



Are customer star ratings and sentiments aligned? A deep learning study of the customer service experience in tourism destinations

Enrique Bigne¹ · Carla Ruiz¹ · Carmen Perez-Cabañero¹ · Antonio Cuenca¹

Received: 11 January 2022 / Accepted: 23 January 2023 / Published online: 6 February 2023
© The Author(s) 2023

Abstract

This study explores the consistency between star ratings and sentiments expressed in online reviews and how they relate to the different components of the customer experience. We combine deep learning applied to natural language processing, machine learning and artificial neural networks to identify how the positive and negative components of 20,954 online reviews posted on TripAdvisor about tourism attractions in Venice impact on their overall polarity and star ratings. Our findings showed that sentiment valence is aligned with star ratings. A cancel-out effect operates between the positive and negative sentiments linked to the service experience dimensions in mixed-neutral reviews.

Keywords Sentiment analysis · Deep learning · Artificial neural networks · Tourism destination · Star rating

1 Introduction

Tourism destination services' practitioners and academics agree that focusing strongly on the customer experience may create a sustainable and unique advantage for tourism brands/destinations (Kim and So 2022; Sorokina et al. 2022). Mature destinations may be unable to differentiate themselves from

✉ Enrique Bigne
enrique.bigne@uv.es

Carla Ruiz
carla.ruiz@uv.es

Carmen Perez-Cabañero
carmen.perez-cabanero@uv.es

Antonio Cuenca
antonio.cuenca@uv.es

¹ Faculty of Economics - Department of Marketing, University of Valencia, Av. Naranjos S/N, 46022 Valencia, Spain

similar vacation options, and the quality of the customer service experience may be affected by, among other factors, the negative consequences caused by the high number of visitors they receive (Yu and Egger 2021). In this context, online reviews are drivers of destination choice (Bigne et al. 2021; Yu and Egger 2021) as they reflect many experiences regarding different aspects of the reviewed service (Siering et al. 2018). Attractions are the core products of mature tourism destinations. Unlike hotels, which have clear attributes (i.e., room size, services, location), attractions are complex and contextual. Customer experience is particularly important for tourism attractions because consumers spend time and money on them and are highly involved in the consumption of the attractions and their complementary services (Simeon et al. 2017).

Review star ratings are the number of stars allocated by reviewers to their online reviews, indicating their assessment of the products/services consumed. Readers very often use star ratings as an important heuristic through which to narrow down their consideration set, thus reducing cognitive effort and search costs (Dhar and Bose 2022). Online reviews have several functions: they express the reviewers' emotions, describe real experiences, offer recommendations and provide diagnostic information to other consumers (Taecharungroj and Mathayomchan 2019). The type of information they provide sets the emotional tone of online reviews. Objective reviews tend to reflect cognitive aspects of the customer experience, while subjective reviews reflect affective aspects. Online reviews provide "subjective opinions" based on personal views, and emotions, about products/services and destinations, and "objective statements" based on facts, evidence and measurable observations, or a mixture of both (De Keyzer et al. 2017; Feldman 2013). Recent research has highlighted the significant influence of the emotional tone of online reviews on perceived review helpfulness (Bigne et al. 2021; Craciun et al. 2020; De Keyzer et al. 2017); however, the relationships between emotions expressed in online reviews and star ratings are still underexplored. Examining the relationship between the emotional tone of the reviews and customers' star ratings can help Destination marketing organizations (DMOs) and tourism companies to better design feedback systems to improve the quality of information received and thus, to enhance their attractions based on customers' online textual reviews and ratings (Li et al. 2019; Yoon et al. 2019; Zhang et al. 2016). The relationship between the emotional tone of online textual reviews and customers' ratings also influences future tourists' demands because customers tend to read both textual reviews and ratings to assess their consistency (Zhao et al. 2019). To bridge this research gap, this paper assesses the consistency between the emotional tone of service reviews and their star ratings.

It is generally assumed that star ratings are a numeric representation of the text of online reviews, and that their valences are consistent. However, this maybe not always be true. Consumers may write negative comments despite awarding 4 or 5 stars in a TripAdvisor review, out of a desire to help other consumers (Valdivia et al. 2019). As the star rating and the valence of textual components may be inconsistent, the exact nature of this relationship should be clarified. This is a very important issue given the effect of sentiments and star ratings on consumers' purchase intentions and attitudes towards tourism services (Bigne et al. 2021; Pike et al. 2021; Sayfuddin et al. 2021), the experiential and intangible nature of services and because

consumers use overall star rating as key heuristic cues when narrowing down service choices (Dhar and Bose 2022; Racherla et al. 2013).

In addition, this study aims to identify the most important service aspects underlying the dimensions of the customer experience for positive (4–5 stars), negative (1–2 stars) and neutral-mixed (3 stars) reviews of tourism destinations. Previous studies have treated service experiences as either positive or negative constructs and have overlooked the relationship between the online reviews and the specific dimensions of the service experience. However, consumers often have mixed feelings about the attributes of a service experience. The customer experience with tourism attractions is derived not only from their core features, but also from their supporting facilities, accommodation, transport and interactions with service employees and other customers (Kandampully et al. 2018). For example, a tourist visiting the Doge's Palace may post positive comments about the architecture and the paintings, but post negative comments about the long queues, local transport and/or the entry fees. Therefore, to gain deeper knowledge of the dimensions of the customer experience included in reviews with different star ratings we pose the following research question: *RQ1. Which are the most important service aspects underlying the dimensions of the customer experience for positive (4–5 stars), negative (1–2 stars) and neutral (3 stars) reviews of tourism destinations?* This question arises because the textual content of the reviews incorporates information that is not always reflected by star ratings (Bigne et al. 2021). For instance, knowing what attraction visitors talk about and the emotional tone of such comments advances our understanding of how customers evaluate the attributes of the service delivered by tourism attractions (e.g., value for money, staff, sensory experiences).

Although service managers have recognised the value of online reviews, they overlook the large number of neutral reviews and, thus, do not always include them in their social media monitoring metrics. Therefore, an important knowledge gap remains regarding the effects of neutral user-generated content (UGC) on consumer behaviour, as previous studies have worked on the basis that neutral reviews are less useful than extreme positive and negative reviews (Liu and Park 2015). However, mixed-neutral reviews (reviews containing both positive and negative aspects) are perceived as diagnostic, that is, they increase the consumer's motivation and ability to process positive and negative reviews to reduce this conflict (Tang et al. 2014). Therefore, a pertinent research question is what is the effect of the individual dimensions of the dimensions of the customer service experience on the overall sentiment polarity of mixed-neutral reviews.

In sum, the influence of consumers' reviews on other consumers' perceptions and decisions prior to purchase has been extensively examined, but there is still a gap in the literature as to how star ratings, which reflect their post-purchase evaluations, are influenced by the emotional tone of reviews (objective versus subjective content) and how they relate to the different components of the consumer service experience. This study, therefore, aims to bridge these gaps in the literature and to offer new academic and practical insights into how best to manage star ratings (e.g., how to obtain a five-star rating) and, in turn, improve the customers' online experience. Specifically, this research analyses: (i) the consistency between the emotional tone of service reviews and star ratings; (ii) the relationships between the sentiment valence

and the overall polarity of online reviews, segmented by number of stars awarded; (iii) identifies the most important service aspects underlying the dimensions of the customer experience for positive (4–5 stars), negative (1–2 stars) and neutral (3 stars) reviews of tourism destinations and (iv) analyses the influence of positive and negative comments expressed about service-based aspects on the overall polarity of mixed-neutral reviews.

We address these issues empirically using consumer review data collected from 20,954 reviews posted on TripAdvisor. Specifically, we examine the relationships between consumers' ratings and the valence of the text components of online reviews of three tourism attractions in a mature destination (Venice), by combining deep learning applied to natural language processing (NLP), machine learning (ML) and artificial neural networks (ANN), to identify how the positive and negative components of online reviews impact on overall polarity and star rating. Deep learning has recently emerged as a powerful ML approach for analysing digital content based on NLP.

1.1 Literature review

1.1.1 Effects of star ratings and review sentiment on consumer behaviour

Information search theories suggest that mechanisms that provide a “signal” as to what a piece of information might contain can be important aids for consumers trying to decide what information to use (Al-Natour and Turetken 2020). A star rating is a signal of the consumer's evaluation of his/her experience with a tourism service (Yoon et al. 2019). Star ratings which provide ordinal information (1 to 5) of consumer satisfaction are easy to process by consumers, who use them as signals to assess the quality of other consumers' service experiences and to make decisions (Yoon et al. 2019). However, written comments about destinations provide tacit context-specific explanations of the reviewer's feelings, experiences and emotions, which go beyond numeric ratings (Bigne et al. 2021).

Recent studies have provided plenty of evidence for the positive effect of rating scores and the sentiment valence of online review content on the profitability of tourism services. Yang et al. (2018) synthesised 25 studies in the tourism and hospitality industry and revealed that the financial performance of hotels is affected by both customer rating scores and the number of online reviews. Nieto-Garcia et al. (2019) found that hotel revenues are affected by different rating attributes, such as staff and facilities. Luca (2016) found that a 1-star increase in customer ratings on Yelp led to an increase of 5–9% in restaurant revenues. Sayfuddin and Chen (2021) found that customer ratings are positively associated with the revenues of hotels listed on TripAdvisor.com. Pike et al. (2021) highlighted the importance of the influence of positive online reviews on attitudinal loyalty towards Dubai as a stopover destination. As star ratings and review content are two major information sources for consumers in their decision-making, many studies have examined the effects of star ratings and sentiments on consumer behaviour. Indeed, star ratings in reviews of tourism services have been found to significantly influence booking intentions (Sayfuddin

et al. 2021) and positive brand attitudes (Yang et al. 2018). The sentiment valence of online reviews (positive or negative) also drives consumers' purchase decisions (Baniya et al. 2021; Pike et al. 2021). However, some scholars have argued that a review's content is more influential than its overall star rating because consumers rely more on sentiment when making purchase decisions and because star ratings do not contain detailed evaluations (Hu et al. 2012; Racherla et al. 2013).

