RESEARCH PAPER

Cultural differences in hospitality service evaluations: mining insights of user generated content

Chong Guan[1] · Yu-Chen Hung[1] · Wenting Liu[1]

Abstract
Online peer-to-peer platforms have been widely utilized to facilitate sharing vast amounts of user-generated content (UGC) in the hospitality industry. However, research on the cultural impacts of the broad spectrum of UGC remains nascent and limited. This study therefore elaborates how the hospitality service dimensions that are reflected in UGC both predict service evaluation and are contingent on a reviewer's prevailing culture. We analyzed a usable sample of 9,257 hotel reviews from 148 countries with latent Dirichlet allocation and aspect-based sentiment analysis algorithms to identify the top service features and corresponding polarity scores. Machine learning identified three dimensions of hospitality service quality: adaptability, reliable delivery, and tangibles. These three dimensions predict customers' overall hotel ratings, and the magnitude and direction of their effects depend on alignments of cultural orientations. Following Hofstede's cultural dimensions, reliable delivery and tangibles align with individualism and uncertainty avoidance cultures. Adaptability aligns with power distance, masculinity, long-term orientation, and indulgence cultures.

Introduction

Online reviews, electronic versions of traditional word of mouth (WOM), have become a potent source of information for consumers. TripAdvisor, one of the most popular travel review websites, offers approximately 859 million annual user reviews in more than eight million listings for hotels, vacation rentals, restaurants, and attractions (TripAdvisor, 2019). Such review sites allow easy access to fellow travelers' opinions before booking on online tourist agent sites or hotel websites. 92% of consumers evaluate the textual content of reviews before making a purchase online (Min et al., 2018). Due to growing scholarly awareness of this influence, increasing research has focused on the factors

that drive reviewers to write reviews (Gonçalves et al., 2018), how they write them (Min et al., 2018), what they write (Nakayama & Wan, 2018), and how companies shall respond to the reviews (Zhu et al., 2021).

Online reviews are generated by reviewers living in different cultural realities. Reviewers' expressions reflect the prevailing values of the culture they belong to (Chu & Choi, 2011). To date, nascent studies on cultural influences in online reviews have shown that culture is a prevalent factor that affects the motivation to review (Min et al., 2018), reviewers' linguistic style (Nakayama & Wan, 2018), and user-generated content (UGC; Mariani and Matarazzo (2020)). However, the scope of current studies has been limited to a single cultural value (e.g., power distance (Gao et al., 2018)) and how overall customer satisfaction is reflected in online ratings (Mariani & Predvoditeleva, 2019). Despite the richness of textual content, few studies have examined the service quality dimensions of hospitality that culturally grounded reviewers use to form evaluations (Kozinets, 2016; Winer & Fader, 2016), with few exceptions (Büschken & Allenby, 2016; Li et al., 2013). Thus, the dimensions

---

✉ Chong Guan
guanchong@suss.edu.sg

[1] Singapore University of Social Sciences, 463 Clementi Road, 599494 Singapore, Singapore

Ⓐ Springer

of reviewers' comments that play more significant roles in their evaluations have yet to be defined.

Although numerous studies have established specific cultural impacts on evaluations of hospitality services by surveying a small sample of customers (Furrer et al., 2000; Mattila, 1999, 2000; Tsaur et al., 2005), recent research has highlighted the need to use online reviews to make service dimensions more relevant to their cross-cultural context amid rapidly changing service practices (Naumov, 2019). Technologies have increasingly become the interfaces where customers interact with service providers (Ivanov & Webster, 2019). Thus, a new conceptualization of service quality dimensions is needed to update traditional views on guest-host interactions, such as SERVQUAL (Parasuraman et al., 1991b; Parasuraman et al., 1985). For example, research on electronic service quality suggests that a customer's evaluation of technology is a distinct process (Zeithaml et al., 2002). The recent surge of automation in service delivery has shifted all hotel guest experiences further from staff-customer interactions (Ostrom et al., 2015). Service quality is no longer the result of service encounters between 'hosts and guests' (Kandampully et al., 2017). These evolving practices, which blend the roles of humans and technologies in service delivery, warrant an updated perspective of service quality dimensions (Naumov, 2019).

The richness of consumer-generated content provides an opportunity to reveal the service dimensions that are relevant to an evolving service context since consumers can freely express themselves without being limited or guided by survey questions (Li et al., 2013). Thus, our research answers two questions. First, what are the service quality dimensions that consumers use to evaluate hotel services when service delivery blends technology and staff? Second, how do cultural values moderate the relative importance of these service dimensions for customer satisfaction? To answer the first question, we draw on the literature on the transformed role of staff in service standardization and customization in relation to technology (Doorn et al., 2016; Larivière et al., 2017; Robinson et al., 2019; Sandoff, 2005; Solnet et al., 2016). We posit that service quality perceptions reflect two kinds of staff behaviors in blended service encounters: high adaptiveness in customized service versus low adaptiveness in standardized service. We refer to the former as the human touch in service and the later as the systemization of service. Using this theoretical framework, machine learning algorithms were deployed to uncover the service dimensions empirically. To answer the second question, we reviewed cultural impacts on service quality evaluations based on Hofstede's cultural dimensions (Hofstede, 1991) and on the SERVQUAL model (Parasuraman et al., 1991b; Parasuraman et al., 1985). Our review produced mixed findings on how Hofstede's six cultural dimensions affect the relative importance of service dimensions. We further examined how these cultural dimensions affect our empirically identified service dimensions in the second phase of analysis.

We collected nearly 10,000 reviews from 148 countries, which were analyzed with machine learning (ML) algorithms, namely, latent Dirichlet allocation (LDA) (Blei et al., 2003), aspect-based sentiment analysis (ABSA) (Jo & Oh, 2011; Liu, 2012) and structural equation modeling (SEM). We conducted the analysis in two phases. First, we identified and extracted the keywords from textual data using LDA's top service features and generated the sentiment score for each feature using ABSA. The loading of the service features was used to establish the distinct service dimensions that impact customers' overall satisfaction. Second, SEM was applied to compare the differential moderating roles of various culture dimensions in how service dimensions impact overall satisfaction.

The first phase of analysis revealed that the service features loaded on three distinct dimensions. These three dimensions were labeled *adaptability, reliable delivery*, and *tangibles*. The adaptability dimension reflected the important role of staff in creating a memorable experience with customization, while the dimensions of reliable delivery and tangibles reflected the systemized integration of technology and staff in standardized service. Hence, consistent with our expectations, adaptability expressed the human touch aspect in service quality, while reliable delivery and tangibles demonstrated the systemization aspect of service quality. The results of the phase 1 analysis indicate that these three service dimensions are used by consumers to evaluate hotel service quality amid blended service delivery.

The second phase of analysis tested how the relative importance of the three dimensions varied with respect to Hofstede's cultural dimensions (Hofstede, 1991). The service dimension of adaptability aligns with cultures that value power distance, masculinity, long-term orientation, and indulgence. These dimensions significantly contribute to favorable service evaluations by members of such cultures. In contrast, adaptability did not result in favorable service evaluations among reviewers from individualist and uncertainty avoidant cultures. Instead, the service dimensions of tangibles and reliable delivery significantly fostered favorable service evaluations from members of cultures that are high in individualism and uncertainty avoidance. The results of the phase 2 analysis indicate that standardized service aligns with individualist and high uncertainty avoidance cultures. The systemized service delivery has more value to consumers who tend to be individualist and uncertainty avoidant. Customized service aligns with cultures comprising power distance, masculinity, long-term orientation, and indulgence. The human touch in service is

critical for consumers who tend to be power distant, masculine, long-term focused, and indulgent.

Our research advances the understanding of how culture influences online reviews in relation to service quality dimensions. Machine learning empirically identified three dimensions of service quality: adaptability, reliable delivery, and tangibles. These three dimensions conceptually relate to two distinct focuses, service standardization or customization, in any blended service delivery by technology and staff. We also identified the alignments between the three identified service dimensions and Hofstede's cultural dimensions. The results suggest that customized service is more aligned with cultures that value power distance, masculinity, long-term orientation, and indulgence than with cultures that emphasize individualism and uncertainty avoidance. Our findings expand the understanding of international service quality dimensions beyond frameworks that are anchored in guest-staff interaction, such as SERVQUAL (Parasuraman et al., 1991b; Parasuraman et al., 1985).

Our findings are of practical importance to hotels and review platforms. Hotels invest to improve their perceived service quality and attract positive reviews. They can afford to be more selective in their technological investments in the service dimensions that are valued by their target audiences. Promotional messages should highlight the dimensions that align with prevailing cultural values. Depending on the locations of website visitors, hotels can present content featuring dimensions that match their cultural values along with testimonials from their compatriots. This matching will also increase the effectiveness of a hotel's digital advertising due to a higher conversion rate. Review platforms are well positioned to advise their advertisement clients on showcasing their service dimensions in a culturally aligned manner. Equipped with such data, review platforms can offer guidelines for agencies to manage their message evaluation to help optimize advertisers' campaign budgets.

Theoretical framework

User-generated content (UGC) is broadly defined as what consumers share on digital platforms, including product reviews on sites, such as Amazon or TripAdvisor, and textual, visual, and video posts on social media (Ayeh et al., 2013). Online reviews have attracted the attention of both academia and industry due to their critical impacts on consumer attitudes and purchase decisions (Bickart & Schindler, 2001; Trusov et al., 2009; Xia & Bechwati, 2008). Chevalier and Mayzlin (2006) and Duan et al. (2008) confirm that online reviews are a valid predictor of company financial performance in terms of revenue and profitability.

UGC serves as an appropriate source for market insights about consumers' experiences with service quality (Tirunillai & Tellis, 2014). Content analysis of online reviews in hospitality identifies the determinants of customer satisfaction. Li et al. (2013) discern six factors that are based on the most frequently mentioned words including logistics/location, facilities, receptionist services, food and beverage, room cleanliness and maintenance, and monetary value. Büschken and Allenby (2016) use topic modeling to identify the latent topics that underpin online reviews of Manhattan hotels, airport hotels, and Italian restaurants. Similar topics emerge for both hotels and restaurants. Reviews frequently mention problems at check-in, nearby attractions, food recommendations, room noise, positive room features, location, transportation, amenities, staff friendliness, and New York experiences.

Cultural impact on online reviews of hospitality

Culture has been shown to affect both the overall ratings and textual content of online reviews. Overall rating has been used as a proxy for customer satisfaction (Radojevic et al., 2017). Gao et al. (2018) find a negative relationship between reviewers' power distance and their online hotel ratings. Specifically, customers with high power distance report being less satisfied more often than their low power distance counterparts. They feel superior to service providers and demand customized service upon request. Their sense of frustration prevails when service staff are unable to adaptively respond to their demands. Mariani and Matarazzo (2020) show that customer satisfaction is higher when service providers and customers speak the same language, suggesting that a cultural understanding of the required level of staff adaptiveness enhances customer satisfaction.

In addition to overall ratings, visitors' generated comments reveal how their dispositions are culturally accustomed. Westerners are inclined to use more positive emotional expressions and analytical narration (Min et al., 2018). After patronizing restaurants, their comments concern service and ambiance more than food quality and price fairness (Nakayama & Wan, 2018). These findings indicate that cultural impacts extend beyond linguistic styles. Textual contents reflect the service quality dimensions that are valued in cross-cultural contexts, warranting further investigation. Echoing this view, Kozinets (2016) and Winer and Fader (2016) articulate the need to incorporate different cultural realities into consumer-generated reviews. Min et al. (2018) also stress the importance of online reviews for determining the service dimensions that are relevant in cross-cultural contexts. In response to the calls for broad cultural understanding of online reviews, our research investigates how service dimensions align with cultural values.

