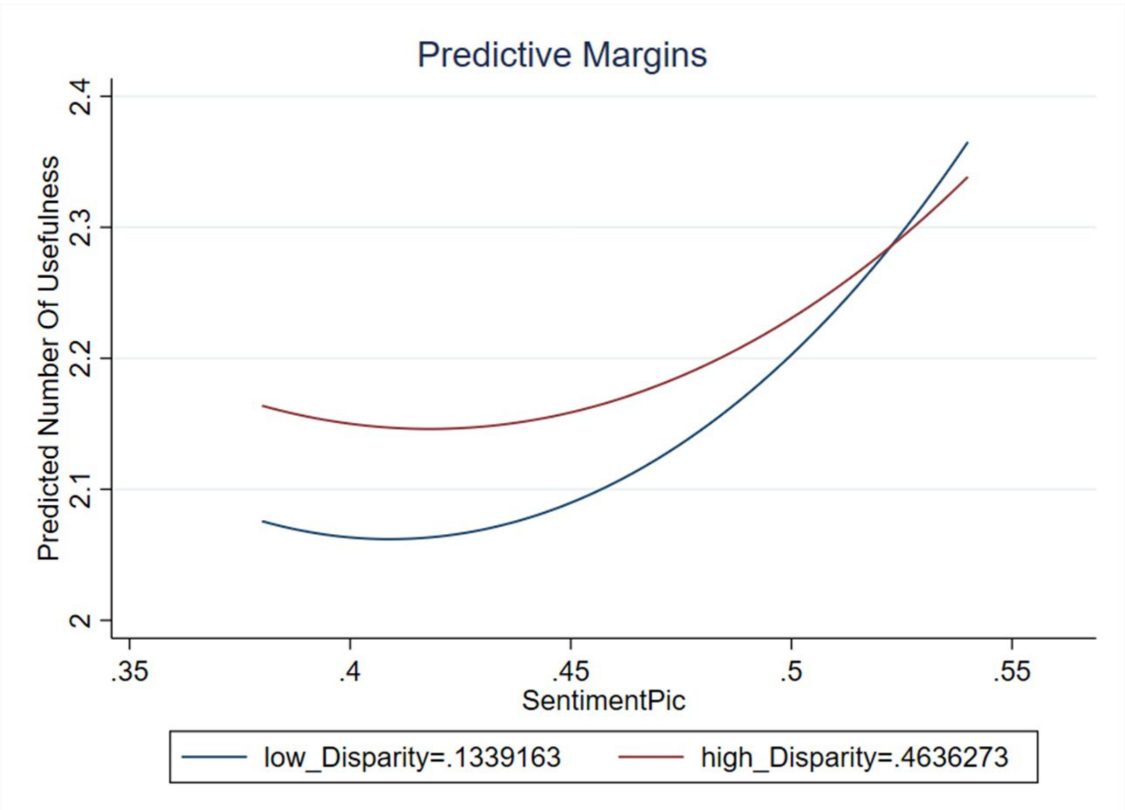

Figure 6. Moderating effect of disparity on review usefulness

4.3 Review photo sentiment and review enjoyment

Table 4 displays estimation results for the impact of review photo sentiment on perceived review enjoyment. Model 4.1 contains the control and independent variable of interest

(*SentimentPic*), with estimation results indicating a linear and significantly positive influence on review enjoyment (coefficient = 0.3335666, $p < 0.01$). Hypothesis 2b was therefore supported.

We further tested the moderating effects of the number of photos in a review (*ReviewPic*) and review text–photo sentiment disparity (*Disparity*); Model 4.2 in Table 4 shows these results. First, the number of photos strengthened the positive effect of review photo sentiment on review enjoyment, as evidenced by a positive interaction term between photo sentiment and photo number (coefficient = 0.2041845, $p < 0.05$). This interaction effect is depicted in Figure 7, which shows a steeper slope for a high number of photos in a review group. Hypothesis 3b was thus supported. Second, we tested the moderating effect of review text–photo sentiment disparity (*Disparity*). The negative coefficient of *(S)* × *Disparity* in Model 4.2 (coefficient = -1.1957257, $p < 0.01$) implies that review text–photo sentiment disparity can weaken the positive relationship between review photo sentiment and review enjoyment. As such, Hypothesis 4b was supported. This interaction effect appears in Figure 8 with a steeper slope for the low review text–photo sentiment disparity group.

Table 4. The Effect of Review Photo Sentiment on Review Enjoyment

	Review Enjoyment	
	Model 4.1	Model 4.2
Constant	-4.7054935***	-4.7951296***
	(-44.52)	(-39.16)
Stars	-0.0061648	-0.0216068***
	(-1.08)	(-3.68)
logLength	0.5398300***	0.5079283***
	(69.09)	(63.02)
logReadability	0.1123738***	0.1086896***
	(14.27)	(13.84)
Date	0.0000005	0.0000110
	(0.08)	(1.55)
1.Elite	0.4202488***	0.4093526***
	(31.85)	(31.05)
logFriends	0.2889381***	0.2838500***
	(86.37)	(84.75)
logFollowers	0.3090447***	0.2995048***
	(75.65)	(73.08)
SentimentPic—linear (*S*)	0.3335666***	0.6878868***
	(4.46)	(4.45)
logReviewPic		0.0616863
		(1.48)
Disparity		0.5984590***
		(3.34)
(*S*) ×logReviewPic		0.2041845**
		(2.27)
(*S*) × Disparity		-1.1957257***
		(-3.09)
Restaurant Fixed Effects	Yes	Yes
Alpha	1.7554892	1.7373580
Likelihood-ratio test of alpha = 0	2.81545e+05	2.80389e+05
	(P = 0.000)	(P = 0.000)
Log Likelihood	-1.60656e+05	-1.60435e+05
LR Chi2	5.33035e+04	5.37456e+04
Pseudo R2	0.142	0.143