1.1.2 Emotional tone of the review and star ratings

Reviewers may report their service experience using an objective (factual) or subjective (emotional) tone. The present study proposes that the tone of online reviews is an element of message content (De Keyzer et al. 2017). Some messages are predominantly objective, that is, they are based on information about specific attributes, such as "*Lines in Doge's Palace can be long and waiting times up to a few hours if you decide to go there between noon and 7 pm*". The arguments used in factual reviews are rational, objective, specific and clear. Other messages are predominantly emotional, focusing on the feelings evoked in the writer by the service experience, with no or little support from verifiable arguments (e.g., "*Loved the 2 h we spent at the Doge's Palace, enjoyed lots of interesting art, design and excellent examples of how powerful the Doge was*"). Emotional messages are often subjective and abstract, containing both relevant and non-relevant information about the service attributes.

Customers writing objective reviews often compare recent experiences with past experiences and, thus, are more rational. Zhao et al. (2019) argued that review subjectivity has a negative effect on star ratings. Customers often consider online review platforms as places to complain about their consumption experiences. Customers writing subjective reviews are more emotional and thus tend to generate more extreme, negative evaluations of tourism services when the product or service offerings don't fulfill their expectations (Schoefer and Ennew 2005). Bad experiences drive reviewers to post more emotional words/details that reflect their negative perceptions, which makes the material posted more subjective. These subjective complaints caused by customer dissatisfaction with attractions are reflected in low star ratings.

The effect of the emotional tone of a review on its star rating is also grounded in schema theory (Brewer and Nakamura 1984). For mature destinations, with well-known attractions, tourists usually have pre-constructed schemas based on their previous experience. Therefore, when they enjoy a good service experience this confirms their schema and, thereafter, they give a high star rating to the tourism attraction and share their experiences based on facts, evidence and measurable observations. However, if they have gone through an unsatisfactory service experience, they share their experience based on their personal opinions, feelings and judgments about the attraction they have visited and give a low star rating (Zhao et al. 2019). Therefore, we propose:

H1a Positive reviews feature a higher (lower) percentage of objective (subjective) comments than do mixed-neutral reviews.

H1b Negative reviews feature a lower (higher) percentage of objective (subjective) comments than do mixed-neutral reviews.

1.1.3 Valence of review sentiment and star ratings

The valence of online reviews, as established by their textual components, and star ratings may not coincide due to their basically different natures: the valence of textual components is descriptive, while star ratings are quantitative. This inconsistency is often observed in practice, as some star ratings are positive, while the valence of the textual content is negative (Valdivia et al. 2019). Some studies have confirmed the existence of inconsistency between star ratings and the underlying sentiments of online reviews (Luo and Xu 2021; Racherla et al. 2013; Sharma et al. 2020; Valdivia et al. 2019). However, other studies (Baniya et al. 2021; Gaur et al. 2021; Geetha et al. 2017; Yoon et al. 2019; Zhu et al. 2020) have found that there is a positive relationship between star ratings and textual content. Table 1 shows how recent research into whether consumers' comments about their service experiences reflect their overall star ratings has reported inconclusive results. Our study aims to resolve this conflict in the literature by exploring the relationship between star ratings and the sentiment of service reviews.

Sentiment analysis assigns polarities, positive, neutral and negative to opinions, and obtains valuable information through text analysis (Bigne et al. 2021). Customer sentiment polarity has been defined as the ratio between the number of positive words and negative words in a given review (Bigne et al. 2021). Customer sentiment is the total amount of sentiment that exists in a text, both positive and negative. Polarity is the direction of this sentiment, that is, positive, negative or neutral. Online ratings given to attractions can be considered consistent when they match the underlying customer sentiments of reviews. As consumers assign star ratings and compose their textual components at the same time, in line with previous studies we expect consumers to be consistent in the ratings and the textual components they provide. That is, the more positive the text of a review is, the more stars it will receive (Baniya et al. 2021; Gaur et al. 2021; Geetha et al. 2017; Yoon et al. 2019; Zhu et al. 2020). Thus, the valence of textual components should have a positive relationship with star ratings. Therefore, we posit:

H2a The overall polarity of a review of a tourism attraction has a positive relation to its star rating.

H2b/c The number of positive (negative) words in a review of a tourism attraction has a positive (negative) relation to its star rating.

Negativity bias theory (Kanouse and Hanson 1987) suggests that people put more weight on negative experiences than they do on positive experiences. Negative events tend to be more vivid in the mind, and are easier to recall, than positive experiences (Sharma et al. 2020). Prospect theory (Kahneman and Tversky 1979) argues that people have a stronger tendency to avoid losses (loss aversion) than to seek

Table 1 Review of recent studies on the relationship between sentiments and star rating

Authors	Research topic	Method	Findings
Geetha, M., Singha, P., Sinha, S. (2017)	To establish a relationship between customer sentiments in online reviews and customer ratings for hotels	Sentiment analysis of reviews of a random sample of 20 hotels located in India in the budget category, and 20 from the premium category	Customer sentiment polarity explains significant variation in customer ratings across premium and budget hotel categories
Yoon, Y., Kim, A. J., Kim, J., Choi, J. (2019)	To identify: (1) how a review's textual components relate to the star rating, (2) how these components influence the ratings given by other reviewers	1900 reviews in the restaurant category from TripAdvisor.com, examining the number of stars that a review receives	A star rating has a positive relationship with the overall review's valence. That is, the more positive a review is, the greater number of stars a review receives. The reviews of other consumers also play a role in determining the star rating of the focal review, suggesting social influence among consumers
Valdivia, A., Hrabova, E., Chaturvedi, I., Luzón, M. V., Troiano, L., Cambria, E., Herrera, F. (2019)	To analyse if inconsistencies exist between user polarity and sentiment analysis methods that automatically extract polarities	88,882 reviews posted during 2012–2016 of six Italian and Spanish tourism attractions	Inconsistencies exist between user polarity and sentiment analysis methods that automatically extract polarities
Hong & Pittman (2020)	To analyse how star ratings, number of reviews and review valence provide cues to consumers, which, in turn influenced the perceived credibility of online reviews	Study 1: 105 students (74% female) recruited from undergraduate courses at a large Midwestern university in the USA Study 2: 99 (99.57% female) non-student participants Experimental design: A 2 (star rating: high vs. low) × 2 (number of reviews: high vs. low) × 2 (product review valence: positive vs. negative) × 2 (multiple messages: nutrition bars vs. multivitamins) mixed subject experimental design	Negatively valenced reviews tend to supersede star ratings, with the number of reviews serving as a proxy for argument strength. When a review was positive, the participants trusted the star ranking system, perceiving a high star ranking to be more credible

Table 1 (continued)

Authors	Research topic	Method	Findings
Sharma et al. (2020)	To analyse the relationship between sentiments and online ratings by introducing the prospect theory	57,036 reviews of 20 US airlines retrieved from TripAdvisor	Negative deviations in ratings had a higher impact on customer sentiment than positive variations of equal magnitude
Siddiqi, U. I., Sun, J., & Akhtar, N. (2020)	To examine: (1) how conflicting customer star ratings (i.e. heuristic cues) and opinions about hotel attributes (i.e. systematic cues) engender attitude ambivalence, (2) how offline interpersonal informational influence moderates the relationship between conflicting reviews and attitude ambivalence, and (3) relational impact on purchase intentions	Convenience sample of 382 inbound tourists in Beijing, China	Positive effect of conflicting star ratings and opinions about hotel attributes on attitude ambivalence, which lowers consumers' purchase intentions. The study also found a significant role of offline interpersonal informational influence as a moderator of the relationship between conflicting reviews and attitude ambivalence
Zhu et al. (2020)	To examine the relationship between guests' sentiments and online ratings in the context of peer-to-peer accommodation	4602 reviews of San Francisco on the Airbnb platform	Positive (negative) sentiment was linked to high (low) ratings
Baniya, R., Dogru-Dastan, H., & Thapa, B. (2021)	To examine visitor experience via sentiment and topic analysis, given the popularity and volume of visitors annually	32,394 online reviews about Angkor Wat posted between January 2015 and October 2019 on TripAdvisor. Separate analyses were conducted for reviews with one, two and three-star ratings, and for four- and five-star ratings	Topic analysis showed that sunset and sunrise experiences, attraction structure, guided tours and temple experiences attracted mainly positive sentiments. Conversely, overcrowding, persistent selling, clothing styles and expense were seen as negative

Table 1 (continued)

Authors	Research topic	Method	Findings
Gaur, L., Afaq, A., Solanki, A., Singh, G., Sharma, S., Jhanjhi, N., D. (2021)	This paper aims to understand consumers' preferred hotel attributes, guest demands and the concerns of global hotel chains	45,500 online reviews of customers of 50 global hotel chains, taken from different online review sites, analysed using machine learning algorithms, text mining, and sentiment analysis techniques	General review star ratings correspond with the opinion (sentiment) scores for the title and the text of the online reviews
Sayfuddin, A. T. M., & Chen, Y. (2021)	To examine whether customer ratings and online reviews affect hotel revenues and, if so, to quantify the effects	376,060 online reviews for 1,348 hotels, based on their monthly taxable revenues between January 2014 and December 2017	The results showed that the reputational effect of a 1-star increase enhances revenues by 2.2–3.0%, whereas the reputational effect of a 1-star increase enhances revenues by 1.5–2.3%

gains; thus, negative online reviews are more diagnostic than positive online reviews because perceived loss has a bigger impact on preferences and evaluations than do perceived gains. The reference-dependent model of loss aversion (Tversky and Kahneman 1991) proposes that loss-framed arguments (i.e., arguments that address losses and disadvantages) have greater impact on preferences than gain-framed arguments (i.e., arguments that address gains and advantages). Based on prospect theory and negativity bias theory, we posit that the influence of the negative words in a service review on its star rating will be higher than the influence of the positive words.

H2d The relationship between number of negative words and star rating is stronger than the relationship between number of positive words and star rating.