Table 1 The relationships between service quality dimensions in SERVQUAL and Hofstede's cultural dimensions

		Individualism/Collectivism	Power Distance	Masculinity/Femininity	Uncertainty Avoidance	Long-Term Orientation	Service Type
Furrer et al. (2000)	Tangibles	+	+	+	-	-	Banking Service
	Reliability	+	-	*	+	+	
	Responsiveness	+	-	-	+	+	
	Assurance	-	+	*	+	-	
	Empathy	-	-	*	+	*	
Donthu & Yoo (1998)	Tangibles	*	*	^	+	*	Banking Service
	Reliability	*	-	^	*	*	
	Responsiveness	*	-	^	*	+	
	Assurance	+	*	^	*	*	
	Empathy	+	*	^	*	*	
(Tsoukatos & Rand, 2007)	Tangibles	^	*	*	*	*	Retail Insurance
	Reliability	^	-	-	+	+	
	Responsiveness	^	-	-	+	*	
	Assurance	^	-	-	+	*	
	Empathy	^	*	*	*	*	
(Kueh & Ho Voon, 2007)	Tangibles	*	-	*	+	+	Restaurants
	Reliability	*	-	*	+	+	
	Responsiveness[a]	*	-	*	+	+	
	Assurance[a]	*	-	*	+	+	
	Empathy[a]	*	-	*	+	+	
(Tsiligiris et al., 2022)	Tangibles	-	*	^	+	+	Higher Education
	Reliability	*	*	^	+	+	
	Responsiveness	*	-	^	*	+	
	Assurance	*	-	^	*	+	
	Empathy	*	*	^	*	+	
(Tsikriktsis, 2002)[b]	Tangibles	*	*	+	*	+	Banking Websites
	Reliability	^	^	^	^	^	
	Responsiveness	*	*	+	*	*	
	Assurance	*	*	*	*	*	
	Empathy	*	*	+	*	+	

"+" = positive relationship/"-" = negative relationship/"*" = nonsignificant relationship/"^" = not tested

"a" = tested as one dimension

"b" = the study used a modified WEBQUAL scale, where tangibles refer to design and visual appeal. No reliability dimension was included. Responsiveness remains the same as SERVQUAL. Assurance refers to trust. Empathy refers to flow-emotional appeal. Modified scales, such as LODGSERV (Knutson et al., 1990) and HOLSERV (Wong Ooi Mei et al., 1999), are used in the accommodation sector (Al-Ababneh, 2017). To the best of our knowledge, no empirical studies have examined their relation to Hofstede's cultural dimensions

Cultural impact on service quality dimensions

One of the widely accepted models of service quality dimensions is the SERVQUAL model of Parasuraman et al. (Parasuraman et al., 1991a; Parasuraman et al., 1985). The model categorizes service quality perceptions across five dimensions: tangibles, reliability, responsiveness, assurance, and empathy. Tangibles comprise physical facilities, equipment, personnel, and communication materials, among others. Reliability represents a staff's capacity to provide promised services dependably and accurately. Assurance denotes the expertise and courtesy of service employees and their ability to convey confidence and trust. Responsiveness is a staff's willingness to help customers and deliver prompt service. Empathy entails the caring, individualized attention provided to a customer. The relative importance of each of these five dimensions is subject to customers' values and beliefs. At the individual level, demographics and psychographics are shown to systematically influence the service quality dimensions that are important for customer satisfaction; at the societal level, culture is critical for shaping consumer beliefs and expectations concerning service norms (Berry et al., 1988).

Any evaluation of service quality is strongly affected by cultural background (Dedj & Pavlovi 2011; Matzler et al., 2006; Torres et al., 2014). An examination, based on Hofstede's cultural dimensions and the five dimensions of SERVQUAL, provides mixed empirical results (Table 1). For example, tangibles are shown to be highly valued by

individualist consumers for banking services in Furrer et al. (2000), but this importance was not supported in other studies. Assurance is shown to be highly valued by individualist consumers in Furrer et al. (2000), but a negative or nonsignificant relationship is found in other studies. These mixed findings are further complicated by concerns regarding whether similar service quality dimensions apply in international contexts (Ladhari, 2009). Moreover, a review by (Taylan Dortyol et al., 2014) identified a range of two to twelve service dimensions that may be relevant to international tourists' reviews.

The cultural dimensions in Hofstede (1980) have been extensively used to investigate cultural differences in recent decades (Beugelsdijk et al., 2017). Hofstede's model has also been widely criticized for its methodology, sample, national focus, and number of dimensions (Sent & Kroese, 2020). Despite such criticism, it remains the most used of the competing cultural dimension models (Inglehart & Baker, 2000; Schwartz, 2006). Scholars have consistently attempted to redefine or update one or more of its dimensions without much consensus (Ababneh & Shrafat, 2014; Stępień & Dudek, 2021). These six cultural dimensions, with Hofstede's (1991) addition of the indulgence dimension ,have been proven to be highly relevant to UGC (Radojevic et al., 2019). Regarding UGC, there is uncertainty concerning which cultural dimensions are more relevant. For example, cultural differences between visitors and hotel staff have significant effects on satisfaction in certain dimensions but none in other dimensions (Radojevic et al., 2017; Radojevic et al., 2018). Given both the relevance and uncertainty of Hofstede's cultural dimensions regarding user-generated content, we further adopt this model to investigate cultural influences.

Service quality dimensions in evolving service contexts

The existing research on the relationship between culture and service quality dimensions, however, has not focused on the relevance of traditional service dimensions to the current service context, which is dominated by technological innovation. Under SERVQUAL, a service encounter has been conceptualized as "the dyadic interaction between a customer and a service provider". Service delivery has thus been characterized by dyadic human interactions between customers and employees. Service quality is determined by the friendly, welcoming, and warm behaviors that service providers offer to customers. However, a customer evaluation amid new technologies is a distinct process. For example, five distinctive e-service dimensions have been developed by Zeithaml et al. (2002), comprising information availability and content, ease of use or usability, website privacy/security, website graphic style, and reliable fulfillment.

Hospitality providers have continuously invested in service automation. Customers have gradually demanded self-reliant and speedy service. Self-service technologies at airports and hotels are popular innovations that improve customer experiences and reduce waiting times (Kattara & El-Said, 2014; Kucukusta et al., 2014). Service encounters currently consist of interactions with apps, contactless check-ins and check-outs, in-room technologies, and smart facilities. Service quality is supported, rather than achieved, by human interactions involving customers and employees. Through these changing practices, it is expected that a personal touch during service encounters becomes a reserved privilege, provided by a few hospitality providers (Naumov, 2019). Since the service quality dimensions concerning people-delivered service are inapplicable to customer evaluations of service delivery enabled by technology (Parasuraman et al., 1985), our first research question is as follows: what service quality dimensions do consumers use in today's evolving service context?

Technological progress transforms the role of staff in service encounters (Leischnig et al., 2018). One significant service practice evolution is automating touchpoints to minimize human intervention (Kannan & Healey, 2011). Automation is deployed to reduce operating costs and standardize service delivery. In standardized service, tasks are highly structured; employees are expected to exercise limited judgment and flexibility. They have little discretion for adopting their services to suit customers' personal needs (Wang et al., 2020). Staff plays a functional role in assuring reliable service delivery. Hence, the reliability of a technological environment with sufficient staff support is critical for service quality perception.

While reliability has been traditionally regarded as a critical influencer of service quality perceptions, customization has also become a more determining influence (Gwinner et al., 2005). Hotel guests expect smart technology to empower staff to improve customized service by providing instant information on room status, catering to personal preferences, and fulfilling special requests (Solnet et al., 2016). In customized service, tasks are less structured than in standardized service. Although technology enables customer-tailored solutions (Chung et al., 2009; Chung et al., 2016), customizing offerings to suit individual customers often requires interpersonal interventions. The execution of customized service relies on employee adaptiveness, i.e., "the deliberate modification of the service offering and/or the employee's interpersonal behavior in a situationally appropriate manner in response to meeting perceived consumer needs" (Gwinner et al., 2005, p. 135). Employ adaptiveness can be a service differentiator, as staff cocreates

Table 2 Proposed alignments between cultural and service quality aspects

	Individualism- Collectivism (IDV)	Uncertainty Avoidance (UAI)	Masculinity-Femininity (MAS)	Power Distance (PDI)	Long Term Orientation (LTO)	Indulgence-Restraint (IND)
Human touch			A (H1)	A (H1)	A (H1)	A (H1)
Systemization	A (H2)	A (H2)				

Note: Culturally aligned service dimensions are highlighted in gray, while culturally misaligned service dimensions are not highlighted

memorable experiences with guests. Customized service requires a proactive attitude, attentiveness, and emotional resonance to exude the warmth of human touch that automation cannot.

We posit that service quality perceptions reflect two kinds of staff roles in blended service encounters: high adaptiveness to customize service in a low-structured task and low adaptiveness to standardize service in a highly structured task (Leischnig et al., 2018). For ease of reference, we identify the former as the human touch in service quality and the later as the systemization of service quality. Regarding the human touch in service quality, staff adaptiveness is a differentiator that shows a personal touch to cocreate a customer experience. An evaluation of this is based on whether staff show personal initiative and willingness to modify their approach during the interaction. Regarding the systemization of service quality, staff functions as a supplement to technology-enabled service delivery. A favorable service quality perception can involve no or little human intervention, and an evaluation is based on whether the service is delivered steadily and smoothly. A similar dichotomy focused on touch or tech has shown that both aspects affect customer satisfaction (Makarem et al., 2009).

Human touch and customer satisfaction

A human touch in service quality should enhance overall customer satisfaction. Customized service offers a tailor-made solution for a customer's benefit (Ding & Keh, 2016). Compared to standardized service, it better matches preferences with service attributes. Thus, customized service has been shown to lead to greater customer satisfaction (Franke et al., 2009), to enhance customer experiences (Franke et al., 2008), and to increase customer loyalty (Coelho & Henseler, 2012).

However, some studies show that service customization does not always result in a better customer experience (Franke et al., 2009; Leischnig et al., 2018; Simonson, 2005). One factor that may adversely affect service quality perception is perceived risk (Ding & Keh, 2016). A higher variability in the performance of service evokes a higher perceived risk. Potential variations in staff discretionary behavior may induce uncertainty about whether a customized offering fits personal preferences. Standardized service, on the other

hand, allows a predictable workflow that assures confidence in consistent service quality. The second factor is consumer reactance. Customized service often requires proactive staff to demonstrate adaptive behavior beyond the usual service level. However, excessive attentiveness by staff may lead to the counterproductive effect called "overservicing" (Ku et al., 2013; Terpstra & Verbeeten, 2014). Excessive service may appear unnecessary and even disturbing. Without a proper cultural understanding, attentiveness to provide personalized service may appear forceful and aggressive to some consumers. A personalized service encounter may create a situation where consumers feel obligated to accept staff favors. This may trigger psychological reactance, especially among customers who value independence and freedom.

Cultures as moderators of human touch for customer satisfaction

Based on previous research, we posit that service quality differs in two aspects—human touch and systemization. A human touch in service will positively relate to service quality evaluation unless cultural values trigger counterproductive effects. Among the cultural dimensions in Hofstede's model, two cultural dimensions, IDV and UAI, are likely to heighten reactance and perceived risk. Due to their increased activation of these two counterproductive factors, we posit that IDV and UAI cultures are not in alignment with high staff adaptiveness. That is, high human touch may negate satisfaction among members of individualist and uncertainty avoidant cultures. In contrast, systemization will mitigate the activation of counterproductive factors, effectively enhancing satisfaction. As counterproductive factors in other cultural dimensions are not increased, human touch aligns with them. High human touch will significantly enhance customer satisfaction. A summary of this relationship is provided in Table 2. Notably, two aspects may consist of more than one dimension. The service dimensions will be informed by topic modeling. We explain the alignment between service aspect and cultural dimension in detail below.

Individualism-Collectivism (IDV) This dimension describes whether people view themselves as independent or identify themselves with groups (Triandis et al., 1988). Individualist

cultures are oriented around the self, value personal freedom and autonomy, and encourage individual decision-making. In contrast, collectivist cultures are characterized by an emphasis on communal goals and group conformity, which should come before individual desires or pursuits.