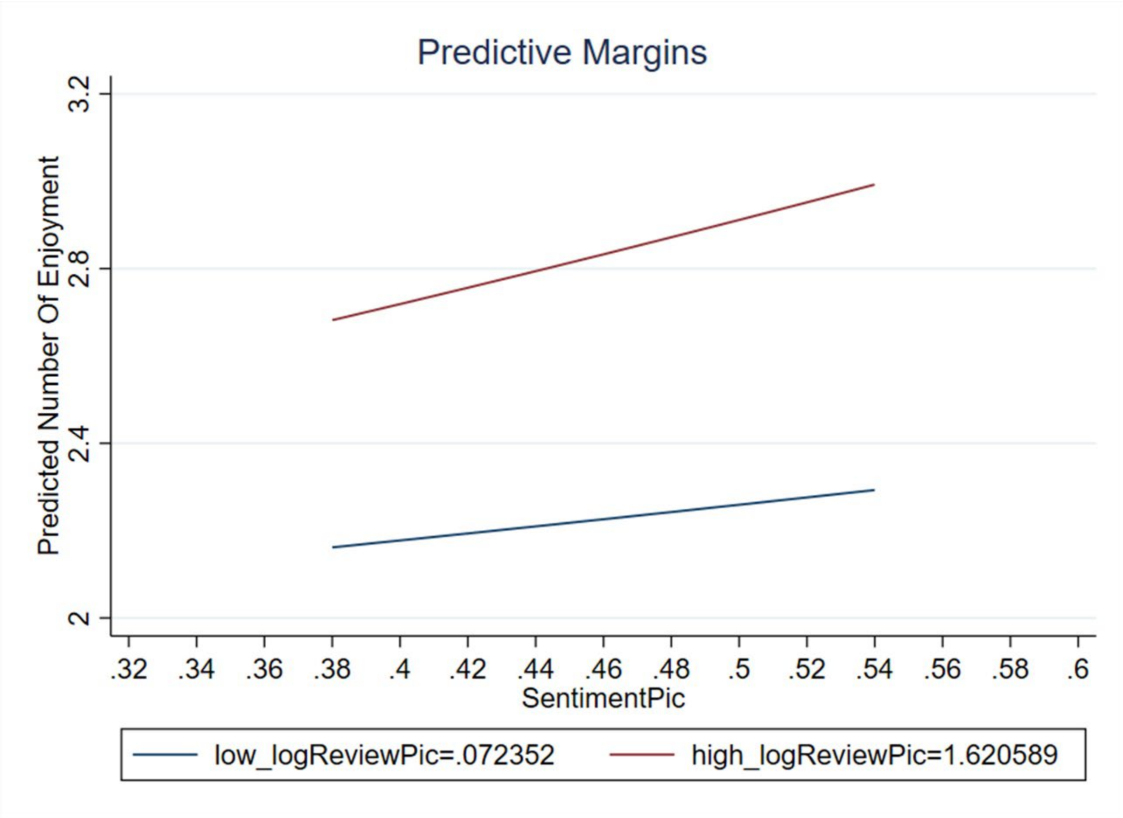

Figure 7. Moderating effect of number of photos in a review on review enjoyment

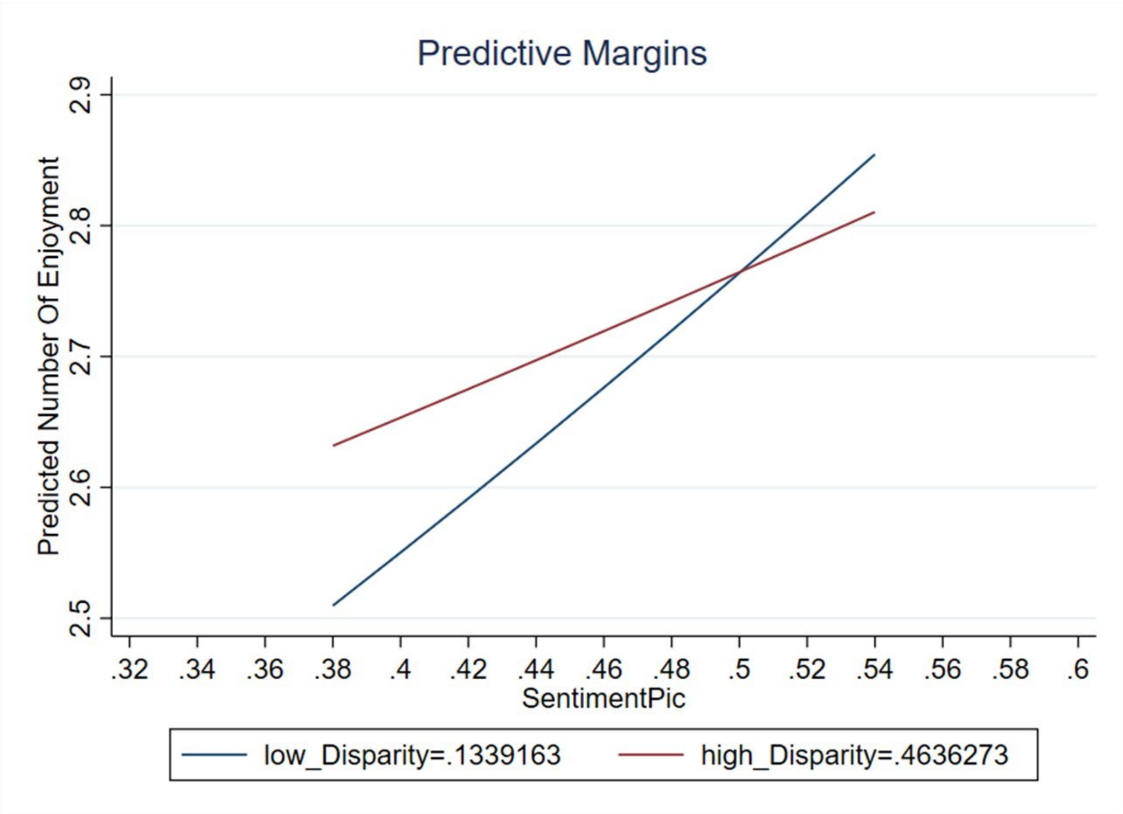

Figure 8. Moderating effect of disparity on review enjoyment

4.4 Robustness check: A tourist review sample analysis

The above empirical results show that the proposed hypotheses were supported in the entire restaurant online review sample in Las Vegas. To further check the robustness of the results, we identified the reviewers' tourist/local resident status and generated a subsample of online reviews written by tourists. A reviewer's "tourist" status was determined by their registration address. For example, if a reviewer registered an account with an address in San Francisco, which is outside of Las Vegas, they were considered a tourist. Our tourist subsample contained 305,379 online reviews, 79,346 of which included photo(s). We next re-analyzed how review photo sentiment affected perceived review usefulness and enjoyment based on a sample of online reviews written by tourists.

Consistent with our prior data analysis using the full review sample, the estimation results of Models 5.1 and 5.3 in Table 5 reveal that reviews with/without photo(s) significantly affected review usefulness and review enjoyment. Specifically, compared with a review without photo(s), a tourist review with photo(s) tended to be deemed more useful (coefficient = 0.3512116, $p < 0.01$) and more enjoyable (coefficient = 0.3796439, $p < 0.01$). In other words, people could obtain more useful and enjoyable information from tourist reviews with photo(s).

In addition, the estimation results of Model 5.2 in Table 5 using the tourist review sample demonstrate a U-shaped relationship between review photo sentiment and review usefulness. The results of Model 5.4 indicate that review photo sentiment had a linear and significantly positive influence on review enjoyment. The relationships between review photo sentiment and review usefulness/enjoyment were moderated by the number of review photos and review text–photo sentiment disparity. In particular, the number of review photos could strengthen this positive relationship, whereas review text–photo sentiment disparity weakened it. These estimation results coincide with our analysis of the full review sample.

Table 5. Empirical Results Using a Tourist Review Sample

	Review Usefulness		Review Enjoyment	
	Model 5.1	Model 5.2	Model 5.3	Model 5.4
Constant	-4.0084830***	-2.9569658***	-4.4750878***	-4.7403722***
	(-83.64)	(-11.55)	(-76.68)	(-33.88)
Stars	-0.1265034***	-0.0923180***	-0.0655146***	-0.0303507***
	(-48.10)	(-16.40)	(-19.94)	(-4.33)
logLength	0.5656484***	0.5171035***	0.5202816***	0.4972994***
	(127.07)	(64.18)	(96.70)	(51.44)