1.1.4 Customer service experience and mixed-neutral reviews

The term ‘experience’ is often used to refer to product offerings in service settings that involve hedonic consumption, for example, in travel, restaurants, hotels and the arts (Holbrook and Hirschman 1982). However, there is still an open debate about the conceptualisation and measures of the customer experience. The experiential concept has been widely examined in tourism studies, given that experience is the core benefit consumers derive from service-oriented offerings (e.g., Baniya et al. 2021; Kandampully et al. 2018; Simeon et al. 2017). Recent hospitality and tourism literature (e.g. Rather et al. 2022) has defined customer experience as a customer’s cognitive, emotional, behavioural, sensorial and social responses to a firm’s offerings during his/her entire purchase journey.

A review of the literature suggests that the customer experience has various components (Rather et al. 2022; Kim and So 2022). Homburg et al. (2017) proposed a five-dimensional customer service experience model comprising sensorial, affective, cognitive, relational and behavioural facets, and Brakus et al. (2009) a four-dimensional model with intellectual, affective, sensory and behavioural experiences. Schmitt (1999) argued that customers may be affected by sensory, affective, behavioural, intellectual and social experiences. Verhoef et al. (2009) viewed the customer experience as a multidimensional construct comprising cognitive experiences (THINK), sensory experiences (SENSE), affective experiences (FEEL), physical experiences, behaviours and lifestyles (ACT) and social-identity experiences (RELATE).

Tourist experiences are recognised, by their complex nature, as being multi-dimensional (Kim and So 2022). Scientific discussions on destination brand experiences evolved when Barnes et al. (2014) and, later, others (e.g. Rather et al. 2022) adopted Brakus et al.’s brand experience scale to measure experiences during visits to tourist destinations. Ketter (2018) analysed the application of the experiential modules in destination marketing campaigns and concluded that Brakus’s approach can serve as an analytical framework through which to analyse destination marketing.

In this study we follow Homburg et al.’s (2017) five-dimensional conceptualisation of an experience construct: sensory (stimulation of any of the five senses: sight, hearing, touch, taste and smell), affective (emotions elicited through interaction with

the attraction that make the experience memorable), behavioural (physical interactions during the visit to the attraction), social (interactions with staff and other customers) and cognitive (problem-solving and intellectual experiences). We posit that the customer experience is derived not only from the core experience at the attraction (cognitive experience), but also from the emotions elicited during the visit, sensory experiences (e.g., vistas, music) and the transport used to get to and move between attractions, and interactions with service employees and other customers when visiting the attraction.

Positive (4–5 stars) and negative (1–2) star ratings reflect attitude extremity, that is, the deviation from the midpoint of an attitude scale. Past research has identified two reasons why customers post a midpoint rating, that is, three stars out of five (Thornton 2011; Tang et al. 2004). A 3-star review might reflect indifference, or a cancel out effect of the positive and negative feelings expressed/felt about the different dimensions of the customer experience. This two-dimensional view provides a theoretical rationale for distinguishing between the two types of neutral UGC: (i) indifferent-neutral, containing neither positive nor negative terms, expressing no dominant attitude or subjective preferences; (ii) mixed-neutral, with similar amounts of positive and negative terms, reflecting balanced evaluations, attitudes, and/or emotions. In both cases, a midpoint rating has been shown to be a legitimate measure of a middle-ground attitude.

The *accessibility-diagnostics* (AD) model (Feldman and Lynch 1988) explains how people form attitudes that guide behaviours (or proximate determinants of behaviours, such as judgments) based on the accessibility and diagnostic characteristics of inputs. In the case of tourist attractions, the service experience may generate conflicting sentiments based on different aspects of the service. For instance, reviewers may be very satisfied with the attraction guide, but may also have negative feelings about the value for money of the attraction, or its long queues, which cancel out the positive feelings, leading to a neutral attitude. The AD model proposes that reviewers' judgments of a service are likely to be influenced by sentiments formed about salient individual service-specific aspects that readily come to mind when making recommendations. In line with the AD model, we posit that positive and negative sentiments related to different aspects of a service are evaluated separately in the consumer's mind and cause him/her to develop both positive and negative evaluations and feelings towards his/her experiences and, thus, develop neutral attitudes (mixed-neutral reviews).

Compensatory models (Johnson and Meyer 1984) also provide theoretical grounds that support the proposition that mixed-neutral reviews may be posted because of the asymmetric effects of positive and negative experiences related to service-specific aspects. Compensatory models propose that consumers make trade-offs, by a type of linear compensation, between attributes when evaluating products/services. In a non-compensatory decision-making process, some attributes are not considered. This perspective assumes that only important attributes are considered because consumers limit the cognitive effort they expend. Extending this approach to online star ratings, in a compensatory model the overall star rating should reflect the summary effect of the content of posts covering different service-specific aspects. However, if consumers follow a non-compensatory process, only some

attributes determine the overall rating. Bigne et al. (2020) analysed compensatory versus non-compensatory rating behaviours and found, under specific conditions, that compensatory approaches had been taken. Guo et al. (2017) analysed 266,544 online reviews, using natural language processing, and found that 5 of 19 main dimensions of customer satisfaction are key determinants of overall customer ratings. Therefore, the valence of the sentiments expressed about service-based aspects might affect numerical star ratings in different ways due to compensatory information processing.

In line with the AD model and compensatory models, we posit that mixed-neutral reviews may contain positive and negative sentiments, linked to different dimensions of the customer service experience, that cancel each other out. Therefore,

H3 For mixed-neutral reviews (3 stars), the positive and negative sentiments related to the (a) cognitive, (b) social interaction, (c) sensory, (d) affective and (e) behavioural dimensions of the service experience impact on the overall sentiment of the review, such that they cancel each other out.

The conceptual model is depicted in Fig. 1.

2 Methodology

2.1 Sample description

We crawled all the English language reviews (20,617) posted on TripAdvisor during 2015–2019 about the three Venetian attractions with the highest number of online reviews, St. Mark’s Square, the Grand Canal and the Doge’s Palace. Tourists rate these as the city’s top 3 attractions, with an overall rating of 4.5/5. The data were crawled during February 2020. For each comment we analysed the full text of the review, individual rating and heading. The stars associated with the reviews of these three attractions are shown in Table 2.

Table 2 shows that the three attractions mostly received positive reviews, 91.8% achieving four or five stars. The number of reviews for each attraction was balanced.

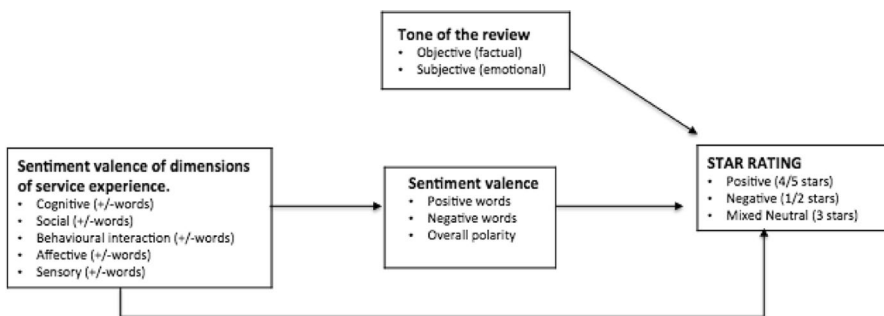


Fig. 1 Conceptual model

Table 2 Number of reviews, by star ratings, given to the three Venetian attractions

Attraction	1–2 stars	3 stars	4–5 stars	Total reviews %
The Grand canal	71	241	6,969	7281 35.3%
St. Mark's square	198	792	6,536	7526 36.5%
The Doge's palace	103	283	5,424	5810 28.2%
Total reviews %	372 1.8%	1316 6.4%	18,929 91.8%	20,617 100%

Star rating: mean = 4.57; standard deviation = 0.722

Most reviews that awarded 3 or less stars were of St. Mark's Square, while the 4- and 5-star reviews were distributed over the three attractions.

2.2 Deep learning

The objective/subjective classification of the texts was addressed as a binary classification problem powered by a deep neural network. Advanced ML methods based on DL algorithms have shown high accuracy in determining the polarity of online reviews (Chang et al. 2020). Recent studies have shown that good results can be achieved in sentiment analysis-focused tourism research by using a combination of DL models and NLP. Kim and Park (2017) applied a DL approach based on Stanford sentiment analysis to analyse the sentiment polarity of reviews of Paris. Al-Smadi et al. (2018) showed that deep recurrent neural network approaches (RNN) outperform other ML methods in polarity identification. Chang et al. (2020) demonstrated that combining DL and NLP tools can assist hotels in decision-making by prioritising which reviews need responses.

The first step in classifying the texts was to generate in-domain word embeddings (Rudkowsky et al. 2018). These word embeddings are fed into a recurrent neural network, a long short-term memory (LSTM) which uses all the words in a given text (i.e., online review), in the order they appear in the text, to build a dense representation of a whole document. The document representation is, thereafter, converted into two values which, after passing through a softmax layer, describe a probability distribution over two possible classes (i.e., being objective vs. being subjective). For training purposes two researchers manually labelled a set of customer comments as being subjective or objective. From these manually labelled comments, the algorithm automatically balanced both classes to avoid biasing the classifier towards the more frequently used class. From the resulting balanced set of labelled comments, 80% were used as training data, while the remaining 20% were retained as evaluation data to assess if the classifier kept learning and improving its predictions about unseen data after each training epoch. In our experiments more than 70% of the test comments (not used for the training) were correctly classified as objective/subjective based on the gold annotations, which is reasonably high considering the subjective nature of the task. When the training was completed, the resulting model was stored

and used to predict the most probable class of new unseen examples that used the outputs of the previous layer to predict the likelihood of the text being objective or subjective. Of the comments, 66.1% were classified as objective, and 33.9% as subjective; Table 3 provides an example of an objective and a subjective comment.

Following the application of the deep-learning technique, we analysed the relationships between the emotional tone of the reviews (objective versus subjective) and their star ratings.