A member of an individualist culture aspires to achieve autonomy and become his or her unique self. Service dimensions that support autonomous self-narration should be advocated for such persons. Self-service that is enabled by service automation is a source of autonomy in service encounters. It is reasonable to assume that a customized service experience is sought after and appreciated by members of individualist cultures due to their need for uniqueness. However, recent research shows a concave relationship between attentive service provision and satisfaction (Ku et al., 2013). Service encounters that exceed standard service protocols may be considered favors that individuals have not requested and are not prepared to reciprocate in a transactional relationship. An overly attentive staff may appear forceful and aggressive, threatening individual freedom of choice and self-determination. Individualists are likely to respond with reactance and distrust. In contrast, collectivists are oriented to understanding staff adaptive behavior through communal relationships and to interpret their intentions to treat customers like their own family or friends accordingly. Therefore, customized service may risk triggering reactance among individualist customers, negating their customer satisfaction. A human touch in service quality may not align with individualism.

Uncertainty Avoidance Index (UAI) This cultural dimension entails how societies differ in their tolerance for risk, unpredictability, and ambiguity (Hofstede, 1991). Cultures with a high UAI value stability, established norms and formal protocols for structured tasks. Formal rules and explicit guidelines are structured in detailed contracts with business partners (Wuyts & Geyskens, 2005). Individuals in cultures with a high UAI try to minimize unknown and unusual circumstances. They exert meticulous efforts on internal and external controls (Hwang, 2005). A service encounter that deviates from expectations may create an ambiguous situation where uncertainty-avoidant individuals are not certain about how to act, what is expected of them, or how to react. Uncertainty elicits anxiety in a social relationship. Consumers with a high UAI use tangible cues in their environment to judge service quality because these visible features reduce uncertainty in a service outcome (Donthu & Yoo, 1998). Hence, a standardized physical environment and service flow enhance service evaluations of uncertainty-avoidant individuals by effectively reducing their anxiety from uncertainty. Deviations from standardized service

are high in unpredictability; however, even when these are pleasurable, their pleasantness is overshadowed by anxiety. Customized service, therefore, may not effectively enhance customer experience (Nakata & Sivakumar, 1996). A human touch in service dimensions may not align with an uncertainty avoidant culture.

Masculinity-Femininity (MAS) The masculinity dimension refers to whether a society values traits such as achievement or nurturing. Masculinity is closely related to societal expectations of differentiated gender roles (Hofstede, 2001). Masculine cultures stress ambition and material success and tend to have clearer distinctions between male and female roles. In contrast, feminine cultures are relationship-oriented and tend to value caring and nurturing behaviors. Individuals from a feminine society are concerned with quality of life and are apt to embrace more fluid gender roles (Hofstede, 1980, 2001). In a masculine culture, the pursuit of achievement is driven by a need for social admiration. Members of masculine cultures seek social cues to demonstrate and advance their achievements. Personalized service signals their worthiness; adaptive behavior by service staff is a testimonial of their social value. Thus, a human touch in the service quality dimension aligns with masculine cultures. Service customization should enhance customer satisfaction. Relative to those of feminine cultures, it plays a more critical role in service quality evaluations among members of masculine cultures.

Power Distance Index (PDI) The power distance dimension describes the acceptance of power that is established in relationships by social institutions (Hofstede, 1991). People with a high PDI are more likely to follow a hierarchy where everyone has a place (Herbig & Miller, 1992). In lower PDI cultures, characterized by more democratic or consultative relations, individuals tend to have more autonomy and are less concerned about status (Hofstede, 2001; Zhang et al., 2018). Consumers from a high PDI culture often feel superior to their service providers (Mattila, 1999). Since customers with a high PDI seek affirmation of their position in a social hierarchy, they deem themselves entitled to extraordinary responsiveness and demand individualized attention from service providers. During a service encounter, an alteration of standard treatment to accommodate personal requests serves as a gesture to reinforce this sense of superiority. Personalized service is a valuable experience for members of a high PDI culture that enhances customer satisfaction. Therefore, a human touch aligns with high PDI cultures. Relative to low PDI cultures, human touch plays a

much more key role in service quality evaluations of members of high PDI cultures.

Long-Term Orientation (LTO) The LTO cultural dimension expresses whether a society values a future-oriented perspective more than pragmatic short-term material/social success or emotional gratification (Hofstede & Bond, 1988; Hofstede, 1991). Populations with a high LTO uphold the virtues of perseverance and thrift. A strong work ethic is highly valued because it produces long-term rewards. Trust and reciprocity are encouraged to build and maintain relationships, reducing opportunistic behaviors (Hallikainen & Laukkanen, 2018; Wang et al., 2015). Short-term oriented societies, however, consider the present more important than the future, value tradition and the current social hierarchy and tend to emphasize rapid results. Individuals in these societies are also more sensitive to instant gratification from pleasurable pursuits (Hofstede, 2001; Yoon, 2009; Zhang et al., 2018). In cultures with a high LTO, customization functions as a dedication to maintaining a long-term relationship through proactive customer communication service provider adaptation. Once a long-term relationship is forged through mutual efforts, other service providers cannot enter into this relationship easily with standardized service. Personalized service should enhance customer satisfaction in a high LTO culture. Therefore, a human touch aligns with LTO cultures. Relative to short-term oriented cultures, human touch plays a much more key role in service quality evaluations of members of high LTO cultures.

Indulgence-Restraint (IND) This dimension entails a culture's tendencies concerning desire fulfillment. As this is a newly coined dimension, the present study is the first to link IND to the latent dimensions of service quality in UGC and with the previously discussed cultural dimensions. Indulgent cultures allow or encourage relatively unrestrained gratification of the fundamental and natural human desires that are related to hedonic experiences and leisure (Hofstede Insights, 2019). Their populations perceive themselves to be in control of their personal lives, consider the freedom of speech important, and deem themselves happy. Conversely, populations from cultures that value restraint tend to suppress their need to self-gratify and are regulated by strict social norms (Hofstede Insights, 2019; Huang & Crotts, 2019). Since indulgent cultures emphasize hedonic gratification, indulgent individuals expect staff to proactively make their stays extraordinarily enjoyable with swift responses and personal care. Customized service that offers a high sensory stimulation that is catered to individual needs is critical for creating a memorable experience for such individuals. Therefore, a human touch aligns with indulgent cultures. Relative to restraint cultures, human touch greatly

impacts service quality among members of indulgent cultures and enhances their customer satisfaction. Accordingly, we suggest the following.

H1: A human touch in service positively affects the review ratings of individuals from cultures with a high level of masculinity, power distance, long-term orientation, and indulgence.

H2: Systemization of service delivery positively affects the review ratings of individuals from cultures that are highly individualist and uncertainty avoidant.

Methods

To determine whether national cultures moderate the effects of the latent service dimensions derived from UGC on overall evaluations of hospitality products and services, an empirical study was conducted using unsupervised ML techniques. Building on Tirunillai & Tellis's work (2014), we automatically extracted a set of key attributes using LDA and ascertained the polarity of each of these aspects via ABSA.

The study comprised two phases. Phase 1 employed exploratory factor analysis (EFA) on 50% of the sample (Sample 1) to establish the dimensionality of service as reflected in the sentiment scores for key service attributes by following Hung and Guan (2020) and Pennebaker and King (1999). The analysis identified 15 top features in three distinct service dimensions. The factor structure was replicated on an independent sample (Sample 2) using confirmatory factor analysis (CFA). Phase 2 tested the hypothesized model using multigroup SEM. We formulated covariance-based structural equation modeling (CB-SEM) instead of partial least squares structural equation modeling (PLS-SEM) because our goal was to delineate the differential effects of UGC on service dimensions that are central to reviewers' cultural backgrounds from those that are peripheral to their cultural backgrounds. A comparison of alternative conditions requires several goodness-of-fit criteria, allowing us to further investigate whether the congruence between national culture and service expectation enhanced the predictive power of UGC for service evaluation. As the algorithm for obtaining PLS-SEM solutions is not based on minimizing the divergence between observed and estimated covariance matrices, the concept of Chi-square-based model fit measures—used in CB-SEM—is not applicable (Hair et al., 2019). In this phase, the analysis compared and contrasted the six dimensions of national culture regarding the impacts of service dimensions on overall evaluations, which manifest in hotel ratings. The two phases of analysis built on each other, supporting the hypothesis that national

culture influences how latent dimensions of UGC impact service evaluations in the hospitality industry.

Empirical investigation

Data

Online consumer reviews were obtained from a leading tourist review platform, TripAdvisor, using a web crawler. This platform was chosen based on its volume (number of unique reviews) and number of reviewers. TripAdvisor is one of the most widely investigated review platforms and has been selected as a source for data collection in numerous extant studies (Ayeh et al., 2013; Banerjee & Chua, 2016). However, cross-cultural heterogeneity in hotel reviews and ratings has not been widely investigated.

We extracted all reviews and reviewer-related information on hotels and resorts in Singapore listed on the Singapore Stock Exchange for the period of 2010–2015 from TripAdvisor. By choosing hotels in a single location, the study could control for any potential confounding factors that are related to different travel destinations (Salkind, 2010). By restricting the sampling frame to only a few of the largest hotel chains, the risk of having hotel-level outliers skew our results could be reduced. Appendix A summarizes the descriptive statistics of the hotels and reviews. A total of 10,004 reviews were drawn from these hotels. For every data entry, six fields were obtained, namely, hotel, review title, review rating, date, review content, and reviewer country of origin. Entries with missing variables were eliminated from the initial data pool. Our final dataset for empirical analysis included 9,257 reviews that were posted by reviewers from 148 countries.

Phase 2 of the study examined a culture's moderating impact on how the latent dimensions of UGC influence an overall hotel rating. Following Hofstede's national cultural framework (Hofstede, 2001), secondary data on cultural scores in six dimensions (IDV, MAS, PDI, UAI, LTO and IND) were collected. Hofstede's framework has been employed by several service quality studies that focus on the individual level (Donthu & Yoo, 1998; Furrer et al., 2000). Building upon these existing studies, this research extends the application of this framework to predict service evaluation-based UGC.

Phase 1

Service dimension extraction and validation were the primary contributions of Phase 1. We first explored the set of service quality attributes that describe hospitality products and services in UGC using LDA. LDA is superior to other extant methods, as it permits exploratory analysis of a text corpus and automatic extraction of the candidate terms. Here, corpus refers to a collection of all texts (documents). Previous studies have used LDA to identify product features expressed in online discussions (Ma et al., 2013; Tirunillai & Tellis, 2014). In line with previous literature, LDA is used to analyze and interpret service attributes contained in UGC. Furthermore, this study ascertains the valence of top features with ABSA (Jo & Oh, 2011; Liu, 2012). In the current context, valence is the expression of a positive or negative performance on an attribute and is termed "sentiment" in text-mining research (Tirunillai & Tellis, 2014). ABSA aims to detect the sentiment polarity that is associated with each aspect in a given text corpus (Jo & Oh, 2011; Liu, 2012; Pontiki et al., 2016). We deploy BERT-base ABSA to extract valence from textual contents for a number of reasons. First, sentiment analysis has been widely adopted to classify texts into binary or trinary categories, such as positive, neutral, or negative polarity, as an overall sentiment prediction of a text corpus (Pang & Lee, 2005; Turney, 2002). However, a full-length text with multiple clauses or sentences tends to contain multifaceted opinions and fuzzy sentiments in different aspects (Chiu, 2004; Wang, 2008). Thus, recent research efforts focus on identifying fine-grained opinion polarity regarding a specific aspect that is associated with a given entity, such as reviewed objects (Yu et al., 2011), comments on local restaurants (Kiritchenko et al., 2014), or question and answer pairs (Saeidi et al., 2016). When comments and reviews refer to more than one aspect, the task of ABSA is to determine these aspects and extract their valence scores. Second, we select the pretrained BERT language model (Devlin et al., 2019) for the ABSA task it is proven to have higher accuracy than other models, which rely heavily on feature engineering (Kiritchenko et al., 2014; Wagner et al., 2014). The BERT-based model, armed with bidirectional transformers, has become more popular among the natural language processing (NLP) community as a basis for various downstream tasks, such as aspect-oriented opinion word extraction (AOWE) (Fan et al., 2019) and aspect term level end-to-end aspect-based sentiment analysis (E2E-ABSA) (Li et al., 2019). We adopted the two-step BERT-base ABSA procedure, where the pretrained BERT language model is fine-tuned with domain-specific textual data that is biased toward more informal language and followed by a sequence-pair classification task (Sun et al., 2019).