logReadability	0.0738719***	0.1149812***	0.0795721***	0.1358730***
	(17.31)	(15.40)	(14.92)	(14.63)
Date	0.0000380***	-0.0000443***	0.0000922***	0.0000005
	(10.75)	(-6.55)	(21.36)	(0.06)
1.Elite	0.4331861***	0.4123520***	0.5077552***	0.4309844***
	(54.56)	(32.53)	(51.56)	(27.76)
logFriends	0.1898682***	0.2464438***	0.2351182***	0.2987619***
	(104.59)	(73.02)	(106.94)	(74.52)
logFollowers	0.2273234***	0.2470433***	0.2757542***	0.2984101***
	(86.00)	(62.07)	(84.02)	(61.53)
1.Picture	0.3512116***		0.3796439***	
	(48.82)		(42.64)	
SentimentPic—linear (*S*)		-4.6558604***		0.5581059***
		(-4.29)		(3.05)
SentimentPic—quadratic (S^2)		5.5163193***		
		(4.56)		
logReviewPic		0.3468922**		0.0319223
		(1.97)		(0.63)
Disparity		-0.7413123		0.6106390***
		(-1.27)		(2.85)
(*S*) × logReviewPic		-1.3226135*		0.2579059**
		(-1.67)		(2.36)
(S^2) × logReviewPic		1.8176079**		
		(2.06)		
(*S*) × Disparity		4.5059224*		-1.2290148***

		(1.69)		(-2.67)
(S^2) × Disparity		-5.9309665**		
		(-1.98)		
Restaurant Fixed Effects	Yes	Yes	Yes	Yes
Alpha	1.0954445	1.0019313	2.1395979	0.5987384
Likelihood-ratio test of alpha = 0	1.80849e+05	9.83931e+04	3.81190e+05	2.07864e+05
	(P = 0.000)	(P=0.000)	(P = 0.000)	(P = 0.000)
Log Likelihood	-3.35253e+05	-1.15058e+05	-3.23851e+05	-1.15631e+05
LR Chi2	1.27287e+05	4.55249e+04	1.06973e+05	3.91756e+04
Pseudo R2	0.160	0.165	0.142	0.145

4.5 Robustness check: A smartphone era sample analysis

The rise of smartphones could be a major factor affecting users' willingness to post reviews with photos. To further test the robustness of our models, we examined the influence of review photo sentiment on perceived review usefulness and enjoyment following the popularity of smartphones. Nielsen Insights indicated that more than half of mobile phone users in the United States have used smartphones since 2012 (Nielsen, 2012), causing smartphones to dominate the market. Therefore, we took review data during the smartphone era as a sample to verify our findings.

Based on reviews with and without photos after 2012, we estimated Models 6.1 and 6.3 (Table 6). Estimation results show that, in the smartphone era, reviews with photo(s) had a positive and significant effect on perceived review usefulness and review enjoyment, consistent with our previous analysis. Moreover, based on online reviews with photos after 2012, we estimated Models 6.2 and 6.4 in Table 6. Similar to the main estimation results in Tables 3 and 4, a U-shaped relationship was identified between review photo sentiment and review usefulness. A positive correlation manifested between review photo sentiment and review enjoyment. In addition, the number of photos in a review strengthened the relationship between review photo sentiment and review usefulness/enjoyment, whereas text–photo sentiment disparity weakened it.