2.3 Sentiment analysis

We carried out an automatic sentiment analysis using deep learning, a class of ML technique applied to natural language processing (Deng and Yu 2014; Timoshenko and Hauser 2019). We used software specifically developed for the service industry. This is based on free open-source tools available in OpeNER (<https://www.opener-project.eu/project/>), an NLP platform (García-Pablos et al. 2016), and deep-learning open-source tools. The sentiment polarity analysis has several steps. First, the data (text) were extracted and cleaned. The languages used were automatically detected and unwanted languages were filtered out to avoid noise. The texts were then segmented into sentences and tokens which, broadly speaking, are words and punctuation marks. Following this, part-of-speech tagging (PoS tagging) and lemmatisation processes were carried out. The PoS tagging process consists of determining the corresponding morphosyntactic category for a word, given its context (e.g., in the sentence “amazing experience”, “experience” is a noun, and “amazing” is an adjective). The lemmatisation process obtains the canonical form of a word as it would appear in a dictionary. Both PoS tagging and lemmatisation help reduce the target vocabulary and simplify later processes. Next, the polarity of each sentence

Table 3 Examples of objective and subjective reviews

OBJECTIVE REVIEW	
Dimensions	<i>This is a “must see before I die” attraction!!! (5 stars)</i>
-Cognitive (Cultural heritage)	This place shows just how much wealth and power there used to be in Venice. Obviously, it’s Venice, so there are loads of tourists everywhere, which means getting into the palace is not an easy task. Lines can be long and waiting times are up to a few hours if you decide to go there between noon and 7 pm. Luckily, it opens early and closes late (I think it’s open till 10 pm), so aiming at early hours or late hours (the risk with the late hours is that your time will be limited if you go too late) would be your best bet!
-Social (Overcrowding)	
-Activities (Behavioural)	
SUBJECTIVE REVIEW	
Dimensions	<i>Midday visit (3 stars)</i>
-Cognitive (Cultural heritage)	Midday visit. No problem with any queue/line, walked straight in and very few in the palace. Loved the 2 h we spent there, enjoyed lots of interesting art, design and excellent examples of how powerful the Doge was. Slight negative in that not all areas were open and if you have mobility issues it could be difficult. A positive to finish, the visit to the prison, feeling the atmosphere and the walk across the bridge
-Social (Overcrowding)	
-Activities (Behavioural)	

was calculated using Bing Liu's (2010) well-known English polarity lexicon. As this lexicon is domain independent, some manual revision was undertaken to discard irrelevant words. In this step, negation is considered as a polarity shifter, which means that words in the scope of a negation particle represent the opposite polarity (e.g., "great" is positive but, in the scope of the negation "not", as in "not great", it becomes negative). The overall sentiment polarity of sentences was calculated using the ratio of positive (or negative) words detected in them, after taking into account negation. A further step applied a similar dictionary-based approach to classify each piece of text into a set of predefined categories/topics.

Following previous works (Homburg et al. 2017; Simeon et al. 2017; Bigne et al. 2021; Yu and Egger 2021; Rather et al. 2022; Kim and So 2022) several topics were chosen to analyse the 5 dimensions (cognitive, emotional, social, sensory and behavioural) of the customer experience: (i) the behavioural dimension captures topics linked to interactions with the physical environment of the attraction (recommendations on activities to undertake or to avoid, and transport); (ii) the social dimension is linked to the social interaction with residents and other customers visiting the attraction; (iii) the cognitive dimension captures intellectual experiences and problem-solving issues (cultural heritage and value for money); (iv) the affective dimension captures emotions elicited through interaction with the attraction that make the experience memorable); (v) the sensory dimension captures topics related to consumers' feelings evoked through the five senses by the attraction. We preferred manually curated resources over automatically generated (e.g., using topic-modelling techniques, such as latent Dirichlet allocation, LDA: Osmani, et al. 2020), because user-generated content, being informal text, is prone to generate a lot of noise and the resulting topics are difficult to control and require additional curation. Each comment was assigned to up to three topics. Table 4 shows the topics linked to the dimensions of the customer experience and provides several words about each topic as examples.

2.4 Artificial neural networks

We applied artificial neural networks to examine the inter-relationships between star ratings and sentiments. Artificial neural networks (ANNs) have been widely used as a statistical technique in services' research (see Palmer et al. 2006). ANN use has greatly increased in the field of business in the last twenty years (Tkáč and Verner 2016). Their key advantages are: (i) they are suitable for analysing inter-relationships between variables in nonlinear relationships (Aakash et al. 2021; Bloom 2005; Phillips et al. 2015; Uysal and El Roubi 1999); (ii) they overcome model misspecification (Aakash et al. 2021), and (iii) they avoid multicollinearity (Coelho et al. 2013). Overall, ANNs perform better than multiple regression models (Uysal and El Roubi 1999). The present study applies ANNs to social media (Phillips et al. 2015). A multilayer perceptron (MLP) is a class of feedforward ANN that uses a supervised learning technique called backpropagation learning, while radial basis function (RBF) depends only on the distance between the input variables and the dependent variable. Using ANNs we estimated the influence of each aspect of the

Table 4 Dimensions of the customer experience: topic, definition and keywords

Identified topic	Definition	Keywords (example)
Dimension 1: cognitive		
Cultural heritage	Visitors' impressions of the cultural and historical background of the attraction	story, shape, bell, tower, museum
Value for money	Visitors' impressions of the value they get in return for the money spent at the attraction	day ticket, worth, euro, expensive
Dimension 2: social interaction		
Social interaction with local residents and staff	Visitors' impressions of the interaction with local residents and staff working in services at the attraction	rude locals, rude Italians, great people, friendly, helpful, pushy
Overcrowding	Visitors' perceptions of space congestion at the attraction generated by high density of people	sea of tourists, tourist hordes, overcrowded, a lot of people, busy
Dimension 3: behavioural interaction with the attraction		
Transport	Visitors' impressions of the transport available to arrive at/move between attractions	public transport, bus, train, ship
Activities	Visitors' recommendations of activities to undertake or avoid when visiting the attraction	walk, book tour, spend time, visit prison
Dimension 4: affective		
Memorable experience	Visitors' emotions experienced throughout their visit to the attraction that make it a memorable experience	amazing, remarkable, unique, disappointing
Dimension 5: sensory		
Sense	Visitors' feelings evoked through the five senses by the attraction	live music, view, feed pigeons, tasty

customer experience with tourism attractions on the overall sentiment of mixed-neutral reviews (3 stars). Analysing these neutral rating scores through ANNs provides helpful insights. ANNs are effective in explaining 3-star ratings scores (neither positive nor negative) because the algorithm learns from the dataset to identify patterns.

3 Results

3.1 Hypotheses testing

To test H1a and H1b we undertook a crosstab analysis and applied the Chi-square statistic to verify if the relationships observed between the variables were statistically significant (see Table 5).

The results (Table 5) showed that the majority of reviews contained objective messages, independent of their ratings. However, this majority was smaller in the negative (1–2 stars) and neutral (3 stars) reviews. Positive reviews (4–5 stars) had a higher percentage of objective comments (66.9%) than did mixed-neutral reviews (57.4%). Positive reviews contained a lower percentage of subjective comments (33.1%) than did the mixed-neutral reviews (42.6%). Thus, H1a is supported. Negative reviews (1–2 stars) contained a lower percentage of objective comments (54%) than did mixed-neutral reviews (57.4%). Negative reviews contained a higher percentage of subjective comments (46%) than did mixed-neutral reviews (42.6%). Thus, H1b is supported.

To test H2 we analysed the relationships between the star ratings of the reviews and their overall polarities (H2a), the number of positive (H2b) and negative (H2c) words they contained and compared the means of the negative, mixed-neutral and positive reviews by applying a one-way ANOVA. The descriptive statistics of these variables are shown in Table 6.

Table 6 shows that overall polarity was higher for positive reviews (4–5 stars) than for mixed-neutral (3 stars) and negative reviews (1–2 stars). There were more positive words in reviews with 4 and 5 stars (mean = 3.685, SD = 2.515), while there were more negative words in reviews with 1 and 2 stars (mean = 2.314, SD = 2.149). Thereafter, we tested if these means were statistically different by means of a one-way ANOVA. First, we checked if there was homogeneity in variances between the

Table 5 Number of star ratings awarded by tone of the reviews

Number of reviews	1–2 stars	3 stars	4–5 stars	Total
Objective reviews	201	755	12,668	13,624
%	54.0%	57.4%	66.9%	66.1%
Subjective reviews	171	561	6,261	6,993
%	46.0%	42.6%	33.1%	33.9%
Total	372	1,316	18,929	20,617
%	100%	100%	100%	100%

Pearson Chi-square: 74.637; degrees of freedom=2; significance=0.000

Table 6 Mean and standard deviation of the overall polarity, positive words and negative words, based on the star ratings of reviews

	1–2 Stars	3 Stars	4–5 Stars
Overall polarity	3.123 (1.099)	3.535 (1.042)	4.134 (.804)
Positive words	2.575 (2.405)	2.942 (2.322)	3.685 (2.515)
Negative words	2.314 (2.149)	1.654 (1.681)	.856 (1.262)
Number of reviews	372	1,316	18,929

(): standard deviation

three different star groups. As the Levene test showed there was no homogeneity of variance (Levene test < 0.05), we used the Welch statistic and the Games-Howell post hoc analysis. This analysis is shown in Table 7.