Preparing text for analysis Text preprocessing is typically a crucial step for NLP applications. It transforms textual data into a more digestible format so that ML algorithms can achieve better performance. Words that are not informative, such as non-English characters, punctuations, and words that contain numbers and common English stop words

Fig. 1 ABSA Process
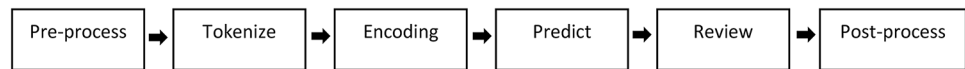
(e.g., "the," "and," "when," "is," "at," "which," "on," "in"), which are used to connect grammatical elements but are not required for connotation, are removed. Following Tirunillai and Tellis (2014), part-of-speech tagging is then applied to retain only nouns, adjectives, or adverbs—that is, words that tend to contain information about a service or service quality. This cleansed set forms the "corpus" of textual content that is used for subsequent analysis (Manning et al., 2008). Each review is treated as a single document. All of these steps are performed for each document in the sample.

Feature and associated valence extraction

Consumers express their opinions using words that represent one or more dimensions of a service quality that they believe are worth sharing via reviews. These dimensions of service quality are unobservable (hidden) for the researchers, while each review is a collection of words, chosen by consumers, that are observable. The latent dimensions can be inferred from the top words derived from the reviews. LDA is deployed for this purpose by using Gensim (Blei, 2012; Blei et al., 2003). Gensim is an open-source library for unsupervised NLP that is implemented in Python. The model identifies the top observed-features (highest-weighted words) in the reviews.

The ABSA is set as a multistep classification problem. A set of fixed aspects, set A, is defined based on the extracted features from LDA, e.g., set A = {pork, chef, kid, family}. Given a sentence S_i, the sentiment polarity $y \in$ {positive, neutral, negative} is predicted over the complete set of aspect terms. The BERT model detects the aspect terms and determines valence polarity y for each term. The BERT-base ABSA model is implemented using an open-source Python library aspect-based sentiment analysis (Rolczynski, 2020). This process is illustrated in Fig. 1. The texts are preprocessed and converted into individual sentences that can be tokenized, encoded, and predicted independently. A review step is added to supervise the prediction process, which dictates a discard if suspicious internal states and outputs are detected. The detailed steps of how the BERT model works are discussed below.

BERT, as a language representation model, explicitly represents sentences into encoded words and contextualized embeddings, which are expressed in vectors in the same continuous space (Devlin et al., 2019). A comparison of words and manipulation of sentences can be realized by simple mathematical operations, such as dot product or vector addition. For a given sentence S_i, the sentence and aspect information is converted into an auxiliary format of $(S_i, a_1), …, (S_i, a_n)$ for the sentence pair classification step using the masked language model and next-sequence prediction, which determines whether sequence A is naturally followed by sequence B (Rietzler et al., 2020). For example,

$$p = soft\max(W \cdot h[CLS] + b) \tag{1}$$

This automatic extraction of the candidate terms and valence scores allows subsequent dimension labeling and interpretation.

Results

Phase 1 results

Exploring the factor structures The first research phase examines the service dimensional structures of the attributes that have been extracted from the reviews. Past studies have attempted to factor analyze language use (Hung & Guan, 2020; Pennebaker & King, 1999). Their results have generally demonstrated the appropriateness of factor analysis for textual data. In the current study, the total sample of 9257 reviews is randomly divided into 2 independent subsamples, Sample 1 and Sample 2, each containing 50% of the reviews. Sample 1 is used to explore and derive the service dimensions that are reflected in the key features and in the associated valence scores that are derived from the UGC, and Sample 2 is evaluated to confirm the factor structure identified using Sample 1.

LDA was first applied to extract the top features of service quality in all the reviews in the sample. On the basis of a number of considerations that were adapted from Pennebaker and King's work (Pennebaker & King, 1999), the 15 most salient terms (with the highest weights) were entered into the ABSA model as seed words (aspects) to generate the associated valence scores. First, features were included only if they did not substantially overlap with other included variables or higher-level categories. For example, words representing "meat" (a higher-level category) came from words representing "pork". The category term "meat", in this case, is not included. Similarly, "shower" and "bathroom" represent analogous room features. Thus, the more generic version of the term, "bathroom", is kept. Second, categories that do not refer to specific service features (e.g., generic terms, such as hotel or day, and perceptual processes,

Table 3 Rotated factor loadings for the exploratory analysis of polarity scores of top features (Sample 1)

Polarity Scores of Top Features	Component		
	Factor 1 *Adaptability* (29.57% variance)	Factor 2 *Tangibles* (21.37% variance)	Factor 3 *Reliable Delivery* (21.02% variance)
pork	0.855		
chef	0.813		
ask	0.766		
kid	0.726		
family	0.647		0.527
check	0.511		
room		0.779	
bed		0.723	
bathroom	0.528	0.583	
pool		0.562	
breakfast		0.520	
staff			0.785
service			0.766
experience			0.686
location			0.551

Note: Only loadings of 0.50 or above are shown; N = 4596

such as see or look) are excluded. Third, words pertaining to the travel destination (i.e., Singapore) are not considered because they do not inherently describe the service quality of a hotel chain. Finally, only words with mean probability scores above 1% are included. The model generates three probability scores for each candidate term: a positive probability score (P), a negative probability score (N), and a neutral probability score. The sum of these three probability scores is 1. We then derive the polarity score of each candidate term by subtracting P from N, which ranges from −1 to 1. The higher the score is, the more likely that the candidate term's polarity classification is positive and vice versa. The polarity scores of the 15 candidate words are included in the initial EFA.

To examine the underlying dimensions of the 15 service variables, we conducted an EFA on Sample 1. The results indicate that a factor model is appropriate for the data: KMO = 0.962; Bartlett's test of sphericity = 54,831, $p < .001$. In addition, all individual measures of sampling adequacy are reasonably high, ranging from 0.924 to 0.985. Thus, the revised factor analysis is deemed appropriate. An examination of the scree plot indicates that a three-factor or four-factor solution would best fit the data. The total variance, explained with a three-factor solution, is 71.97%, close to the 75% of accounted variance suggested by Pett et al. (Pett et al., 2003). Therefore, the three-dimensional model is adopted, as the framework elicits a relatively more parsimonious and interpretable set of latent dimensions of service quality than the other models. Principal-component analysis extracts three factors, and varimax rotation is used to

facilitate interpretation of the factors. For the three-factor solution, all variables had communalities above 0.59.

The extracted dimensions and the corresponding rotated factor loadings are shown in Table 3. The keywords with the highest factor loadings relating to each dimension facilitate interpreting the characteristics that each dimension represents. Loadings on the first factor (rotated eigenvalue = 4.44) include terms associated with anticipating special dietary requests and making specific arrangements for families and children. Thus, "*adaptability*" is chosen as an appropriate label for this dimension. Following similar logic, the second factor (rotated eigenvalue = 3.21) includes attributes that are related to the physical facilities of hotels (e.g., room, bed). This factor is labeled the "*tangibles*" dimension. The third factor (rotated eigenvalue = 3.15) characterizes "*reliable delivery*" because it consists of words that convey the attributes of hotel service staff (e.g., "staff," "service," and "experience"). One of the limitations of the LDA method is that for certain dimensions, the automatic extraction of candidate words by weight scores may not express the words' entire connotation, especially when they are taken out of context. Such cases entail manual labeling and interpretation of the dimensions through human intervention. For each extracted dimension, ten documents regarding each feature for manual analysis are randomly selected. An in-depth manual investigation of the qualitative data not only enhances the understanding of context but also offers more insight into the cause or nature of an associated dimension. Exemplary quotes regarding each service attribute and its associated service dimension are summarized in Appendix B.

Table 4 Estimated construct correlation matrix from the CFA of Sample 2

		CR	AVE	1	2	3
1	Adaptability	0.93	0.67	*0.83*	0.87	0.89
2	Tangibles	0.85	0.54	0.04	*0.73*	0.90
3	Reliable Delivery	0.87	0.63	0.07	0.19	*0.79*

Notes: CR: composite reliability; AVE: average variance extracted; value on the diagonal is the square root of AVE. Below-diagonal values represent correlations between the latent variables. The above-diagonal values represent HTMT values between the latent variables

Table 5 Descriptive statistics of the factor scores and ratings across different culture groups (Sample 2)

		Individualism-Collectivism (IDV)	Masculinity-Femininity (MAS)	Power Distance (PDI)	Uncertainty Avoidance (UAI)	Long Term Orientation (LTO)	Indulgence-Restraint (IND)
Mean		IDV: High	MAS: High	PDI: High	UAI: High	LTO: High	IND: High
Rating		*4.36[a](0.93)*	*4.31(0.95)*	*4.10(1.02)*	*4.31(0.94)*	*4.16(1.02)*	*4.34(0.92)*
	Adaptability	0.04 [a] (1.00)	0.01(1.00)	-0.06(1.00)	0.05(0.99)	-0.03(1.01)	0.03(0.99)
	Tangibles	0.08(0.98)	0.04(0.98)	-0.09(1.02)	0.06(0.99)	-0.04(1.01)	0.06(0.99)
	Reliable Delivery	0.09(0.96)	0.05(0.97)	-0.10(1.03)	0.07(0.97)	-0.05(1.02)	0.07(0.97)
		IDV: Low	MAS: Low	PDI: Low	UAI: Low	LTO: Low	IND: Low
Rating		*4.08(1.03)*	*4.11(1.02)*	*4.36(0.93)*	*4.12(1.03)*	*4.31(0.94)*	*4.05(1.05)*
	Adaptability	-0.06(1.00)	-0.01(1.00)	0.05(0.99)	-0.07(1.10)	0.03(0.99)	-0.04(1.01)
	Tangibles	-0.10(1.01)	-0.06(1.02)	0.08(0.98)	-0.09(1.01)	0.04(0.99)	-0.10(1.02)
	Reliable Delivery	-0.11(1.04)	-0.08(1.04)	0.09(0.96)	-0.11(1.03)	0.05(0.98)	-0.11(1.04)

Note: [a] Mean (standard deviation in parentheses)

Confirmatory factor analysis To determine whether factor structures are replicable across independent samples, we employ CFA to examine the three-factor model that was identified using Sample 2. We run a three-factor CFA using Mplus 8. The results of our analysis indicate that the model fit indices are adequate, χ^2 (76) = 1567.81, CFI = 0.974, TLI = 0.965, RMSEA = 0.065, and SRMR = 0.024, and meet the standards recommended by Bagozzi and Yi (Bagozzi & Yi, 2012) and Hu and Bentler (Hu & Bentler, 1999).

Next, we examine the model's reliability and validity. First, the composite reliabilities of the measures are 0.60 or above, and all the average variances extracted are greater than or equal to 0.50, meeting the reliability standards recommended by Bagozzi and Yi 2012 and Fornell and Larcker (Fornell & Larcker, 1981). Second, all standardized item loadings are greater than 0.5, with the lowest at 0.629, affirming the convergent validity of the measures (Stevens, 2001). Third, discriminant validity is demonstrated when the square roots of the average variances extracted are greater than the correlations between the corresponding latent constructs (Fornell & Larcker, 1981). In addition to the Fornell-Larcker criterion of discriminant validity, heterotrait-monotrait ratios of correlations (HTMT) (Henseler et al., 2015) are also reported in Table 4. With values below or equal to 0.90, the HTMT criterion indicates that discriminant validity has been met (Henseler et al., 2015). The estimated construct correlation matrix from the confirmatory factor model of Sample 2 is shown in Table 4.

We also tested measurement invariance between each pair of national culture groups (high vs. low, based on a midpoint split) by constraining the factor loadings to be equal across the groups. The differences in the χ^2 fit statistic, attributed to the constraints, are not significant ($p > .10$). Therefore, the factor loadings are considered stable across the cultural groups. Overall, the measurement model is acceptable in terms of reliability and validity and is adopted in the subsequent analysis for hypothesis testing. Table 5 shows the descriptive statistics of factor scores and ratings for the different culture groups.