Table 6. Empirical Results Using a Smartphone Era Review Sample

	Review Usefulness		Review Enjoyment	
	Model 6.1	Model 6.2	Model 6.3	Model 6.4
Constant	-4.0983281***	-3.4508587***	-4.4106581***	-4.8897438***
	(-62.43)	(-10.34)	(-56.09)	(-30.29)
Stars	-0.1420599***	-0.1028110***	-0.0796477***	-0.0289955***
	(-50.23)	(-18.96)	(-22.83)	(-4.30)
logLength	0.5592793***	0.5153329***	0.5090910***	0.4976531***
	(114.21)	(65.45)	(87.14)	(53.19)
logReadability	0.0978871***	0.1242562***	0.0968883***	0.1394458***
	(18.69)	(15.88)	(15.06)	(14.51)
Date	0.0000747***	-0.0000002	0.0001476***	0.0000287**
	(11.50)	(-0.02)	(18.93)	(2.44)
1.Elite	0.5227248***	0.4461773***	0.5831510***	0.4437194***
	(57.97)	(35.45)	(53.00)	(28.99)
logFriends	0.1816245***	0.2313654***	0.2171623***	0.2754140***
	(96.57)	(72.76)	(96.92)	(73.66)
logFollowers	0.3049029***	0.2863109***	0.3643024***	0.3420939***
	(102.17)	(72.66)	(99.84)	(72.00)
1.Picture	0.3891289***		0.4138826***	
	(52.58)		(46.29)	
SentimentPic—linear (*S*)		-2.8888010**		0.5574129**
		(-1.99)		(2.35)
SentimentPic—quadratic (S^2)		3.3877918**		
		(2.08)		
logReviewPic		0.9543640***		0.0129464
		(4.02)		(0.19)
Disparity		-1.3389265**		0.8515118***
		(-2.46)		(4.20)
(*S*) × logReviewPic		-4.0705189***		0.4651890***
		(-3.78)		(3.21)
(S^2) × logReviewPic		5.1343543***		
		(4.24)		
(*S*) × Disparity		7.6290819***		-1.8370072***

		(3.05)		(-4.16)
(S^2) × Disparity		-9.7282510***		
		(-3.42)		
Restaurant Fixed Effects	Yes	Yes	Yes	Yes
Alpha	1.1027449	1.0223674	2.0486338	1.8029188
Likelihood-ratio test of alpha = 0	1.73697e+05	1.14159e+05	3.30933e+05	2.23491e+05
	(P = 0.000)	(P = 0.000)	(P = 0.000)	(P = 0.000)
Log Likelihood	-2.88607e+05	-1.24381e+05	-2.77153e+05	-1.24261e+05
LR Chi2	1.30500e+05	5.20221e+04	1.08973e+05	4.43590e+04
Pseudo R2	0.184	0.173	0.164	0.151

5. Discussion and Implications

5.1 Conclusion and discussion

Given the importance of visual content, leveraging such content to boost consumer engagement is essential for businesses. Consumer engagement with online reviews mainly manifests in review helpfulness and enjoyment votes, which play a pivotal role in consumers' information processing, online review adoption, and purchase behaviors (Choi & Leon, 2020; Filieri et al., 2018; Hlee et al., 2019; Lee & Hong, 2019; Lee & Choeh, 2018; Park & Nicolau, 2015) Emotions reflected in online reviews could reinforce consumer engagement (Bigne et al., 2021; Chua & Banerjee, 2016; Siering et al., 2018). Review sentiment can be conveyed through text (Alaei et al., 2019; Ma et al., 2018), photos, or videos (Campos et al., 2015; Deng & Li, 2018). The effects of UGPs have recently attracted increasing attention (e.g., An et al., 2020; Cai & Chi, 2020; Marder et al. 2021). Even so, little is known about the role of review photo sentiment or about the effects of the interactions or congruence between review text and photos. We sought to gain an in-depth understanding of customers' engagement with reviews (i.e., review usefulness and enjoyment) based on UGP characteristics (i.e., review photo sentiment). We adopted a deep learning method to extract photo sentiment from hundreds of thousands of heterogeneous review data (i.e., numerical, text, photo) on Yelp, combined with econometric models. Results address a relevant knowledge gap by identifying the effects of review photo sentiment on review usefulness and enjoyment. Our conclusions are summarized below.

First, we found that online reviews including UGPs were perceived to be more useful and enjoyable than those without. These outcomes show that UGPs can complement potential customers' perceived review usefulness (An et al., 2020; Bigne et al., 2021; Ma et al., 2018). Li et al. (2019) discovered that sensory cues (i.e., visual, auditory, and haptic) in review text could contribute to review enjoyment. Our study further demonstrates that reviews containing photos, which function as a rich source of sensory cues, enhance review enjoyment.

Second, the estimation results present a U-shaped relationship between review photo sentiment and review usefulness, with review photo sentiment having a linear positive effect on review enjoyment. The positive or negative sentiment reflected in reviews is thought to be more helpful than neutral online reviews (Bigne et al., 2021; Siering et al., 2018) due to unclear opinions in neutral reviews. As Wang (2018) mentioned, photos offer another way to convey customers' sentiments about their experiences. Our results confirm a non-linear relationship between review photo sentiment and review usefulness. Additionally, in line with findings from psychologists, pictures can more effectively arouse viewers' emotional resonance (Ito et al., 1998; Uhrig et al., 2016). Review photo sentiment could evoke similar emotional responses among consumers.