Table 7 shows that the means in the groups are different from one another, as all the mean differences displayed are significant. Therefore, the post-hoc analysis confirms that overall polarity is higher in positive reviews (mean = 4.134, $SD = 0.8044$) than in mixed-neutral (mean = 3.535, $SD = 1.042$) and negative reviews (mean = 3.123, $SD = 1.099$). Similarly, it is confirmed that the mean of positive words is higher in positive reviews (mean = 3.685, $SD = 2.515$) than in mixed-neutral (mean = 2.924, $SD = 2.322$) and negative reviews (mean = 2.575, $SD = 2.405$), while the mean of negative words is higher in negative reviews

Table 7 Post-hoc analysis

	Welch (Sig.)	Stars (i)	Stars (j)	Mean difference (i-j)	St. Error	Sig
Overall polarity <i>Games-Howell</i>	356.539 (.000)	1–2 stars	3 stars	-.412***	.063	.000
			4–5 stars	-1.010***	.057	.000
		3 stars	1–2 stars	.412***	.063	.000
			4–5 stars	-.598***	.029	.000
			1–2 stars	1.010***	.057	.000
			3 stars	.598***	.029	.000
Positive words <i>Games-Howell</i>	100.032 (.000)	1–2 stars	3 stars	-.349***	.140	.035
			4–5 stars	-1.109***	.126	.000
		3 stars	1–2 stars	.349***	.140	.035
			4–5 stars	-.760***	.066	.000
			1–2 stars	1.109***	.126	.000
			3 stars	.760***	.066	.000
Negative words <i>Games-Howell</i>	224.702 (.000)	1–2 stars	3 stars	.660***	.120	.000
			4–5 stars	1.459***	.111	.000
		3 stars	1–2 stars	-.660***	.120	.000
			4–5 stars	.799***	.047	.000
			1–2 stars	-1.459***	.111	.000
			3 stars	-.799***	.047	.000

*** $p < 0.001$; *df* degrees of freedom; *St.* standard, *Sig.* significanceOverall polarity, Levene statistic: 265,911, *df1* 2, *df2* 20,614, *Sig.* 0.000Positive words, Levene statistic: 3522, *df1* 2, *df2* 20,614, *Sig.* 0.030Negative words, Levene statistic: 195,804, *df1* 2, *df2* 20,614, *Sig.* 0.000

(mean = 2.314, $SD = 2,149$) than in the mixed-neutral (mean = 1.654, $SD = 1,681$) and positive reviews (mean = 0.854, $SD = 1.262$). Therefore, hypotheses H2a, H2b and H2c are supported.

To test H2d we provide a crosstab analysis that examines the correlation between the categories of the variables star rating, negative words and positive words (See Table 8). We also calculated the associated Chi-square statistic and the bivariate Pearson correlation coefficient.

The results in Table 8 show that more 1 and 2 star reviews include three or more negative words than do 3 and 4 star reviews. For instance, 17.1% of 1-star reviews and 18.3% of 2-star reviews included three negative words, while 13.0% of 4-star reviews and 9.1% of 5-star reviews included three negative words. Reviews with 1 or 2 stars usually include more negative words than do reviews with 4 or 5 stars, probably to support that low star rating. However, this trend does not hold for reviews with one or two negative words, as reviews with four and five stars contain more reviews featuring one or two negative words than do reviews with one or two stars. It should be noted that 54.4% of all reviews include at least one negative word. As shown in Table 8, the Chi-square statistic suggests a significant relationship exists between these two variables (Chi-square = 566.433, sig. < 0.001). We also provide the bivariate Pearson correlation coefficient ($r = -0.235$; sig. 2-tailed < 0.001). This coefficient indicates that the volume of negative words in a review is associated with lower star ratings, and vice versa. In addition, the results showed that more 4 and 5 star reviews included three or more positive words than did reviews with 1 or 2 stars. Some 20.8% of reviews with 4 stars, and 21.1% of reviews with 5 stars, included three positive words, whereas 18.3% of reviews with 1 star, and 18.7% of reviews with 2 stars, included three positive words. Again, this trend does not hold in reviews with one or two positive words, as the percentage of reviews with 1 or 2 stars including a positive word (19.3% and 23.7%, respectively) or two positive words (31.2% and 26.3%, respectively) is higher than the percentage of reviews with 4 stars and 5 stars that include a positive word (13.5% and 11.7%, respectively) or two positive words (23.1% and 18.6%, respectively). This means that more than 1 and 2 star reviews include one or two positive words than do reviews with 4 and 5 stars, the latter two usually including more positive words, probably to support the additional stars. Moreover, the impact of including just one or two positive words might not be important for those reviews with a low star rating. It is noteworthy that 67.1% of reviews include 3, or more, positive words. The Chi-square statistic suggests that a significant relationship exists between these two variables (Chi-square = 311.624, sig. < 0.001). In addition, the bivariate Pearson correlation coefficient between star rating and positive words) indicates a significant relationship exists between them ($r = 0.111$; sig. 2-tailed < 0.001). Consequently, the higher the number of positive words in a review, the higher the star rating, and vice versa.

Finally, the results showed that the relationship between the number of negative words and star rating ($r = -0.235$, $p = 0,000$) of a review is stronger than the relationship between the number of positive words and star rating ($r = 0.111$, $p = 0,000$). Thus, H2d is supported.

Table 8 Crosstab analysis between star rating, negative words and positive words

Number of negative words	1 star	2 stars	3 stars	4 stars	5 stars	Total
1 negative word	40	63	369	1,434	3,729	5,635
%	32.5%	33.0%	38.5%	49.6%	60.1%	54.4%
2 negative words	21	54	284	794	1,483	2,636
%	17.1%	28.3%	29.6%	27.5%	23.9%	25.4%
3 negative words	21	35	153	377	563	1,149
%	17.1%	18.3%	16.0%	13.0%	9.1%	11.1%
4 negative words	16	21	82	148	228	495
%	13.0%	11.0%	8.6%	5.1%	3.7%	4.8%
5 negative words	8	7	29	72	95	211
%	6.5%	3.7%	3.0%	2.5%	1.5%	2.0%
6 negative words	6	5	23	30	44	108
%	4.9%	2.6%	2.4%	1.0%	0.7%	1.0%
7 or more negative words	11	6	19	34	63	133
%	8.9%	3.1%	1.9%	1.2%	1.0%	1.3%
Total reviews	123	191	959	29,889	6,205	10,367
%	100%	100%	100%	100%	100%	100%
Number of positive words	1 star	2 stars	3 stars	4 stars	5 stars	Total
1 positive word	21	47	237	631	1,591	2,527
%	19.3%	23.7%	20.1%	13.5%	11.7%	12.8%
2 positive words	34	52	275	1,080	2,526	3,967
%	31.2%	26.3%	23.3%	23.1%	18.6%	20.1%
3 positive words	20	37	240	973	2,865	4,135
%	18.3%	18.7%	20.4%	20.8%	21.1%	21.0%
4 positive words	9	23	176	760	2,391	3,359
%	8.3%	11.6%	14.9%	16.3%	17.6%	17.0%
5 positive words	8	16	108	522	1,603	2,257
%	7.3%	8.1%	9.2%	11.2%	11.8%	11.4%
6 positive words	7	12	52	282	1,039	1,392
%	6.4%	6.1%	4.4%	6.0%	7.7%	7.1%
7 or more positive words	10	11	90	427	1,550	2,088
%	9.2%	5.6%	7.6%	9.1%	11.4%	10.6%
Total reviews	109	198	1,178	4,675	13,565	19,725
%	100%	100%	100%	100%	100%	100%

Negative words/star rating

Pearson Chi-square = 5666.433; degrees of freedom = 72; significance < 0.001

Bivariate Pearson correlation coefficient $r = -0.235$; significance (2-tailed) < 0.001

Positive words/star rating

Pearson Chi-square = 311.624; degrees of freedom = 112; significance < 0.001

Bivariate Pearson correlation coefficient $r = 0.111$; significance (2-tailed) < 0.001

Content analysis.

To address RQ1 we carried out a content analysis of the online reviews, using KH coder, in three phases, parser configuration, frequency analysis and categorisation. We divided the comments into three groups: negative (1–2 stars), neutral (3 stars)

and positive (4–5 stars). Content analysis is usually based on a word-frequency count because, despite its flaws, it is assumed that the words mentioned most frequently reflect the posters’ greatest concerns (Marine-Roig 2017). We carried out a preliminary frequency analysis and a co-occurrence analysis of composite words for each group of reviews (see Figs. 2, 3 and 4, and Table 9).

Figure 2 depicts the co-occurrence analysis of composite words and the frequency of use of words for the one- and two-star reviews. The dimensions of the customer experience are organised into four main independent clusters relating to the words “tour”, “square”, “euro” and “sit”. These four clusters represent simultaneously the words most used and their links with other words frequently used in the 1- and 2-star reviews. The cognitive dimension features words related to visitors’ perceptions of the historical and cultural heritage of the attraction (e.g., museum, secret room), connected to perceptions of value for money (e.g., ticket, money, buy); the behavioural dimension features in its two categories transport (e.g., gondola) and recommended activities (with connected words such as visit-square, spend-time, book-tour); the social dimension features words related to interactions with the staff at the attraction (e.g., tour guides), and which are also connected to the cognitive dimension (secret tour), and words related to overcrowding (e.g., busy, long queue/line); the emotional values dimension is represented by the words “nice” and “experience”. Words related to the sensory

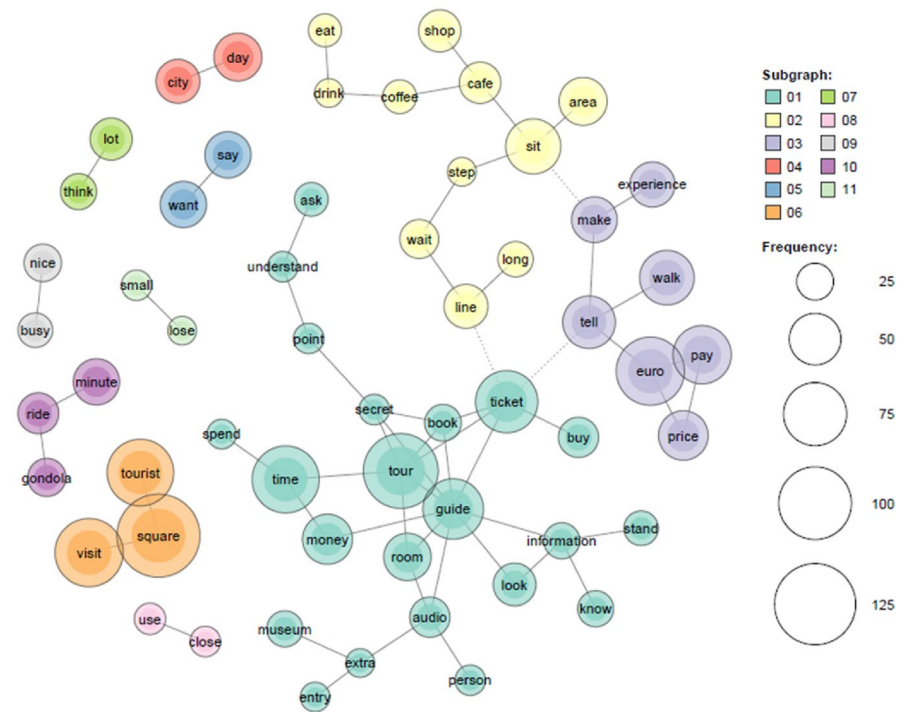


Fig. 2 Co-occurrence analysis for negative reviews (1–2 stars)