Phase 2

As discussed above, a consumer forms an expectation about services that is characterized by his or her national culture. We have classified a service dimension as aligned or misaligned with each national culture orientation based on the literature review (see Table 2). Users are thus apt to form their evaluations in accordance with these classifications. Phase 2 of the study examines a culture's moderating impact on how UGC predicts review ratings. The review dataset includes each reviewer's self-reported country of origin, and the national culture variables are derived from each review's country of origin. The values of the six national cultural dimensions are based on Hofstede (2019).

First, we estimate a one-group structural equation model (SEM), where all path coefficients are not differentiated between cultural groups. The one-group model

Table 6 Standardized coefficient estimates of the six two-group SEMs (Sample 2)

	Model I	Model II	Model III	Model IV	Model V	Model VI
Construct	Individualism -Collectivism (IDV)	Masculinity-Femininity (MAS)	Power Distance (PDI)	Uncertainty Avoidance (UAI)	Long Term Orientation (LTO)	Indulgence-Restraint (IND)
	IDV: High	MAS: High	PDI: High	UAI: High	LTO: High	IND: High
Rating*(R²)*	*0.456*[a]	*0.296*	*0.304*	*0.475*	*0.295*	*0.280*
Adaptability	-0.930*** [b]	0.463***	0.537***	-1.011***	0.442***	0.506***
Tangibles	0.444***	-0.063	-0.099	0.478***	-0.065	-0.078
Reliable Delivery	1.101***	0.150*	0.113	1.153***	0.172**	0.102!
	IDV: Low	MAS: Low	PDI: Low	UAI: Low	LTO: Low	IND: Low
Rating*(R²)*	*0.307*[a]	*0.524*	*0.458*	*0.300*	*0.472*	*0.550*
Adaptability	0.596*** [b]	-1.316***	-0.940***	0.575***	-0.925***	-1.240***
Tangibles	-0.101	0.577***	0.445***	-0.059	0.481***	0.561***
Reliable Delivery	0.055	1.392***	1.111***	0.029	1.078***	1.352***

Notes: [a] Variance explained

[b] $p < .1$; *: $p < .05$; **: $p < .01$; ***: $p < .001$ (two-tailed)

exhibits a reasonable model fit: $\chi^2 (89) = 3030.187$, p < .001; CFI = 0.952, TLI = 0.935, RMSEA = 0.084, SRMR = 0.031. Both the *adaptability* and *reliable delivery* dimensions of UGC are found to exert positive predictive effects on review ratings (for *adaptability*: $\beta = 0.139$, $p = .050$; for *reliable delivery*: $\beta = 0.483$, $p < .001$). However, the *tangibles* dimension shows no significant effect ($p < .1$), which may imply a more complex underlying moderating effect that may have partially negated its impacts, leading to nonsignificant results.

Hypothesis testing To test our research hypotheses, we formulate six multigroup models (Models I to VI) that incorporate the moderating relationships of each cultural dimension into the conceptual framework, and we use Mplus 8 to implement maximum likelihood estimation on the proposed models. Each model is organized into two groups, where each group represents a high or a low score in a particular cultural dimension based on a midpoint split. Table 6 shows the estimated results of the path models. For each model, in addition to comparing the directions and magnitudes of the path coefficients of *adaptability*, *tangibles*, and *reliable delivery* regarding review ratings, we use the Wald chi-squared difference test to determine if these three paths are the same for each high- and low-scoring culture group. Specifically, for each model, we first estimate an unconstrained model where all path coefficients are allowed to vary across the two conditions. We then introduce an alternative model by constraining *adaptability*, *tangibles*, and *reliable delivery* to be equal and conduct a contrast test on them.

UGC concerning service dimensions that are aligned with reviewers' cultural backgrounds are posited to predict review ratings. The coefficient estimates pertaining to the effects

of *adaptability* are first examined. *Adaptability* is expected to align with MAS, PDI, LTO and IND, as per H1. The model with MAS as a grouping variable (Model II) exhibits a reasonably good model fit: $\chi^2 (206) = 3957.27$, p < .001; CFI = 0.940, TLI = 0.930, RMSEA = 0.089, SRMR = 0.037. The analysis results show that the effects of *adaptability* on review rating vary across the high MAS and low MAS groups ($\chi^2 = 54.965$, df = 1, $p < .001$). Sentiments on *adaptability* only positively affect review ratings in cultures with a high MAS ($\beta = 0.463$, $p < .001$), not in feminine cultures ($\beta = -1.316$, $p < .001$). Similarly, high PDI cultures tend to value *adaptability* in service encounters. The two-group model by PDI (Model III) demonstrates a satisfactory model fit: $\chi^2 (206) = 3982.781$, p < .001; CFI = 0.940, TLI = 0.930, RMSEA = 0.089, SRMR = 0.039. The change in the χ^2 fit statistic, attributed to the grouping variable PDI, is significant for *adaptability* ($\chi^2 = 80.092$, df = 1, $p < .001$). The results show that *adaptability* has a positive effect on review ratings when PDI is high ($\beta = 0.537$, $p < .001$), but the effect is negative when PDI is low ($\beta = -0.940$, $p < .001$). Additionally, the fit indices of Model V that involve LTO as a grouping variable are acceptable ($\chi^2 (206) = 3946.535$, p < .001; CFI = 0.940, TLI = 0.930, RMSEA = 0.088, SRMR = 0.037), and the predictive effect of *adaptability* on review ratings is positive for the high LTO group ($\beta = 0.442$, $p < .001$) and negative for the low LTO group ($\beta = -0.925$, $p < .001$), and the χ^2 difference statistic that is attributed to this grouping is significant ($\chi^2 = 68.178$, df = 1, $p < .001$). Last, Model VI, which involves IND as the grouping variable, indicates a good fit: $\chi^2 (206) = 4084.846$, p < .001; CFI = 0.938, TLI = 0.928, RMSEA = 0.090, SRMR = 0.038. The change in the χ^2 fit statistic when the grouping constraint is applied to *adaptability* is significant ($\chi^2 = 60.931$, df = 1, $p < .001$): under high IND, the effect of *adaptability* is positive and

Appendix A Sampled hotel groups and subsidiary hotels

Hotel Groups	Subsidiary Hotels	No. of Reviews
Amara Holdings Limited	Amara Hotel	681
	Amara Sanctuary Resort Sentosa	730
Hotel Properties Limited (HPL)	Concorde Hotel Singapore	1032
	Four Seasons Hotel Singapore	1039
	Hilton Hotels Singapore	1458
Overseas Union Enterprise Limited (OUE)	Marina Mandarin Singapore	3219
	Crowne Plaza Changi Airport Hotel	2595
	Mandarin Orchard Singapore	5686
Global Premium Hotels Limited	Fragrance Hotel - Various Outlets	2974
	Parc Sovereign – Albert St	447
	Parc Sovereign Hotel - Tyrwhitt	217
Mandarin Oriental International Limited	Mandarin Oriental Singapore	3058
Hotel Grand Central Limited	Hotel Grand Central	186
Hotel Royal Limited	Hotel Royal	257
	Hotel Royal @ Queens	478
Shangri-La Asia Limited	Shangri-La Apartment, Singapore	90
	Shangri-La Hotel, Singapore	2880
	Shangri-La's Rasa Sentosa Resort & Spa	1216
	Hotel Jen Orchard Gateway Singapore	210
	Hotel Jen Tanglin Singapore	1789
Pan Pacific Hotels Group Limited	Pan Pacific Singapore	4092
	Pan Pacific Orchard	1182
	Pan Pacific Serviced Suites Beach Road	261
	Pan Pacific Serviced Suites Orchard	108

significant ($\beta = 0.506$, $p < .001$); under low IND, the effect of *adaptability* is negative ($\beta = -1.24$, $p < .001$). The above results support H1.

Next, *tangibles* and *reliable delivery* are posited to be culturally aligned with IDV and UAI, as per H2. Model I, with IDV as the grouping variable, exhibits decent fit: χ^2 (206) = 3991.971, p < .01; CFI = 0.940, TLI = 0.930, RMSEA = 0.089, SRMR = 0.039. As expected, when IDV is high, both *tangible* and *reliable delivery* are found to predict review ratings (for *tangibles*: $\beta = 0.444$, $p < .001$; for *reliable delivery:* $\beta = 1.101$, $p < .001$), but not when IDV is low ($p > .1$). Similarly, Model IV, where UAI acts as the grouping variable, shows a reasonable fit: χ^2 (206) = 3935.624, p < .001; CFI = 0.941, TLI = 0.931, RMSEA = 0.088, SRMR = 0.038. Both *tangibles* and *reliable delivery positively* affect review ratings when UAI is high (for *tangibles*: $\beta = 0.478$, $p < .001$; for *reliable delivery:* $\beta = 1.153$, $p < .001$). However, these effects are not significant when UAI is low ($p > .1$). Thus, the results regarding *tangibles* and *reliable delivery* support H2.

In addition to the proposed effects, we have observed a systematic negative relationship between *adaptability* and review ratings when MAS, PDI, LTO, and IND are low and when IDV and UAI are high. This could be due to a perception of "overgenerous" service delivery (Estelami & Maeyer, 2002; Imrie, 2005). Surprisingly, overattentive service has been found to negatively impact satisfaction.

Service encounters that exceed most service protocols may even be construed as sufficiently forceful and aggressive to threaten individuals' freedom of choice and self-determination (Ku et al., 2013). The *adaptability* service dimension emphasizes anticipating customer needs, which might impair the freedom of action and personal space of high IDV customers. Cultures that score high in UAI prefer clear and structured guidelines and may feel threatened by ambiguous situations (Akdeniz & Talay, 2013). Anticipating guests' needs and tailoring services to suit their special circumstances or requirements tend to cause service variability, but high UAI cultures have created beliefs and institutions that attempt to prevent variability. Feminine cultures (or those with low MAS) value a fluid social structure, mutual respect and environmental protection. Low PDI societies embrace a more equitable, collaborative culture and are more likely to believe that customers are equal to service staff in status. Societies with a low LTO prefer to maintain time-honored norms and view changes with suspicion. Restrained societies (low IND) are more likely to believe that gratification needs to be curbed and regulated by strict norms. Thus, people from these cultural orientations are more likely to consider unsolicited services superfluous or even disconcerting.

Appendix B Exemplary quotes for each service attribute and associated service dimension

Service Dimension	Top Features in UGC	Exemplary Quotes from UGC
Adaptability	Pork	*"When he learnt we had some views on the barbecue pork he offered to ask the chef to have a more slim one. ... To our even more surprise when Mr. NG Chong-kim the restaurant manager learnt our views on the barbecue pork he sincerely offered us a free desert."*
		"Staff actually told us that all food and meat supplies are Halal as we are Muslim except for pork bacon. Really appreciate her help"
	Chef	*"we called for in room dinning and ordered a set of steak sandwich... which was not found in the menu. Few minutes after placing our order we received a phone call and was told that the chef refused to take our order. We were told to refer to menu instead."*
		"a charming breakfast chef who took time to explain some of the dishes as well as asking how I was enjoying my stay"
	Ask	*"Assisted with highchair without me asking even offered to help me take some food so that my daughter is accompanied."*
		"A staff by the name of Nurain came to greet us with her warm smile asking my daughter on why is she crying"
	Kid	*"there are promotions like free breakfast or massage for every stay discounts on food or spa services or fun activities for the kids."*
		"Not very kid friendly as most guests are elderly here."
	Family	*"The staff were not so sensitive to help those who are in need though as there was a family that came with handicap (wheel-chair)."*
		"Before I arrived staff sent email as well as sms to ask me what I need some special things for my child/family and prepared everything in good conditions."
	Check	*"When I went down to the concierge to complain about the situation receptionist Dana helped me to check for room."*
		"we were late for SONG OF THE SEA PROGRAMME and were checking with them for the fastest way to reach beach. To our surprise hotel staff dropped us there in their buggy"
Tangibles	room	*"The rooms do need some renovation and touch ups."*
		"The furniture in the room were really old looking the room doesn't smell refreshing and the room lighting were not adequate."
	bed	*I stayed in 3 different rooms and they all seem to vary some have larger beds some have larger and newer tv's.*
		"The bed was huge and comfortable!"
	bathroom	*"The bathroom was huge and clean though."*
		"The bathroom was moldy and dirty sink was clogged"
	pool	*"Pool deck also need to replace with new wood."*
		"facilities - there a great outdoor pool spread over pretty landscaping."
	breakfast	*I always welcome breakfast to have some hot and fresh cook station and Shutter offered more than just the regular fresh eggs station.*
		"The breakfast buffet selection was not too bad with a mix of oriental and western selections."
Reliable Delivery	staff	*"There were delays during check-in as there were only 2 counter staff doing check-in. What i did not comprehend was there was another staff standing at the third counter but apparently he was only in charge of check-outs.*
		"Check in was a bit disorganized due to the crowd. We were fortunate that the staff serving us was very efficient unlike the others"
	service	*"the 30 minute wait for valet service only for them to bring the wrong car"*
		"The disappointment continued when I tried to place calls to Front Desk and Room Service but nobody picked up - even the Operator wouldn't pick up calls."
	experience	*Our experience concludes for us that we do not want to pay so much fora room that does not guarantee to deliver as promised.*
		"The check in experience was truly an awful one and its staff should be aware that having basic manners is necessary in the hospitality industry."
	location	*"Location was not the best as although there was a shuttle bus."*
		"Things need improvement is the staff knowledge about the facility. I asked the housekeeping staff about the gym location and the staff told me to go to front office and check."