Third, our study introduces the concept of review text–photo sentiment disparity and tests its moderating effect. Shin et al. (2020) found that blog image–text content similarity (based on cosine similarity) positively influences users' reblogs and likes. By contrast, An et al. (2020) demonstrated that the effect of photo caption–text content congrence (based on cosine similarity) on review helpfulenss is insignificant. Different from these studies, we examined the interaction effect of social media text and photos from a sentiment perspective. Results show that sentiment disparity between review text and photo(s) can reduce the impact of review photo sentiment due to the uncertainty of text–photo sentiment inconsistency. The effects of review photo sentiment on review helpfulness and enjoyment are correspondingly mitigated.

Lastly, we found that the number of UGPs can strengthen the effects of review photo sentiment on review usefulness and enjoyment. Supported by cue summation theory (Severin, 1967), review content richness reinforces review helpfulness and enjoyment by adding information and sensory cues (Cheng & Ho, 2015; Yang et al., 2017). More photos could provide more imagery-evoking cues (i.e., perceptual and sensory cues), thus strengthening the impact of review photo sentiment on customer engagement (i.e., "useful," "funny," and "cool" review votes). This outcome further clarifies the role of UGPs' content richness.

5.2 Theoretical implications

This study makes several important theoretical contributions. First, existing literature on online review sentiment has been solely focusing on sentiment analysis of review texts (e.g., Alaei et al., 2019; Geetha et al., 2017; Ma et al., 2018), with sentiment of review photos being completely overlooked. Our work is among the earliest to address the effects of photo sentiment on consumers' perceived review helpfulness and review enjoyment. Specifically, we identified a U-shaped relationship between photo sentiment and review helpfulness along with a positive relationship between photo sentiment and review enjoyment. Our study also provides further evidence of the positive impact of review photos on review helpfulness, consistent with conclusions from An et al. (2020), Cheng and Ho (2015), Lee (2018), and Ma et al. (2018).

Existing studies reported mixed findings regarding the impacts of review photos on review helpfulness (Hlee et al., 2019; Lee, 2018). By introducing photo sentiment and text–photo sentiment disparity, our study helps to reconcile these contradictory findings. Besides whether or not including photos, the sentiment of review photos and the congruence between text and photo sentiment should be considered while studying the effects of review photos on consumers' perceived review helpfulness and enjoyment.

Second, we tested the boundary effects of photo sentiment on review helpfulness and review enjoyment (i.e., under what conditions the impacts of photo sentiment on review helpfulness and review enjoyment become stronger or weaker). Our findings indicate that the number of review photos can strengthen the effects of photo sentiment on review helpfulness and enjoyment, and text–photo sentiment disparity can weaken the impacts of photo sentiment on review helpfulness and enjoyment. Our findings on text–photo sentiment disparity extend existing research testing the text–photo interaction effect on social media (An et al., 2020; Shin et al., 2020), whose results were inconsistent regarding the impact of text–image congruence in content. Our work further expands such literature by evaluating other aspects of text–image congruence, i.e., sentiment congruence, and responds to a call for more studies on the interactions between textual and visual content in online reviews (Li et al., 2021).

Third, this study adds to the literature on the hedonic value of online reviews. Hedonic value of online reviews is crucial for engaging consumers and shaping their attitudes and behaviors (Park & Nicolau, 2015). However, few studies have investigated what contributes to the hedonic value or perceived enjoyment of online reviews (Hlee et al., 2019; Li et al., 2019; Park & Nicolau, 2015; Yang et al., 2017). This study contributed to the literature by uncovering the role of photo sentiment on review enjoyment along with the moderating effects of the number of review photos and text–photo sentiment disparity on the relationship between photo sentiment and review enjoyment.

Last but not least, by focusing on the role of UGPs in online reviews, this study enriched the literature on online reviews and consumers' online information search. Visual information, including review photos, are playing an increasingly important role in consumer decision-makings (An et al., 2020; Li et al., 2021; Liu et al., 2022). Existing studies mainly focused on the effects of whether online reviews include photos and the number of review photos (An et al., 2020; Cheng & Ho, 2015; Hlee et al., 2019; Lee, 2018; Li et al., 2021; Ma et al., 2018), the present study offered a unique perspective on how review photos affect consumer decision-makings by investigating the influences of review photo sentiments and text–photo sentiment disparity.

5.3 Practical implications

This study also offers practical contributions. People have become increasingly reliant on online reviews when purchasing products, especially experiential items in hospitality and tourism (Bei et al., 2004; Senecal & Nantel, 2004; Wu et al., 2016). As photo uploads become more convenient, photos' effects represent a core topic in marketing. The practical implications of our study are as follows.

First, our work provides an avenue for restaurant managers to deepen their understanding of customers' opinions derived from the vast amount of UGC (e.g., review text and photos). Sentiment represents an essential way to understand consumers' attitudes and opinions, and UGPs often lead to initial impressions and efficient messaging. Scholars have argued that photos play a major role in consumers' perceptions of virtual products or services (Ren et al., 2020). Therefore, practitioners should not ignore photos while focusing on reviewers' written commentary. Given the real-time nature of reviews, businesses can better capture consumers' emotional changes by integrating UGPs in review data analysis. This study can enhance managers' awareness of photo sentiment and text–photo sentiment disparity, both of which promote knowledge of market trends and performance management.

Second, for online review platforms, our findings indicate that online review websites should develop scalable image- and text-processing algorithms to detect the most useful and enjoyable reviews. Reviews could then be sorted accordingly to ensure that consumers are exposed to more positive reviews and positive sentiment photos. The rise of a large volume of information on the Internet not only increases consumers' search costs but also presents a challenge for managers. Users' information-filtering process is often considered part of the consumption experience. By proposing a user-oriented image- and text-processing method to recommend the most useful and enjoyable reviews, consumers can quickly obtain valuable feedback. This benefit can enhance a platform's competitive advantage. The extracted information also serves as a reference for businesses to efficiently identify consumers and market trends. Firms can then formulate appropriate advertising and development strategies.