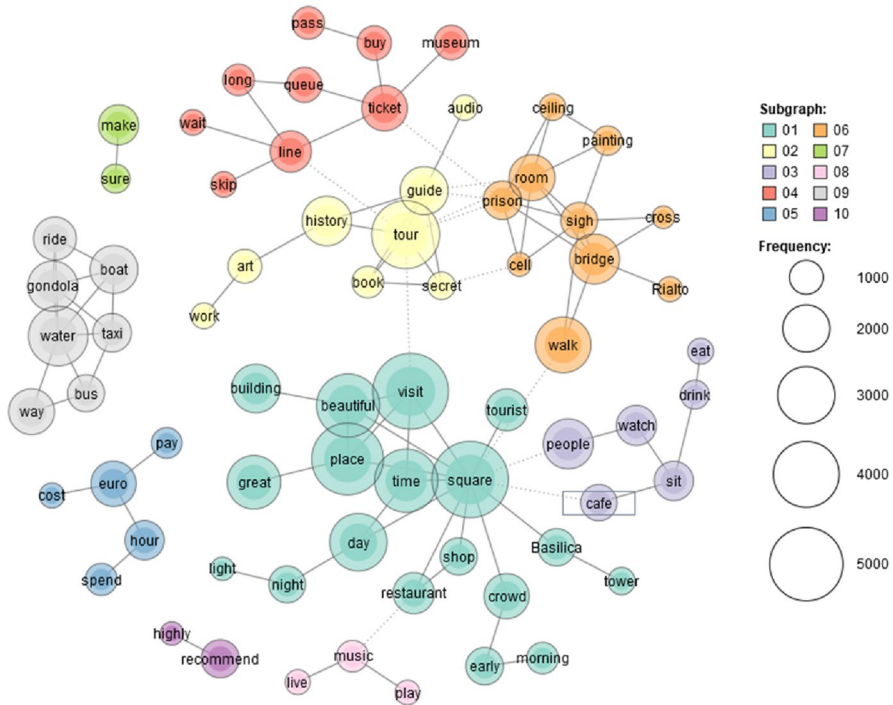


Fig. 4 Co-occurrence analysis for positive reviews (4–5 stars)

Table 9 Fifteen most frequently used words in the reviews, by star rating

Rank	1—2 stars		3 stars		4—5 stars	
	Keyword	Count	Keyword	Count	Keyword	Count
1	Square	131	Square	600	Square	5442
2	Tour	106	Visit	348	Visit	5353
3	People	99	Place	343	Place	4794
4	Place	90	Tour	255	Tour	4279
5	Euro	88	People	252	Time	3689
6	Visit	88	Tourist	236	Beautiful	3683
7	Time	85	Time	230	Water	3259
8	Tourist	84	Crowd	209	Day	3027
9	Ticket	73	Beautiful	202	Walk	2848
10	Guide	69	Walk	199	Great	2624
11	Pay	64	Euro	174	Gondola	2281
12	Sit	57	Lot	163	People	2271
13	Crowd	56	Day	153	Amazing	2270
14	Walk	56	Restaurant	153	Bridge	2267
15	Tell	52	Guide	151	History	2216

review segment (e.g., gondola, boat, water). The cluster linking “pigeon” and “try” belongs to the sensory dimension.

Positive reviews are 91.8% of the total sample. Figure 4 shows seven clusters, all with at least five connected words. All the dimensions under study feature in Fig. 4, particularly the cognitive (with related words such as building, history, basilica, tower and prison, and euro and pay). Emotions belonging to the affective dimension are represented by the words “great” and “beautiful”, connected to place and square, respectively.

The content analysis was complemented with an analysis of the most frequently used words in each star group (see Table 9). The four most frequently used words in the three types of reviews relate to the cognitive and behavioural dimensions: “square”, “tour”, “visit” and “place”. In negative reviews (low score), the most frequently used words relate to value for money, “euro”, “ticket” and “pay”, and to the social dimension (people, tourist, guide and crowd). In neutral reviews, there are, in addition, frequently used words related to the social dimension (people, tourist, crowd and guide) and to the affective dimension (beautiful, nice). The positive reviews have more words related to all the dimensions of the customer experience, with more words being related to the affective dimension of the experience than to the other groups (beautiful, great, amazing). All the customer experience dimensions under analysis are given in Table 9.

Thereafter, we analysed the influence of the positive and negative words related to the five dimensions of the customer experience on the overall polarity of neutral reviews. First, we carried out a descriptive analysis of the number of positive and negative words linked to each service dimension category. As Table 10 shows, 3-star reviews have positive and negative words linked to all the service experience dimensions. Therefore, it seems that consumers are not indifferent to their experiences when visiting Venetian tourism attractions; on the contrary, when they assess their experiences, they make positive and negative comments linked to different service aspects of the attraction which, in turn, lead them to award balanced evaluations (star ratings). The total number of positive words in the 3-star reviews was 3849, while the number of negative words was 2177. The number of positive words linked to specific dimensions of the service experience was higher than the number of negative words, exceptions being the categories value for money (the positive and negative comments were almost balanced) and overcrowding. The neutral score of these reviews can be explained by negativity bias effect, that is, as negative comments are perceived as more diagnostic than positive, consumers include more positive sentiments (positive words) to cancel out the negative sentiments expressed (negative words) and make a final balanced assessment (3-star rating). The higher number of negative words in the categories overcrowding and value for money reflect specific perceptions that Venice is a mature world heritage destination (Table 10). Second, using ANNs, we assessed the weight of the positive and negative words related to the five dimensions of the customer experience. As we did not make any initial assumptions in the specific model, we re-ran the models, changing the activation function and the training function (Table 11). Table 11 shows the MLP models performed better than did the RBF, and the best model was the MLP with one hidden

Table 10 Number of positive and negative words in 3-star reviews, by topic dimensions

Topic dimensions	Number of words	Min./Max
Dimension 1: cognitive	714	0–10
Cultural heritage positive	357	0–6
Cultural heritage negative	446	0–10
Value for money positive	462	0–11
Value for money negative		
Dimension 2: social interaction	150	0–8
Social interaction positive	119	0–6
Social interaction negative	510	0–11
Overcrowding positive	580	0–10
Overcrowding negative		
Dimension 3: behavioural interaction	330	0–10
Transport positive	225	0–11
Transport negative	888	0–15
Activities positive	659	0–11
Activities negative		
Dimension 4: affective	487	0–8
Memorable positive	235	0–6
Memorable negative		
Dimension 5: sensory	417	0–11
Sense positive	222	0–5
Sense negative		
Total positive words	3849	0–20
Total negative words	2177	0–13

Max. Maximum, *Min.* Minimum

Table 11 Artificial Neural Network models for the 3-star reviews

Model	ANN type	Number of hidden layers	Activation function	Training function	Percentage of correct classifications 3 stars
1	MLP	1	Hyperbolic tangent	Scale conjugate gradient	75.2
2	MLP	2	Hyperbolic tangent	Scale conjugate gradient	73.7
3	MLP	1	Hyperbolic tangent	Gradient descent	74.9
4	MLP	2	Hyperbolic tangent	Gradient descent	73.7
5	MLP	1	Sigmoid	Scale conjugate gradient	75
6	MLP	2	Sigmoid	Scale conjugate gradient	75.2
7	MLP	1	Sigmoid	Gradient descent	75
8	MLP	2	Sigmoid	Gradient descent	75.2
9	RBF	–	normalised	–	30.7
10	RBF	–	ordinary	–	22

ANN = Artificial Neural Network

Table 12 Importance and normalised importance of the topics on the overall sentiment for the 3-star reviews

Topic	3 stars	
	Importance	Normalised importance
Negative overcrowding	0.104	100%
Negative value for money	0.103	100%
Positive activities	0.094	91%
Positive overcrowding	0.079	76%
Negative activities	0.078	75%
<i>Positive cultural heritage</i>	0.073	71%
<i>Positive value for money</i>	0.062	60%
<i>Negative transport</i>	0.060	58%
<i>Negative social interactions</i>	0.059	57%
<i>Positive sense</i>	0.055	53%
Negative sense	0.046	45%
Positive transport	0.041	40%
Positive memorable experience	0.039	38%
Positive social interactions	0.038	36%
Negative cultural heritage	0.035	34%
Negative memorable experience	0.032	31%

layer, with a Sigmoid activation function and scale conjugate gradient as the training function.

Table 12 shows the importance (Imp.) and normalised importance (Norm. imp.) of the topics linked to the customer experience dimensions for the overall sentiment of extreme and neutral reviews.

Two criteria were used to analyse the impact of the dimensions of the service experience on mixed-neutral reviews: (i) levels of influence, weak, from 0–50%; medium, from 51–74%; (in *italics*), and strong, from 75–100% (in bold); (ii) the closeness of negative versus positive influence for the same type of comment. The mixed-neutral reviews (3 stars) included negative words, with strong impacts on overall polarity, about overcrowding (100%), value for money (100%) and activities (75%), and positive comments, with strong impacts on overall polarity, about activities (91%) and overcrowding (76%). Overcrowding attracted both negative and positive comments, the influence of the negative was higher than the positive (100% vs. 76%), and the influence of the positive comments about activities was higher than the influence of negative comments about activities (91 vs. 75%): thus, it can be argued that negative comments about overcrowding and value for money, and positive comments about activities, best explain the polarity of mixed-neutral reviews. Furthermore, negative comments about memorable experiences, cultural heritage and positive comments about social interactions, memorable experiences and transport exerted only a weak influence on overall polarity. Therefore, there is a cancel-out effect

between the sentiments of different customer experience dimensions. Therefore, H3 is supported.