General discussion

Online reviews have begun to demonstrate how culture is a prevalent factor that affects consumers' motivation to review (Min et al., 2018), linguistic style (Nakayama & Wan, 2018), and generated content (Mariani and Matarazzo (2020)). The richness of consumer-generated content provides an opportunity to reveal the service dimensions that

are relevant in today's rapidly evolving service contexts. Our research addresses two relevant questions in two phases of study. The first regards the service quality dimensions that are relevant to hospitality in blended technology and staff service delivery. The second concerns how the importance of service dimensions for customer satisfaction vary in cross-cultural contexts.

The first phase of this study establishes a three-dimensional model of service quality in UGC. These three service dimensions, adaptability, reliable delivery and tangibles, are found to be important predictors of overall service evaluations. The NLP analysis, in the first phase of the research, demonstrates the effectiveness of using textual UGC to reflect and measure consumers' preferences and the cues that affect their judgment of service quality. The adaptability dimension reflects the human touch in blended service, while the reliability and tangibles dimensions reflect systemization in standardized service.

The second phase of research validates that overall, service evaluations depend on an alignment between cultural values and the three dimensions of the ART model. The results in the second phase demonstrate this. UGC regarding service dimensions stem from reviewers' cultural backgrounds and predict their review ratings. Specifically, adaptability has a positive effect on review ratings in cultures with high MAS, PDI, LTO, and IND, while tangibles and reliable delivery positively impact review ratings in cultures with high IDV and UAI. The findings affirm our predictions that a human touch in service is less aligned with IDV and UAI than with other cultural dimensions due to a propensity of high IDV and UAI individuals to demand autonomy and certainty.

Theoretical contributions

The findings from the two phases of this research jointly inform how consumers' cultural backgrounds can enhance or attenuate the importance of certain service dimensions. To date, international research on service quality dimensions has undermined whether similar service quality dimensions can apply in international contexts (Ladhari, 2009), given the mixed findings of established frameworks, such as SERVQUAL (Parasuraman et al., 1991b; Parasuraman et al., 1985), and their links with Hofstede's cultural dimensions. A comparison with SERVQUAL leads to the following observations: First, tangibles and reliability remain important. However, based on our findings, reliability requires technology and staff to supplement each other to assure customers of timely responses. Second, adaptability indicates the continuous relevance of the human factor in a blended service context. Technology should augment human touch throughout the customer journey by personalizing

service encounters with a proactive rather than reactive approach, such as empathic responses after complaints. The ART three-dimensional model provides a parsimonious framework to explain when adaptability is valued, based on a cultural orientation.

Managerial contribution

Managerially, the findings show that as a result of their cultural orientations, consumers have varying perceptions of each of the dimensions of service quality; thus, UGC related to such service dimensions has a higher predictive power for service evaluations. Customizing service touchpoints that are congruent with customers' national cultural values may improve their overall service quality perceptions, increase repeat customers and ultimately generate favorable WOM. Specifically, and recently, UGC has been widely utilized by consumers to conduct prepurchase research and to reduce postpurchase cognitive dissonance. Online reviews have been found to affect their purchase decisions (Dellarocas et al., 2007; Zhu & Zhang, 2010). It is important to evaluate the customer journey and ensure that a proactive approach is taken for managing service encounters to ensure that they cater to different individual needs by considering cultural variations.

Our findings can inform marketers how to effectively customize and manage service touchpoints. Furthermore, they can help service employees better internalize customers' expectations and gain a deeper understanding of their needs, wants, and expectations. This in turn could result in better designed tourist products/services and destinations in general and service deliveries in particular. Beyond hotel and accommodation services, the research findings can be extended to UGC for other hospitality products and services, such as food and beverage outlets, theme parks, and transportation. These service environments consist of different proportions of subjective and objective attributes, but their service dimensions can be identified using a similar approach. Brand managers can derive cues that reflect customers' expectancies from mining UGC.

Our results provide platform managers with practical guidelines for crafting analytical algorithms on new classification techniques that are based on content generated by disparate national, regional, or ethnic cultures. With reference to Hofstede's national cultural framework, managers can classify customers into distinct segments and identify effective value propositions that cater to each segment. With these features in mind, merchants who participate in online platforms can benefit from their built-in recommendation systems and push targeted promotional messages through these platforms.

Limitations and future research

As one of the first studies that investigates the effects of cultural orientation on UGC regarding service encounters, this paper has some limitations that offer paths for future research. First, it is likely that the congruency effect between culture and service attributes influences expectations and leads to differences in purchase behaviors. The current research did not capture the totality of these impacts. Future research can examine whether cultural orientations affect both subjective and objective attributes of experience goods and consumption in the service sector. Second, this study uses country-level cultural dimensions as its main explanatory variables for UGC. A national culture may not fully capture the considerable variations in behaviors of individuals or ethnic subgroups within a country (Srite & Karahanna, 2006). Therefore, a possible extension of this research is to examine how the unique characteristics of subcultures and lifestyles interact with service environments and influence consumer responses. Furthermore, consumer expectations could be shaped by personal, cultural, and situational factors (Kopalle et al., 2010). It is plausible that other personal factors (e.g., prior knowledge) or hotel characteristics (e.g., branding or pricing) can also influence consumers' expectations and evaluation. A pricing strategy may alter expectations among customers through distinctive dimensions. Customization is associated with premium pricing. Premium pricing may heighten expectations concerning tangibles, reliability or adaptability in hotel services. However, the actual price that a customer pays can differ significantly at the same hotel for a stay due to seasonality or promotional pricing. Many customers may pay with their accumulated points rather than cash. Our data did not capture advertised prices or the actual prices that customers pay. Future studies can examine how pricing affects customer expectations and experiences using experimental designs.

Third, the focus of the current study is on hotel-level service features and dimensions, and thus we selected hotels in a single location to minimize the confounding factors that relate to the features of any destination (Salkind, 2010). Future research can investigate hotels at different travel destinations, such as places known for their cosmopolitan beauty or natural landscapes, and analyze how the key considerations of hotel level service dimensions interact with destination features to affect overall evaluation. Finally, our sample comprises only listed companies that are well known, chain-branded hotels. A possible extension of this research is to investigate how consumer traits and hotel characteristics interact with cultural dimensions and influence consumer reviews. For example, if a consumer visits an obscure boutique hotel, he or she might pay more attention to tangibles and reliable delivery. In addition, home-sharing has been designated the future of tourism in the worldwide sharing economy (Lim et al., 2021). Future research could evaluate the differences in service evaluation frameworks between traditional hotels and home-shares.

Electronic supplementary material

References

Ababneh, H. T., & Shrafat, F. D. (2014). Analysing the role of culture in IS studies. *International Business Research*, *7*(11), 165–173. https://doi.org/10.5539/ibr.v7n11p165

Akdeniz, M. B., & Talay, M. B. (2013). Cultural variations in the use of marketing signals: a multilevel analysis of the motion picture industry. *Journal of the Academy of Marketing Science*, *41*(5), 601–624. https://doi.org/10.1007/s11747-013-0338-5

Al-Ababneh, M. (2017). *Service quality in the hospitality industry*. Available at SSRN 3633089. https://doi.org/10.2139/ssrn.3633089

Ayeh, J. K., Au, N., & Law, R. (2013). "Do we believe in TripAdvisor?" Examining credibility perceptions and online travelers' attitude toward using user-generated content. *Journal of Travel Research*, *52*(4), 437–452. https://doi.org/10.1177/0047287512475217

Bagozzi, R. P., & Yi, Y. (2012). Specification, evaluation, and interpretation of structural equation models. *Journal of the Academy of Marketing Science*, *40*(1), 8–34. https://doi.org/10.1007/s11747-011-0278-x

Banerjee, S., & Chua, A. Y. K. (2016). In search of patterns among travellers' hotel ratings in TripAdvisor. *Tourism Management*, *53*, 125–131. https://doi.org/10.1016/j.tourman.2015.09.020

Berry, L. L., Parasuraman, A., & Zeithaml, V. A. (1988). SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, *64*, 12–40

Beugelsdijk, S., Kostova, T., & Roth, K. (2017). An overview of Hofstede-inspired country-level culture research in international business since 2006. *Journal of International Business Studies*, *48*, 30–47. https://doi.org/10.1057/s41267-016-0038-8

Bickart, B., & Schindler, R. M. (2001). Internet forums as influential sources of consumer information. *Journal of Interactive Marketing*, *15*(3), 31–40. https://doi.org/10.1002/dir.1014

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77–84. https://doi.org/10.1145/2133806.2133826

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022

Büschken, J., & Allenby, G. M. (2016). Sentence-Based text analysis for customer reviews. *Marketing Science*, *35*(6), 831–998. https://doi.org/10.1287/mksc.2016.0993

Chevalier, J. A., & Mayzlin, D. (2006). The effect of Word of Mouth on sales: online book reviews. *Journal of Marketing Research*, *43*(3), 345–354. https://doi.org/10.1509/jmkr.43.3.345

Chiu, C. M. (2004). Towards a hypermedia-enabled and web-based data analysis framework. *Journal of Information Science*, *30*(1), 60–72. https://doi.org/10.1177/0165551504041679

Chu, S. C., & Choi, S. M. (2011). Electronic Word-of-Mouth in social networking sites: a cross-cultural study of the United States and China. *Journal of Global Marketing*, *24*(3), 263–281. https://doi.org/10.1080/08911762.2011.592461

Chung, T. S., Rust, R. T., & Wedel, M. (2009). My mobile music: an adaptive personalization system for digital audio players. *Marketing Science*, *28*(1), 52–68. https://doi.org/10.1287/mksc.1080.0371

Chung, T. S., Wedel, M., & Rust, R. T. (2016). Adaptive personalization using social networks. *Journal of the Academy of Marketing Science*, *44*, 66–87. https://doi.org/10.1007/s11747-015-0441-x

Coelho, P. S., & Henseler, J. (2012). Creating customer loyalty through service customization. *European Journal of Marketing*, *46*(3/4), 331–356. https://doi.org/10.1108/03090561211202503

Dedj, G., & Pavlovic, D. K. (2011). A tale of two nations - empirical examination of influence of national culture on perceived service quality. *International Journal of Management Cases, 13*, 92–104

Dellarocas, C., Zhang, X. Q., & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: the case of motion pictures. *Journal of Interactive Marketing*, *21*(4), 23–45. https://doi.org/10.1002/dir.20087

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*, Minneapolis, Minnesota

Ding, Y., & Keh, H. T. (2016). A re-examination of service standardization versus customization from the consumer's perspective. *Journal of Services Marketing, 30*(1), 16–28. https://doi.org/10.1108/JSM-02-2015-0088

Donthu, N., & Yoo, B. (1998). Cultural Influences on service quality expectations. *Journal of Service Research*, *1*(2), 178–186. https://doi.org/10.1177/109467059800100207