Third, businesses should pay close attention to reviews with negative review photo sentiment (as well as reviews featuring many photos and high photo–text sentiment congruence), as these reviews are perceived as extremely useful. The negative sentiments of UGPs reflect consumers' negative attitudes. If not given sufficient attention, these reviews could adversely affect businesses. Managers should address these restaurant aspects to improve consumers' dining experiences and online evaluations.

5.4 Limitations and future research

Several limitations apply to this research and should be investigated in the future. First, we employed a consumer-based photo sentiment calculation method; future studies could further validate the accuracy of photo sentiment calculation through other approaches, such as Sentibank (Borth et al., 2013). In addition, as the information context is of great significance to information processing, further studies should consider the role of multimodal information integration, i.e., review text and photo together, when calculating the review sentiment. Second, we only focused on the moderating effects of the number of photos in a review and review text–photo sentiment disparity. Subsequent work can assess moderating effects of other review text and review photo characteristics. Third, the data used in our study were limited to a single city (i.e., Las Vegas) and one product (i.e., restaurants). Follow-up studies could test whether our findings are generalizable based on online review data for different products and offerings in other regions, such as hedonic versus utilitarian goods and restaurants versus hotels versus tourist attractions.

References:

Alaei, A. R., Becken, S., & Stantic, B. (2019). Sentiment analysis in tourism: Capitalizing on big data. *Journal of Travel Research*, *58*(2), 175-191.

Amato, L. H., & Amato, C. H. (2007). The effects of firm size and industry on corporate giving. *Journal of Business Ethics*, *72*(3), 229-241.

An, Q. X., Ma, Y. F., Du, Q. Z., Xiang, Z., & Fan, W. G. (2020). Role of user-generated photos in online hotel reviews: an analytical approach. *Journal of Hospitality & Tourism Management*, *45*, 633-640.

Aydin, G. (2020). Social media engagement and organic post effectiveness: A roadmap for increasing the effectiveness of social media use in hospitality industry. *Journal of Hospitality Marketing & Management*, *29*(1), 1-21.

Babin, B. J., Darden, W. R., & Griffin, M. (1994). Work and/or fun: Measuring hedonic and utilitarian shopping value. *Journal of Consumer Research*, *20*(4), 644-656.

Bei, L.-T., Chen, E. Y., & Widdows, R. (2004). Consumers' online information search behavior and the phenomenon of search vs. experience products. *Journal of Family Economic Issues*, *25*(4), 449-467.

Bigne, E., Chatzipanagiotou, K., & Ruiz, C. (2020). Pictorial content, sequence of conflicting online reviews and consumer decision-making: The stimulus-organism-response model revisited. *Journal of Business Research*, *115*, 403-416.

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.

Bullas, J. (2017). *6 powerful reasons why you should include images in your marketing–Infographic*. Retrieved from http://www.jeffbullas.com//05/28/6-powerful-reasons-why-you-should-include-images-in-your-marketing-infographic/

Cai, R., & Chi, C. G.-Q. (2020). A recipe for food promotion: Effects of color brightness on food evaluations and behavioral intentions. *International Journal of Contemporary Hospitality Management*, *13*(12), 3925-3947.

Campos, V., Salvador, A., Giró-i-Nieto, X., & Jou, B. (2015). Diving deep into sentiment: Understanding fine-tuned CNNs for visual sentiment prediction. *In Proceedings of the 1st International Workshop on Affect & Sentiment in Multimedia* (pp. 57-62).

Carleton, R. N. (2016). Into the unknown: A review and synthesis of contemporary models involving uncertainty. *Journal of Anxiety Disorders*, *39*, 30-43.

Chen, Z., & Lurie, N. H. (2013). Temporal contiguity and negativity bias in the impact of online word of mouth. *Journal of Marketing Research*, *50*(4), 463-476.

Cheng, Y.-H., & Ho, H.-Y. (2015). Social influence's impact on reader perceptions of online reviews. *Journal of Business Research*, *68*(4), 883-887.

Choi, H. S., & Leon, S. (2020). An empirical investigation of online review helpfulness: A big data perspective. *Decision Support Systems*, *139*, 113403.

Chua, A. Y., & Banerjee, S. (2016). Helpfulness of user-generated reviews as a function of review sentiment, product type and information quality. *Computers in Human Behavior*, *54*, 547-554.

Deng, N., & Li, X. R. (2018). Feeling a destination through the "right" photos: A machine learning model for DMOs' photo selection. *Tourism Management*, *65*, 267-278.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:.1810.04805*.

Dwyer, F. M. (1978). *Strategies for improving visual learning: Instructor's manual*. Learning Services.

Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. *Journal of Big Data*, *2*(1), 1-14.

Filieri, R. (2016). What makes an online consumer review trustworthy? *Annals of Tourism Research*, *58*, 46-64.

Filieri, R., McLeay, F., Tsui, B., & Lin, Z. (2018). Consumer perceptions of information helpfulness and determinants of purchase intention in online consumer reviews of services. *Information & Management*, *55*(8), 956-970.

Geetha, M., Singha, P., & Sinha, S. (2017). Relationship between customer sentiment and online customer ratings for hotels-An empirical analysis. *Tourism Management*, *61*, 43-54.

Geise, S., & Baden, C. (2015). Putting the image back into the frame: Modeling the linkage between visual communication and frame-processing theory. *Communication Theory*, *25*(1), 46-69.

Ghasemaghaei, M., Eslami, S. P., Deal, K., & Hassanein, K. (2018). Reviews' length and sentiment as correlates of online reviews' ratings. *Internet Research*, *28*(3), 544-563.

Glenberg, A. M., & Langston, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. *Journal of Memory and Language*, *31*(2), 129-151.

Guan, Y., Tan, Y., Wei, Q., & Chen, G. (2019). The dark side of images: effect of customer generated images on product assessment. *International Conference on Information Systems, ICIS 2019*

Gunning, R. (1969). The fog index after twenty years. *Journal of Business Communication*, *6*(2), 3-13.

Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., & Lew, M. S. (2016). Deep learning for visual understanding: A review. *Neurocomputing*, *187*, 27-48.

Ham, J., Lee, K., Kim, T., & Koo, C. (2019). Subjective perception patterns of online reviews: A comparison of utilitarian and hedonic values. *Information Processing & Management*, *56*(4), 1439-1456.

Heckler, S. E., & Childers, T. L. (1992). The role of expectancy and relevancy in memory for verbal and visual information: What is incongruency? *Journal of Consumer Research*, *18*(4), 475-492.

Hlee, S., Lee, H., Koo, C., & Chung, N. (2021). Fake reviews or not: exploring the relationship between time trend and online restaurant reviews. *Telematics and Informatics*, *59*, 101560.

Hlee, S., Lee, J., Yang, S.-B., & Koo, C. (2019). The moderating effect of restaurant type on hedonic versus utilitarian review evaluations. *International Journal of Hospitality Management*, *77*, 195-206.

Houston, M. J., Childers, T. L., & Heckler, S. E. (1987). Picture-word consistency and the elaborative processing of advertisements. *Journal of Marketing Research*, *24*(4), 359-369.

Instagram. (2021). *#travel hashtag on Instagram photos and videos*. Retrieved from https://www.instagram.com/explore/tags/travel/?hl=en/

Ito, T. A., Cacioppo, J. T., & Lang, P. J. (1998). Eliciting affect using the International Affective Picture System: Trajectories through evaluative space. *Personality and Social Psychology Bulletin*, *24*(8), 855-879.

Krishna, A. (2012). An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior. *Journal of Consumer Psychology*, *22*(3), 332-351.

Kuan, K. K., & Hui, K.-L. (2015). What makes a review voted? An empirical investigation of review voting in online review systems. *Journal of the Association for Information Systems*, *16*(1), 48-71.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1999). International affective picture system (IAPS): Instruction manual and affective ratings. *The center for research in psychophysiology, University of Florida*.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436-444.

Lee, I. (2018). Usefulness, funniness, and coolness votes of viewers: An analysis of social shoppers' online reviews. *Industrial Management & Data Systems*, *118*(4), 700-713.

Lee, J., & Hong, I. B. (2019). Consumer's electronic word-of-mouth adoption: The trust transfer perspective. *International Journal of Electronic Commerce*, *23*(4), 595-627.

Lee, S., & Choeh, J. Y. (2018). The interactive impact of online word-of-mouth and review helpfulness on box office revenue. *Management Decision*, *56*(4), 849-866.

Lee, S., & Choeh, J. Y. (2020). The impact of online review helpfulness and word of mouth communication on box office performance predictions. *Humanities and Social Sciences Communications*, *7*(1), 1-12.

Lee, Y. H., & Mason, C. (1999). Responses to information incongruency in advertising: The role of expectancy, relevancy, and humor. *Journal of Consumer Research*, *26*(2), 156-169.

Li, C., Kwok, L., Xie, K., & Liu, J. (2021). Let photos speak: the effect of user-generated visual content on hotel review helpfulness. *Journal of Hospitality Tourism Research*, Advance online publication.

Li, H., Liu, H., & Zhang, Z. (2020). Online persuasion of review emotional intensity: a text mining analysis of restaurant reviews. *International Journal of Hospitality Management*, *89*(5), 102558.

Li, H., Wang, C. R., Meng, F., & Zhang, Z. (2019). Making restaurant reviews useful and/or enjoyable? The impacts of temporal, explanatory, and sensory cues. *International Journal of Hospitality Management*, *83*, 257-265.

Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management*, *68*, 301-323.

Li, Y., & Xie, Y. (2020). Is a picture worth a thousand words? An empirical study of image content and social media engagement. *Journal of Marketing Research*, *57*(1), 1-19.

Liu, H., Feng, S., & Hu, X. S. (2022). Process vs. outcome: Effects of food photo types in online restaurant reviews on consumers' purchase intention. *International Journal of Hospitality Management*, *102*, 103179.

Liu, Z., & Park, S. (2015). What makes a useful online review? Implication for travel product websites. *Tourism Management*, *47*, 140-151.

Lochbuehler, K., Mercincavage, M., Tang, K. Z., Tomlin, C. D., Cappella, J. N., & Strasser, A. A. (2018). Effect of message congruency on attention and recall in pictorial health warning labels. *Tobacco Control*, *27*(3), 266-271.

Lopez, A., & Garza, R. (2021). Do sensory reviews make more sense? The mediation of objective perception in online review helpfulness. *Journal of Research in Interactive Marketing*.

Lu, X., He, S., Lian, S., Ba, S., & Wu, J. (2020). Is user-generated content always helpful? The effects of online forum browsing on consumers' travel purchase decisions. *Decision Support Systems*, *137*, 113368.

Ma, Y. F., Xiang, Z., Du, Q. Z., & Fan, W. G. (2018). Effects of user-provided photos on hotel review helpfulness: an analytical approach with deep leaning. *International Journal of Hospitality Management*, *71*, 120-131.

Marder, B., Erz, A., Angell, R., & Plangger, K. (2021). The role of photograph aesthetics on online review sites: effects of management-versus traveler-generated photos on tourists' decision making. *Journal of Travel Research*, *60*(1), 31-46.

McGrath, T. (2017). Social listening meets image sharing: A picture is worth 1,000 words. *Marketing Science Institute*.

Nielsen. (2012). *Nielsen Insights Reveal Smartphones account for half of all mobile phones, dominate new phone purchases in the US*. New York, NY. Retrieved from https://www.nielsen.com/us/en/insights/article/2012/smartphones-account-for-half-of-all-mobile-phones-dominate-new-phone-purchases-in-the-us/

Oliveira, B., & Casais, B. (2019). The importance of user-generated photos in restaurant selection. *Journal of Hospitality and Tourism Technology*, *10*, 2-14.

Paivio, A. (1990). *Mental representations: a dual coding approach*. Oxford University Press.