4 Discussion

To formulate effective strategies to improve the customer experience, destination marketing organisations (DMOs) must thoroughly understand why consumers award particular star ratings, and the underlying service dimensions of tourist attractions. Mining information from service reviews is an inexpensive and effective way for DMOs to extract advice from tourists about how to improve the quality of services provided in their destinations. Yet, little is known about the relationship between the star ratings and textual components of reviews. Practitioners must begin to understand this relationship to improve star ratings. The present study has shown, by combining deep learning, machine learning and artificial neural networks, that a star rating–sentiment relationship exists.

The analysis of the interplay between star ratings and the emotional content of reviews shows that positive reviews are more objective (factual) than mixed-neutral and negative reviews, thus the emotional tone of reviews is linked to tourists' schemas in relation to attractions. In response to the call by Barger et al. (2016) that an examination should be made of the effects of content factors on consumer behaviour in social media, the current study analysed the impact of the factual tone vs. emotional tone of the online star ratings of tourism attractions. The findings suggest that consumers evaluate eWOM communications differently depending on their emotional tone and the type of service evaluated. The results showed that factual messages reinforce positive message valence. In line with De Keyzer et al. (2017), our findings showed that consumers may feel more confident giving high star ratings to hedonic services (tourism attractions) when they have been provided with clear, positive factual arguments. This result also supports Zhao et al. (2019), who found a negative effect of review subjectivity on hotel ratings.

Our results also showed that the valence of textual components has a positive relationship with star ratings, which confirms the findings of some recent studies that examined similar relationships (Geetha et al. 2017; Yoon et al. 2019; Zhao et al. 2019), and validates the proposition that sentiment analysis polarity can be used as a complement to, or substitute for, star ratings (Al-Natour and Turetken 2020; Valdivia et al. 2019). The greater influence of negative sentiments on overall polarity supports the argument that consumers have a stronger tendency to avoid loss (loss aversion) than to seek gains (Mellinas et al. 2019).

The findings of our content analysis showed that all the customer experience dimensions are present in the three review segments, but the importance of their constituent elements differ. For example, focusing on the cognitive dimension, negative reviews (1–2 stars) feature more words about value for money than about cultural heritage, while the opposite is true for positive reviews (4–5 stars). The affective and sensory dimensions are also salient in positive reviews. The most frequently used words in the three review types relate to the cognitive and behavioural dimensions. In negative reviews (1–2 stars), the most frequently used words

relate to value for money and to the social dimension. In general, the mixed-neutral and, in particular, the positive reviews (4–5 stars), feature more words related to the affective dimension of the experience than do the negative reviews.

We used artificial neural networks to identify the major reasons why customers assign mixed-neutral ratings (3 stars) to attractions by examining how service-based aspects influence overall polarity. The different influence of positive and negative words related to value for money and overcrowding on the 3-star rating segment can be explained by mental associations. As Venice is a mature destination perceived as expensive and overcrowded, it is possible that price perceptions will already be baked into the tourists' mental associations and, thus, do not influence their evaluations of the experience in positive reviews. As regards the weak influence of negative words about cultural heritage (34%), a possible explanation is that Venice is a destination with an important cultural component, so negative sentiments expressed about cultural facets (e.g., the dungeon was dark, the paintings were poorly maintained) may have less influence on overall assessments of the experience than negative sentiments expressed about other service attributes. This result may also be explained by source credibility, as consumers perceive their peers to be more trustworthy, but less expert, sources than DMOs (Ismagilova et al. 2020). The stronger effect of negative than of positive words about service providers in the 3-star reviews relate to competence (e.g., if reviewers felt that the staff had a bad attitude, or they couldn't solve problems), and highlights the importance of triadic encounters and the need to take into account the roles of frontline employees for improving the customer experience (Nguyen and Menezes 2021).

4.1 Theoretical contributions and implications.

The present study makes three contributions to the body of knowledge of online reviews and online consumer behaviour. First, we analyse the consistency between the emotional tone of reviews and their star ratings. According to recent research (Barger et al. 2016; De Keyzer et al. 2017) the tone of the content of online reviews affects consumers' behavioural responses. The present study goes a step forward towards answering De Keyzer's (2017) call by assessing how the emotional tone of online reviews affect consumers' star ratings. Second, we provide support to prospect theory (Kahneman and Tversky 1979) and negativity bias theory (Kanouse and Hanson 1987) in the social media context by demonstrating that the influence of the negative words in a service review on its star rating is higher than the influence of the positive words. Third, this study evidences that examinations into the content categories of online reviews should go beyond overall valence (Bigne et al. 2021, Quiao et al. 2022, Yang et al. 2019). This research extends previous studies (Geetha et al. 2017; Hong and Pittman 2020) by considering the valence of both star ratings and textual components, and examines the relationship between the two factors at a more granular level, that is, of the individual dimensions of the customer experience. This study disaggregates overall sentiment valence into the valence of the specific components of the service experience. As consumers' perceptions of the

components of the service experience are different, our more granular approach to examining the different valences provides further insights for academics and managers into the dimensions of customer service experience.

This study challenges the dominant one-dimensional view that regards positive and negative sentiments as two ends of a continuum. Instead, we adopt a two-dimensional view, that is, positive and negative components can coexist and exert independent impacts simultaneously. Tourist experiences arise in a variety of settings where consumers search for, visit and interact with tourism attractions. Accordingly, this study demonstrated that the customer experience with tourism services can be broken down into five dimensions (sensory, affective, cognitive, social and behavioural), each evoked by different attractions, which supports the multidimensional measurement approach of customer service experiences (Kim and So 2022; Homburg et al. 2017). This study extends research (e.g. Tang et al. 2014) that has emphasised the importance of mixed-neutral user-generated content. Tang et al (2014) showed that mixed-neutral UGC amplifies the effects of positive and negative UGC, but they do not examine the sentiment associated with the different dimensions of the customer experience in mixed-neutral reviews. The present study focuses on the influence of the positive and negative sentiments of each service-based aspect on the overall polarity of mixed-neutral reviews. The present study also extends previous research into the accessibility-diagnostics model (Feldman and Lynch 1988) and compensatory models (Johnson and Meyer 1984) by demonstrating that mixed-neutral reviews may contain positive and negative sentiments, linked to the different dimensions of the customer service experience, that cancel each other out; thus, consumers' overall sentiment towards attractions is a collective expression of the sentiments they feel about distinct service aspects, with particular emphasis on those aspects perceived by consumers as salient during the service experience (Bigne et al. 2020). It can be argued that consumers posting mixed-neutral reviews weigh the dimensions of the service experience differently, because they attribute more importance to some service-based aspects than to others. Based on this rationale, the explanatory and predictive analyses conducted to identify the drivers of the overall sentiment of mixed-neutral reviews showed that this is a valid approach to capturing the cognitive, affective, social, sensorial and behavioural dimensions of the consumer's experience of tourism services.

This study also offers some practical contributions. Reviews with high star ratings are more likely to be expressed in a factual tone than is the case for mixed-neutral and low star rating reviews. Consequently, we advise reviewers and marketers soliciting positive reviews to back up their eWOM with facts and to avoid emotional terminology. Through the analysis of the sentiments that visitors express about specific service aspects DMOs can obtain an informed understanding and detailed insights that cannot be gained by examining only overall ratings. Of the UGC metrics, sentiment valence and star ratings are the most important, as they have a fundamental impact on all types of service business performance. However, service managers tend to focus only on positive and negative UGC, thus they may ignore the influence of neutral UGC. Because mixed-neutral UGC is not, in fact, truly neutral (Tang et al. 2014), and can influence business performance both directly and indirectly, we

caution against this approach and offer specific improvement proposals for managers and practitioners.

First, DMOs should differentiate between mixed and indifferent neutrality. As indifferent-neutral UGC offers no evaluative opinions (Tang et al. 2014), we recommend DMOs improve the design of their online review systems to encourage posters to provide reviews expressing both positive and negative opinions about different features of a tourism attraction. For example, they might ask consumers to evaluate the tourism attraction from both positive and negative perspectives, such as by posing the questions “what’s great about it”, and “what’s not so great.” In addition, DMOs might solicit consumers’ evaluations of multiple dimensions, including transport, cultural heritage, sensory and emotional experiences, activities, staff competence, social interaction with residents and other tourists and value for money.

Second, we suggest that, for the ancillary services surrounding all attractions (cafes, restaurants, etc.), DMOs should display on their opening web pages, first, objective reviews and, second, clear price information and promotions, as both positive and negative words about value for money are drivers of the overall sentiment of mixed-neutral reviews. By offering price reductions for the early morning, and nearer to closing times, attractions can reduce overcrowding during the middle hours of the day. DMOs/providers might also address overcrowding problems by providing detailed information about opening hours, the best itineraries and the best times to visit an attraction. To improve activity-based sentiment DMOs might offer virtual tours, for example, with historical reconstructions, so that potential visitors can view pictorial and textual information before visiting attractions. In addition, DMOs should train the staff working on the attractions to be friendly and to provide during tours interesting stories about the lifestyles of previous generations and how they did things in the past. We also recommend that, to engage their visitors, DMOs should emphasise the cultural value of attractions, providing information about the architectural styles of buildings/structures, as positive comments about cultural heritage may improve the satisfaction of consumers. Information about transport options should also be provided.

4.2 Limitations and future research lines

Despite the importance of the contributions of our study, and its large sample size, we were limited by our use of a single type of tourism service (attractions). Future research might confirm our results using other information sources and other services. The present study showed that the overall sentiment polarity of reviews has a positive effect on customer ratings of attractions. Nonetheless, at the same time, this influence is not totally responsible for customer ratings. Other factors, such as review length, might explain the ratings. Future research might address this.