Doorn, J. V., Mende, M., Noble, S. M., Hulland, J., Ostrom, A. L., Grewal, D., & Petersen, J. A. (2016). Domo Arigato Mr. Roboto: emergence of automated social presence in organizational frontlines and customers service experiences. *Journal of Service Research*, *20*(1), 43–58. https://doi.org/10.1177/1094670516679272

Duan, W., Gu, B., & Whinston, A. B. (2008). Do online reviews matter? — An empirical investigation of panel data. *Decision Support Systems*, *45*(4), 1007–1016. https://doi.org/10.1016/j.dss.2008.04.001

Estelami, H., & Maeyer, P. D. (2002). Customer reactions to service provider overgenerosity. *Journal of Service Research*, *4*(3), 205–216. https://doi.org/10.1177/1094670502004003004

Fan, Z., Wu, Z., Dai, X. Y., Huang, S., & Chen, J. (2019). Target-oriented opinion words extraction with target-fused neural sequence labeling. *Proceedings of NAACL-HLT 2019*, Minneapolis, Minnesota

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50. https://doi.org/10.1177/002224378101800104

Franke, N., Keinz, P., & Schreier, M. (2008). Complementing mass customization toolkits with user communities: how peer input improves customer self-design. *Journal of Product Innovation Management, 25*(6), 546–559. https://doi.org/10.1111/j.1540-5885.2008.00321.x

Franke, N., Keinz, P., & Steger, C. J. (2009). Testing the value of customization: when do customers really prefer products tailored to their preferences? *Journal of Marketing*, *73*(5), 103–121. https://doi.org/10.1509/jmkg.73.5.103

Furrer, O., Liu, B. S. C., & Sudharshan, D. (2000). The relationships between culture and service quality perceptions: basis for cross-cultural market segmentation and resource allocation. *Journal of Service Research*, *2*(4), 355–371. https://doi.org/10.1177/109467050024004

Gao, B., Li, X., Liu, S., & Fang, D. (2018). How power distance affects online hotel ratings: the positive moderating roles of hotel chain and reviewers' travel experience. *Tourism Management*, *65*, 176–186. https://doi.org/10.1016/j.tourman.2017.10.007

Gonçalves, H. M., Silva, G. M., & Martins, T. G. (2018). Motivations for posting online reviews in the hotel industry. *Psychology & Marketing, 35*(11), 807–817. https://doi.org/10.1002/mar.21136

Gwinner, K. P., Bitner, M. J., Brown, S. W., & Kumar, A. (2005). Service customization through employee adaptiveness. *Journal of Service Research*, *8*(2), 131–148. https://doi.org/10.1177/1094670505279699

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Hallikainen, H., & Laukkanen, T. (2018). National culture and consumer trust in e-commerce. *International Journal of Information Management*, *38*(1), 97–106. https://doi.org/10.1016/j.ijinfomgt.2017.07.002

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Herbig, P. A., & Miller, J. C. (1992). Culture and technology: does the traffic move in both directions? *Journal of Global Marketing, 6*(3), 75–104. https://doi.org/10.1300/J042v06n03_05

Hofstede, G., & Bond, M. H. (1988). 1988/03/01/). The Confucius connection: from cultural roots to economic growth. *Organizational Dynamics*, *16*(4), 5–21. https://doi.org/10.1016/0090-2616(88)90009-5

Hofstede, G. H. (1980). *Culture's Consequences*. Sage Publications

Hofstede, G. H. (1991). *Culture and Organisations: software of the Mind*. McGraw Hill

Hofstede, G. H. (2001). *Culture's Consequences: somparing Values, Behaviors, Institutions and Organizations Across Nations*. Sage Publications

Hofstede Insights (2019). *National culture*. Hofstede Insights. Retrieved 15 December from https://hi.hofstede-insights.com/national-culture

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huang, S., & Crotts, J. (2019). Relationships between Hofstede's cultural dimensions and tourist satisfaction: a cross-country cross-sample examination. *Tourism Management*, *72*, 232–241. https://doi.org/10.1016/j.tourman.2018.12.001

Hung, Y. C., & Guan, C. (2020). Winning box office with the right movie synopsis. *European Journal of Marketing*, *54*(3), 594–614. https://doi.org/10.1108/EJM-01-2019-0096

Hwang, Y. (2005). Investigating enterprise systems adoption: uncertainty avoidance, intrinsic motivation, and the technology acceptance model. *European Journal of Information Systems*, *14*(2), 150–161. https://doi.org/10.1057/palgrave.ejis.3000532

Imrie, B. C. (2005). Beyond disconfirmation: the role of generosity and surprise. *International Marketing Review*, *22*(3), 369–383. https://doi.org/10.1108/02651330510602259

Inglehart, R., & Baker, W. E. (2000). Modernization, Cultural Change, and the Persistence of Traditional Values. *American Sociological Review, 65*(1), 19–51. https://doi.org/10.2307/2657288

Ivanov, S., & Webster, C. (2019). *Robots, Artificial Intelligence, and Service Automation in Travel, Tourism and Hospitality*. Emerald Publishing Limited. https://doi.org/10.1108/978-1-78756-687-320191014

Jo, Y., & Oh, A. H. (2011). Aspect and sentiment unification model for online review analysis. *WSDM '11:Proceedings of the fourth ACM international conference on Web search and data mining*, 814–824

Kandampully, J., Zhang, T., & Jaakkola, E. (2017). Customer experience management in hospitality: a literature synthesis, new understanding, and research agenda. *International Journal of Contemporary Hospitality Management*, 30(1), 21–56. https://doi.org/10.1108/IJCHM-10-2015-0549

Kannan, P. K., & Healey, J. (2011). Service Customization Research: a Review and Future Directions. In S. J. Demirkan, & H. Krishna V. (Eds.), *The Science of Service Systems. Service Science: Research and Innovations in the Service Economy*. Springer. https://doi.org/10.1007/978-1-4419-8270-4_17

Kattara, H. S., & El-Said, O. A. (2014). Customers' preferences for new technology-based self-services versus human interaction services in hotels. *Tourism & Hospitality Research*, 13(2), 67–82. https://doi.org/10.1177/1467358413519261

Kiritchenko, S., Zhu, X., Cherry, C., & Mohammad, S. M. (2014). Nrc-canada-2014: detecting aspects and sentiment in customer reviews. *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, Dublin, Ireland

Knutson, B., Stevens, P., Wullaert, C., Patton, M., & Yokoyama, F. (1990). Lodgserv: a service quality index for the lodging industry. *Hospitality Research Journal*, 14(2), 277–284. https://doi.org/10.1177/109634809001400230

Kopalle, P. K., Lehmann, D. R., & Farley, J. U. (2010). Consumer expectations and culture: the effect of belief in karma in India. *Journal of Consumer Research*, 37(2), 251–263. https://doi.org/10.1086/651939

Kozinets, R. V. (2016). Amazonian forests and trees: multiplicity and objectivity in studies of online consumer-generated ratings and reviews, a commentary on de langhe, fernbach, and lichtenstein. *Journal of Consumer Research*, 42(6), 834–839. https://doi.org/10.1093/jcr/ucv090

Ku, H. H., Kuo, C. C., & Chen, M. (2013). Is maximum customer service always a good thing? Customer satisfaction in response to overattentive service. *Managing Service Quality: An International*, 23(5), 437–452. https://doi.org/10.1108/MSQ-10-2012-0142

Kucukusta, D., Heung, V. C. S., & Hui, S. (2014). Deploying self-service technology in luxury hotel brands: perceptions of business travelers. *Journal of Travel and Tourism Marketing*, 31(1), 55–70. https://doi.org/10.1080/10548408.2014.861707

Kueh, K., & Ho Voon, B. (2007). Culture and service quality expectations. *Managing Service Quality: An International Journal*, 17(6), 656–680. https://doi.org/10.1108/09604520710834993

Ladhari, R. (2009). A review of twenty years of SERVQUAL research. *International Journal of Quality and Service Sciences, 1*(2), 172–198. https://doi.org/10.1108/17566690910971445

Larivière, B., Bowen, D., Andreassen, T. W., Kunz, W., Sirianni, N. J., Voss, C., Wünderlich, N. V., & De Keyser, A. D. (2017). "Service Encounter 2.0": an investigation into the roles of technology, employees and customers. *Journal of Business Research*, 79, 238–246. https://doi.org/10.1016/j.jbusres.2017.03.008

Leischnig, A., Kasper-Brauer, K., & Thornton, S. C. (2018). Spotlight on customization: an analysis of necessity and sufficiency in services. *Journal of Business Research*, 89, 385–390. https://doi.org/10.1016/j.jbusres.2017.12.038

Li, H., Ye, Q., & Law, R. (2013). Determinants of customer satisfaction in the hotel industry: an application of online review analysis. *Asia Pacific Journal of Tourism Research*, 18(7), 784–802. https://doi.org/10.1080/10941665.2012.708351

Li, X., Bing, L., Zhang, W., & Lam, W. (2019). Exploiting BERT for end-to-end aspect-based sentiment analysis. *Proceedings of the 5th Workshop on Noisy User-generated Text (W-NUT 2019)*, Hong Kong, China

Lim, W. M., Yap, S. F., & Makkar, M. (2021). Home sharing in marketing and tourism at a tipping point: what do we know, how do we know, and where should we be heading? *Journal of Business Research*, 122, 534–566. https://doi.org/10.1016/j.jbusres.2020.08.051

Liu, B. (2012). *Sentiment Analysis and Opinion Mining* (Vol. 5). Morgan & Claypool Publishers. https://doi.org/10.2200/S00416ED1V01Y201204HLT016

Ma, B., Zhang, D., Yan, Z., & Kim, T. (2013). An Lda and Synonym Lexicon Based Approach to Product Feature Extraction from Online Consumer Product Reviews. *Journal of Electronic Commerce Research*, 14(4), 304–314. https://doi.org/10.3846/16111699.2012.665383

Makarem, S. C., Mudambi, S. M., & Podoshen, J. S. (2009). Satisfaction in technology-enabled service encounters. *Journal of Services Marketing*, 23(3), 134–144. https://doi.org/10.1108/08876040910955143

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press

Mariani, M. M., & Matarazzo, M. (2020). Does cultural distance affect online review ratings? Measuring international customers' satisfaction with services leveraging digital platforms and big data. *Journal of Management and Governance*, 1–22. https://doi.org/10.1007/s10997-020-09531-z

Mariani, M. M., & Predvoditeleva, M. (2019). How do online reviewers' cultural traits and perceived experience influence hotel online ratings? An empirical analysis of the muscovite hotel sector. *International Journal of Contemporary Hospitality Management*, 31(12), 4543–4573. https://doi.org/10.1108/IJCHM-11-2018-0927

Mattila, A. S. (1999). The Role of culture in the service evaluation process. *Journal of Service Research*, 1(3), 250–261. https://doi.org/10.1177/109467059913006

Mattila, A. S. (2000). The impact of culture and gender on customer evaluations of service encounters. *Journal of Hospitality and Tourism Research*, 24(2), 263–273. https://doi.org/10.1177/109634800002400209

Matzler, K., Renzl, B., & Rothenberger, S. (2006). Measuring the relative importance of service dimensions in the formation of price satisfaction and service satisfaction: a case study in the hotel industry. *Scandinavian Journal of Hospitality and Tourism*, 6(3), 179–196. https://doi.org/10.1080/15022250600872850

Min, K. J., Jun, M., & Kim, C. K. (2018). The effects of culture on consumers' consumption and generation of online reviews. *Journal of Interactive Marketing, 43,* 134–150. https://doi.org/10.1016/j.intmar.2018.05.002

Nakata, C., & Sivakumar, K. (1996). National culture and new product development: an integrative review. *Journal of Marketing*, 16(1), 61–72. https://doi.org/10.1016/0737-6782(96)82491-X

Nakayama, M., & Wan, Y. (2018). Is culture of origin associated with more expressions? An analysis of yelp reviews on japanese restaurants. *Tourism Management, 66*, 329–338. https://doi.org/10.1016/j.tourman.2017.10.019