Pan, Y., & Zhang, J. Q. (2011). Born unequal: A study of the helpfulness of user-generated product reviews. *Journal of Retailing*, *87*(4), 598-612.

Park, S., & Nicolau, J. L. (2015). Asymmetric effects of online consumer reviews. *Annals of Tourism Research*, *50*, 67-83.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, *19*, 123-205.

PowerReviews. (2017). *The role of visual content and how to leverage it*. Retrieved from https://www.powerreviews.com/blog/the-role-of-visual-content-and-how-to-leverage-it/

Ray, A., & Bala, P. K. (2021). User generated content for exploring factors affecting intention to use travel and food delivery services. *International Journal of Hospitality Management*, *92*, 102730.

Ren, M., Vu, H. Q., Li, G., & Law, R. (2020). Large-scale comparative analyses of hotel photo content posted by managers and customers to review platforms based on deep learning: Implications for hospitality marketers. *Journal of Hospitality Marketing & Management*, 1-24.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, *5*(4), 296-320.

Sánchez-Fernández, R., & Iniesta-Bonillo, M. Á. (2007). The concept of perceived value: A systematic review of the research. *Marketing Theory*, *7*(4), 427-451.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, *61*, 85-117.

Schuckert, M., Liu, X., & Law, R. (2015). Hospitality and tourism online reviews: Recent trends and future directions. *Journal of Travel & Tourism Marketing*, *32*(5), 608-621.

Senecal, S., & Nantel, J. (2004). The influence of online product recommendations on consumers' online choices. *Journal of Retailing*, *80*(2), 159-169.

Severin, W. (1967). Another look at cue summation. *AV Communication Review*, *15*(3), 233-245.

Shin, D., He, S., Lee, G. M., Whinston, A. B., Cetintas, S., & Lee, K.-C. (2020). Enhancing social media analysis with visual data analytics: A deep learning approach. *MIS Quarterly*, *44*(4), 1459-1492.

Siering, M., Muntermann, J., & Rajagopalan, B. (2018). Explaining and predicting online review helpfulness: The role of content and reviewer-related signals. *Decision Support Systems*, *108*, 1-12.

So, K. K. F., Kim, H., & Oh, H. (2020). What makes Airbnb experiences enjoyable? The effects of environmental stimuli on perceived enjoyment and repurchase intention. *Journal of Travel Research*, 0047287520921241.

Stackla. (2021a). *Bridging the gap: Consumer & marketing perspectives on content in the digital age*. Retrieved from https://stackla.com/resources/reports/bridging-the-gap-consumer-marketing-perspectives-on-content-in-the-digital-age/

Stackla. (2021b). *User-generated content for travel and hospitality brands-Transform your travel marketing with user-generated content*. Retrieved from https://stackla.com/industry-solutions/travel-hospitality/

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., . . . & Rabinovich, A. (2015). Going deeper with convolutions. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1-9).

Topaloglu, O., & Dass, M. (2021). The impact of online review content and linguistic style matching on new product sales: The moderating role of review helpfulness. *Decision Sciences*, *52*(3), 749-775.

Uhrig, M. K., Trautmann, N., Baumgärtner, U., Treede, R.-D., Henrich, F., Hiller, W., & Marschall, S. (2016). Emotion elicitation: A comparison of pictures and films. *Frontiers in Psychology*, *7*, 180.

Van Rompay, T. J., De Vries, P. W., & Van Venrooij, X. G. (2010). More than words: On the importance of picture–text congruence in the online environment. *Journal of Interactive Marketing*, *24*(1), 22-30.

van Rompay, T. J. L., & Pruyn, A. T. (2008). Brand visualization: Effects of "product shape-typeface design" congruence on brand perceptions and price expectations. *Advances in Consumer Research*, *35*, 825-826.

Vogt, C. A., & Fesenmaier, D. R. (1998). Expanding the functional information search model. *Annals of Tourism Research*, *25*(3), 551-578.

Wang, Y. (2018). *Sensing human sentiment via social media images: Methodologies and applications* Arizona State University.

Wang, Y., & Song, J. (2020). Image or text: Which one is more influential? A deep learning approach for visual and textual data analysis in the digital economy. *Communications of the Association for Information Systems*, *47*.

Wu, L., Mattila, A. S., Wang, C.-Y., & Hanks, L. (2016). The impact of power on service customers' willingness to post online reviews. *Journal of Service Research*, *19*(2), 224-238.

Wu, P. F. (2013). In search of negativity bias: An empirical study of perceived helpfulness of online reviews. *Psychology & Marketing*, *30*(11), 971-984.

Xu, H., Liu, B., Shu, L., & Yu, P. S. (2019). Bert post-training for review reading comprehension and aspect-based sentiment analysis. *arXiv preprint arXiv:.02232*.

Yang, S. B., Shin, S. H., Joun, Y., & Koo, C. (2017). Exploring the comparative importance of online hotel reviews' heuristic attributes in review helpfulness: A conjoint analysis approach. *Journal of Travel & Tourism Marketing*, *34*(7), 963-985.

Yoo, J., & Kim, M. (2014). The effects of online product presentation on consumer responses: A mental imagery perspective. *Journal of Business Research*, *67*(11), 2464-2472.

Zhang, K., Chen, Y., & Li, C. (2019). Discovering the tourists' behaviors and perceptions in a tourism destination by analyzing photos' visual content with a computer deep learning model: The case of Beijing. *Tourism Management*, *75*, 595-608.

Zhang, M., & Luo, L. (2018). Can user generated content predict restaurant survival: Deep learning of yelp photos and reviews. *Social Science Research Network*, *3108288*.

Zhang, Y., Gao, J., Cole, S., & Ricci, P. (2020). How the spread of user-generated contents (UGC) shapes international tourism distribution: using agent-based modeling to inform strategic UGC marketing. *Journal of Travel Research*, 1-23.