The present study collected data only from TripAdvisor, which might have a platform bias. Therefore, future research might draw on reviews taken from other platforms. The results of this study can be generalised only to well-known cultural

- Craciun G, Zhou W, Shan Z (2020) Discrete emotions effects on electronic word-of-mouth helpfulness: The moderating role of reviewer gender and contextual emotional tone. *Decis Support Syst* 130:113226
- De Keyser F, Dens N, De Pelsmacker P (2017) Don't be so emotional! How tone of voice and service type affect the relationship between message valence and consumer responses to WOM in social media. *Online Inf Rev* 41(7):905–920
- Deng L, Yu D (2014) Deep learning: methods and applications. *Found Trends in Signal Process* 7(3–4):197–387
- Dhar S, Bose I (2022) Walking on air or hopping mad? Understanding the impact of emotions, sentiments and reactions on ratings in online customer reviews of mobile apps. *Decis Support Syst* 162:113769
- Feldman R (2013) Techniques and applications for sentiment analysis. *Commun ACM* 56(4):82–89
- Feldman J, Lynch J (1988) Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *J Appl Psychol* 73(3):421–435
- García-Pablos A, Cuadros M, Linaza M (2016) Automatic analysis of textual hotel reviews. *Inf Technol Tour* 16(1):45–69
- Gaur L, Afaq A, Solanki A, Singh G, Sharma S, Jhanjhi N, Hoang M, Le D (2021) Capitalizing on big data and revolutionary 5G technology: extracting and visualizing ratings and reviews of global chain hotels. *Comput Electr Eng* 95:107374
- Geetha M, Singha P, Sinha S (2017) Relationship between customer sentiment and online customer ratings for hotels. *Empir Anal Tour Manag* 61:43–54
- Guo Y, Barnes SJ, Jia Q (2017) Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tour Manage* 59:467–483
- Holbrook MB, Hirschman EC (1982) The experiential aspects of consumption: consumer fantasies, feelings, and fun. *J Consum Res* 9(2):132–140
- Homburg C, Jozić D, Kuehnl C (2017) Customer experience management: toward implementing an evolving marketing concept. *J Acad Mark Sci* 45(3):377–401
- Hong S, Pittman M (2020) eWOM anatomy of online product reviews: Interaction effects of review number, valence, and star ratings on perceived credibility. *Int J Advert* 39(7):892–920
- Hu N, Bose I, Koh NS, Liu L (2012) Manipulation of online reviews: an analysis of ratings, readability, and sentiments. *Decis Support Syst* 52(3):674–684
- Ismagilova E, Slade E, Rana N, Dwivedi Y (2020) The effect of characteristics of source credibility on consumer behaviour: a meta-analysis. *J Retail Consum Serv* 53:101736
- Johnson E, Meyer R (1984) Compensatory choice models of noncompensatory processes: The effect of varying context. *J Consum Res* 11(1):528–541
- Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. *Econometrica* 47:263–291
- Kandampully J, Zhang TC, Jaakkola E (2018) Customer experience management in hospitality: a literature synthesis, new understanding and research agenda. *Int J Contemp Hosp Manag* 30(1):21–56
- Kanouse DE, Hanson LR Jr (1987) Negativity in evaluations. Preparation of this paper grew out of a workshop on attribution theory held at University of California. Lawrence Erlbaum Associates, Inc, Los Angeles
- Ketter E (2018) It's all about you: destination marketing campaigns in the experience economy era. *Tourism Rev* 73(3):331–343
- Kim D, Park B (2017) The moderating role of context in the effects of choice attributes on hotel choice: A discrete choice experiment. *Tour Manage* 63:439–451
- Kim H, So K (2022) Two decades of customer experience research in hospitality and tourism: A bibliometric analysis and thematic content analysis. *Int J Hosp Manag* 100:103082
- Li X, Wu C, Mai F (2019) The effect of online reviews on product sales: A joint sentiment-topic analysis. *Inf Manage* 56(2):172–184
- Liu B (2010) Sentiment analysis and subjectivity. *Handb Nat Lang Process* 2:627–666
- Liu Z, Park S (2015) What makes a useful online review? Implication for travel product websites. *Tour Manage* 47:140–151
- Luca M (2016) Reviews, reputation, and revenue: The case of Yelp. com. Com (March 15, 2016). Harvard Business School NOM Unit Working Paper, (12–016).
- Luo Y, Xu X (2021) Comparative study of deep learning models for analyzing online restaurant reviews in the era of the COVID-19 pandemic. *Int J Hosp Manag* 94:102849

- Mellinas J, Nicolau L, Park S (2019) Inconsistent behavior in online consumer reviews: the effects of hotel attribute ratings on location. *Tour Manage* 71:421–427
- Nguyen NB, Menezes J (2021) The thirty-year evolution of customer-to-customer interaction research: a systematic literature review and research implications. *Serv Bus* 15:391–444
- Nieto-Garcia M, Resce G, Ishizaka A, Occhiocupo N, Viglia G (2019) The dimensions of hotel customer ratings that boost RevPAR. *Int J Hosp Manag* 77:583–592
- Osmani A, Mohasefi J, Gharehchopogh F (2020) Enriched latent dirichlet allocation for sentiment analysis. *Expert Syst*. <https://doi.org/10.1111/exsy.12527>
- Palmer A, Montano J, Sesé A (2006) Designing an artificial neural network for forecasting tourism time series. *Tour Manage* 27(5):781–790
- Phillips P, Zigan K, Silva M, Schegg R (2015) The interactive effects of online reviews on the determinants of Swiss hotel performance: A neural network analysis. *Tour Manage* 50:130–141
- Pike S, Pontes N, Kotsi F (2021) Stopover destination attractiveness: a quasi-experimental approach. *J Destin Mark Manag* 19:100514
- Qiao T, Shan W, Zhang M, Wei Z (2022) More than words: Understanding how valence and content affect review value. *Int J Hosp Manag* 105:103274
- Racherla P, Connolly DJ, Christodoulidou N (2013) What determines consumers' ratings of service providers? An exploratory study of online traveler reviews. *J Hosp Market Manag* 22(2):135–161
- Rather RA, Hollebeek L, Rasoolimanesh S (2022) First-time versus repeat tourism customer engagement, experience, and value cocreation: an empirical investigation. *J Travel Res* 61(3):549–564
- Rudkowsky E, Haselmayer M, Wastian M, Jenny M, Emrich S, Sedlmair M (2018) More than bags of words: sentiment analysis with word embeddings. *Commun Methods Meas* 12(2–3):140–157
- Sayfuddin AT, Chen Y (2021) The signaling and reputational effects of customer ratings on hotel revenues: evidence from TripAdvisor. *Int J Hosp Manag* 99:103065
- Schmitt B (1999) Experiential marketing. *J Mark Manag* 15(1–3):53–67
- Schoefer K, Ennew C (2005) The impact of perceived justice on consumers' emotional responses to service complaint experiences. *J Serv Mark* 19(5):261–270
- Sharma A, Park S, Nicolau JL (2020) Testing loss aversion and diminishing sensitivity in review sentiment. *Tour Manage* 77:104020
- Siddiqi UI, Sun J, Akhtar N (2020) The role of conflicting online reviews in consumers' attitude ambivalence. *Serv Ind J* 40(13–14):1003–1030
- Siering M, Deokar A, Janze C (2018) Disentangling consumer recommendations: explaining and predicting airline recommendations based on online reviews. *Decis Support Syst* 107:52–63
- Simeon MI, Buonincontri P, Cinquegrani F, Martone A (2017) Exploring tourists' cultural experiences in Naples through online reviews. *J Hosp Tour Technol* 8(2):220–238
- Sorokina E, Wang Y, Fyall A, Lugosi P, Torres E, Jung T (2022) Constructing a smart destination framework: a destination marketing organization perspective. *J Destin Mark Manag* 23:100688
- Taecharunroj V, Mathayomchan B (2019) Analysing TripAdvisor reviews of tourist attractions in Phuket, Thailand. *Tour Manage* 75:550–568
- Tang T, Fang E, Wang F (2014) Is neutral really neutral? The effects of neutral user-generated content on product sales. *J Mark* 78(4):41–58
- Thornton JR (2011) Ambivalent or indifferent? Examining the validity of an objective measure of partisan ambivalence. *Polit Psychol* 32(5):863–884
- Timoshenko A, Hauser J (2019) Identifying customer needs from user-generated content. *Mark Sci* 38(1):1–20
- Tkáč M, Verner R (2016) Artificial neural networks in business: two decades of research. *Appl Soft Comput* 38:788–804
- Tversky A, Kahneman D (1991) Loss aversion in riskless choice: a reference-dependent model. *Q J Econ* 106(4):1039–1061
- Uysal M, El Roubi M (1999) Artificial neural networks versus multiple regression in tourism demand analysis. *J Travel Res* 38(2):111–118
- Valdivia A, Hrabova E, Chaturvedi I, Luzón MV, Troiano L, Cambria E, Herrera F (2019) Inconsistencies on TripAdvisor reviews: a unified index between users and sentiment analysis methods. *Neurocomputing* 353:3–16
- Verhoef PC, Lemon KN, Parasuraman A, Roggeveen A, Tsiros M, Schlesinger LA (2009) Customer experience creation: determinants, dynamics and management strategies. *J Retail* 85(1):31–41
- Yang Y, Park S, Hu X (2018) Electronic word of mouth and hotel performance: a meta-analysis. *Tour Manage* 67:248–260

- Yang M, Ren Y, Adomavicius G (2019) Understanding user-generated content and customer engagement on facebook business pages. *Inf Syst Res* 30(3):839–855
- Yoon Y, Kim A, Kim J, Choi J (2019) The effects of eWOM characteristics on consumer ratings: evidence from TripAdvisor.com. *Int J Advert* 38(5):684–703
- Yu J, Egger R (2021) Tourist Experiences at Overcrowded Attractions: A Text Analytics Approach. In: Wörndl W, Koo Ch, Stienmetz (eds) *Information and Communication Technologies in Tourism 2021*. Springer, Cham, pp 231–243
- Zhang X, Yu Y, Li H, Lin Z (2016) Sentimental interplay between structured and unstructured user-generated contents: an empirical study on online hotel reviews. *Online Inf Rev* 40(1):119–145
- Zhao Y, Xu X, Wang M (2019) Predicting overall customer satisfaction: big data evidence from hotel online textual reviews. *Int J Hosp Manag* 76:111–121
- Zhu L, Lin Y, Cheng M (2020) Sentiment and guest satisfaction with peer-to-peer accommodation: when are online ratings more trustworthy? *Int J Hosp Manag* 86:102369

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.