Naumov, N. (2019). The impact of robots, artifiial intelligence, and service automation on service quality and service experience in hospitality. In S. Ivanov & C. Webster (Eds.), *Robots, Artificial Intelligence, and Service Automation in Travel, Tourism and Hospitality* (pp. 123–133). Emerald Publishing Limited. https://doi.org/10.1108/978-1-78756-687-320191007

Ostrom, A. L., Parasuraman, A., Bowen, D. E., Patrício, L., & Voss, C. A. (2015). Service research priorities in a rapidly changing context, 18(2),127–159. https://doi.org/10.1177/1094670515576315

Pang, B., & Lee, L. (2005). Seeing stars: exploiting class relationships for sentiment categorization with respect to rating scales. *ACL '05: Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, Ann Arbor, Michigan

Parasuraman, A., Berry, L. L., & Zeithaml, V. A. (1991a). Refinement and reassessment of the SERVQUAL scale. *Journal of Retailing*, 67(4), 420–450. https://doi.org/10.1021/nl0492436

Parasuraman, A., Berry, L. L., & Zeithaml, V. A. (1991b). Understanding customer expectations of service. *Sloan Management Review*, *32*(6), 39–48. https://doi.org/10.1021/ja00787a040

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A conceptual model of service quality and its implications for future research. *Journal of Marketing*, *49*(4), 41–50. https://doi.org/10.1177/002224298504900403

Pennebaker, J. W., & King, L. A. (1999). Linguistic Styles: language Use as an Individual Difference. *Journal of Personality & Social Psychology*, *77*(6), 1296–1312. https://doi.org/10.1037/0022-3514.77.6.1296

Pett, M. M., Lackey, N. R., & Sullivan, J. (2003). *Making Sense of Factor Analysis*. SAGE Publications

Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., Al-Smadi, Al-Ayyoub, M. M. M., Zhao, Y., Qin, B., De Clercq, O., Hoste, V., Apidianaki, M., Tannier, X., Loukachevitch, N., Kotelnikov, E., Bel, N., Jiménez-Zafra, S. M., & Eryiğit, G. (2016). SemEval-2016 task 5: Aspect based sentiment analysis. *International Workshop on Semantic Evaluation*, 19–30. https://doi.org/10.18653/v1/S16-1002

Radojevic, T., Stanisic, N., & Stanic, N. (2017). Inside the rating scores: a multilevel analysis of the factors influencing customer satisfaction in the hotel industry. *Cornell Hospitality Quarterly*, *58*(2), 134–164. https://doi.org/10.1177/1938965516686114

Radojevic, T., Stanisic, N., & Stanic, N. (2019). The culture of hospitality: from anecdote to evidence. *Annals of Tourism Research*, *79*, https://doi.org/10.1016/j.annals.2019.102789

Radojevic, T., Stanisic, N., Stanic, N., & Davidson, R. (2018). The effects of traveling for business on customer satisfaction with hotel services. *Tourism Management, 67*, 326–341. https://doi.org/10.1016/j.tourman.2018.02.007

Rietzler, A., Stabinger, S., Opitz, P., & Engl, S. (2020). Adapt or get left behind: domain adaptation through BERT language model finetuning for aspect-target sentiment classification. *Proceedings of the 12th Language Resources and Evaluation Conference*, Marseille, France. https://doi.org/10.48550/arXiv.1908.11860

Robinson, S., Orsingher, C., Alkire, L., Keyser, A. D., Giebelhausen, M., Papamichail, K. N., Shams, P., & Sobhy Temerak, M. (2019). Frontline encounters of the AI kind: an evolved service encounter framework. *Journal of Business Research*, *16*, 366–376. https://doi.org/10.1016/j.jbusres.2019.08.038

Rolczynski, R. (2020). *Aspect Based Sentiment Analysis: Transformer & Interpretability (TensorFlow)*. Python Package Index. https://pypi.org/project/aspect-based-sentiment-analysis/

Saeidi, M., Bouchard, G., Liakata, M., & Riedel, S. (2016). SentiHood: targeted aspect based sentiment analysis dataset for urban neighbourhoods. *Proceedings of the 26th International Conference on Computational Linguistics (COLING)*, Osaka, Japan

Salkind, N. J. (2010). *Encyclopedia of Research Design Qualitative Research*. SAGE Publications, Inc. https://doi.org/10.4135/9781412961288.n350

Sandoff, M. (2005). Customization and standardization in hotels - A paradox or not? *International Journal of Contemporary Hospitality Management*, *17*(6), 529–535. https://doi.org/10.1108/09596110510612167

Schwartz, S. H. (2006). A theory of cultural value orientation: explication and applications. *Comparative Sociology*, *5*(2–3), 137–182. https://doi.org/10.1163/156913306778667357

Sent, E. M., & Kroese, A. L. J. (2020). Commemorating geert hofstede, a pioneer in the study of culture and institutions. *Journal of Institutional Economics*, 1–13. https://doi.org/10.1017/S174413742000051X

Simonson, I. (2005). Determinants of customers' responses to customized offers: conceptual framework and research propositions. *Journal of Marketing*, *69*(1), 32–45. https://doi.org/10.1509/jmkg.69.1.32.55512

Solnet, D., Baum, T., Robinson, R. N. S., & Lockstone-Binney, L. (2016). What about the workers? Roles and skills for employees in hotels of the future. *Journal of Vacation Marketing, 22*(3), 212–226. https://doi.org/10.1177/1356766715617403

Srite, M., & Karahanna, E. (2006). The role of espoused national cultural values in technology acceptance. *MIS Quarterly*, *30*(3), 679–704. https://doi.org/10.2307/25148745

Stępień, M., & Dudek, M. (2021). Toward the three–level power distance concept: expanding geert hofstede's power distance beyond cross–cultural context. *Studia Socjologiczne*, *1*, 61–87. https://doi.org/10.24425/sts.2021.136279

Stevens, J. P. (2001). *Applied Multivariate Statistics for the Social Sciences* (4th ed.). Psychology Press

Sun, C., Huang, L., & Qiu, X. (2019). Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence. *The 2019 Conference of the North American Chapter of the Association for Computational Linguistics*, Minneapolis, Minnesota

Taylan Dortyol, I., Varinli, I., & Kitapci, O. (2014). How do international tourists perceive hotel quality? *International Journal of Contemporary Hospitality Management*, *26*(3), 470–495. https://doi.org/10.1108/IJCHM-11-2012-0211

Terpstra, M., & Verbeeten, F. H. M. (2014). Customer satisfaction: cost driver or value driver? Empirical evidence from the financial services industry. *European Management Journal*, *32*(3), 499–508. https://doi.org/10.1016/j.emj.2013.07.001

Tirunillai, S., & Tellis, G. J. (2014). Mining marketing meaning from online chatter: strategic brand analysis of big data using latent dirichlet allocation. *Journal of Marketing Research*, *51*(4), 463–479. https://doi.org/10.1509/jmr.12.0106

Torres, E. N., Fu, X., & Lehto, X. (2014). Examining key drivers of customer delight in a hotel experience: a cross-cultural perspective. *International Journal of Hospitality Management*, *36*, 255–262. https://doi.org/10.1016/j.ijhm.2013.09.007

Triandis, H. C., Bontempo, R., Villareal, M. J., Asai, M., & Lucca, N. (1988). Individualism and collectivism: cross-cultural perspectives on self-ingroup relationships. *Journal of Personality and Social Psychology*, *54*(2), 323–338. https://doi.org/10.1037/0022-3514.54.2.323

TripAdvisor (2019). *2019 TripAdvisor Review Transparency Report*

Trusov, M., Bucklin, R. E., & Pauwels, K. (2009). Modelling effects of word-of-mouth versus traditional marketing findings from an internet social networking site. *Journal of Marketing*, *73*(5), 90–102. https://doi.org/10.1509/jmkg.73.5.90

Tsaur, S. H., Lin, C. T., & Wu, C. S. (2005). Cultural differences of service quality and behavioral intention in tourist hotels. *Journal of Hospitality and Leisure Marketing, 13*(1), 41–63. https://doi.org/10.1300/J150v13n01_04

Tsikriktsis, N. (2002). Does culture influence web site quality expectations?: an empirical study. *Journal of Service Research*, *5*(2), 101–112. https://doi.org/10.1177/109467002237490

Tsiligiris, V., Kéri, A., & Cheah, J. E. T. (2022). Exploring the relationship between student individual culture and service quality expectations in higher education. *Quality Assurance in Education*, *30*(1), 51–72. https://doi.org/10.1108/QAE-11-2020-0137

Tsoukatos, E., & Rand, G. K. (2007). Cultural influences on service quality and customer satisfaction: evidence from greek insurance. *Managing Service Quality: An International Journal*, *17*(4), 467–485. https://doi.org/10.1108/09604520710760571

Turney, P. D. (2002). Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics ACL*, Philadelphia. https://doi.org/10.48550/arXiv.cs/0212032

Wagner, J., Arora, P., Cortes, S., Barman, U., Bogdanova, D., Foster, J., & Tounsi, L. (2014). DCU: aspect-based polarity classification for SemEval task 4. *Proceedings of the 8th International*

Workshop on Semantic Evaluation (SemEval 2014), Dublin, Ireland

Wang, C. L., Shi, Y., & Barnes, B. R. (2015). The role of satisfaction, trust and contractual obligation on long-term orientation. *Journal of Business Research*, *68*(3), 473–479. https://doi.org/10.1016/j.jbusres.2014.09.016

Wang, E. J., Berthon, P. R., & Bechwati, N. N. (2020). The effect of mindfulness on service encounter quality. *Journal of Services Marketing*. https://doi.org/10.1108/JSM-05-2020-0170

Wang, J. (2008). *Encyclopedia of Data Warehousing and Mining*. Information Science reference

Winer, R. S., & Fader, P. S. (2016). Objective vs. online ratings: are low correlations unexpected and does it matter? a commentary on de langhe, fernbach, and lichtenstein. *Journal of Consumer Research*, *42*(6), 846–849. https://doi.org/10.1093/jcr/ucv092

Wong, O., Mei, A., Dean, A. M., & White, C. J. (1999). Analysing service quality in the hospitality industry. *Managing Service Quality: An International Journal, 9*(2), 136–143. https://doi.org/10.1108/09604529910257920

Wuyts, S., & Geyskens, I. (2005). The formation of buyer-supplier relationships: detailed contract drafting and close partner selection. *The Journal of Marketing*, *69*(4), 103–117. https://doi.org/10.1509/jmkg.2005.69.4.103

Xia, L., & Bechwati, N. N. (2008). Word of mouse: the role of cognitive personalization in online consumer reviews. *Journal of Interactive Advertising*, *9*(1), 3–13. https://doi.org/10.1080/15252019.2008.10722143

Yoon, C. (2009). The effects of national culture values on consumer acceptance of e-commerce: online shoppers in china. *Information and Management*, *46*(5), 294–301. https://doi.org/10.1016/j.im.2009.06.001

Yu, J., Zha, Z. J., Wang, M., & Chua, T. S. (2011). Aspect ranking: identifying important product aspects from online consumer reviews. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, Portland, Oregon

Zeithaml, V. A., Parasuraman, A., & Malhotra, A. (2002). 2002/09/01). Service quality delivery through web sites: a critical review of extant knowledge. *Journal of the Academy of Marketing Science*, *30*(4), 362. https://doi.org/10.1177/009207002236911

Zhang, Y., Weng, Q., & Zhu, N. (2018). The relationships between electronic banking adoption and its antecedents: a meta-analytic study of the role of national culture. *International Journal of Information Management*, *40*, 76–87. https://doi.org/10.1016/j.ijinfomgt.2018.01.015

Zhu, F., & Zhang, X. (2010). Impact of online consumer reviews on sales: the moderating role of product and consumer characteristics. *Journal of Marketing*, *74*(2), 133–148. https://doi.org/10.1509/jm.74.2.133

Zhu, J. J., Chang, Y. C., Ku, C. H., Li, S. Y., & Chen, C. J. (2021). Online critical review classification in response strategy and service provider rating: algorithms from heuristic processing, sentiment analysis to deep learning. *Journal of Business Research*, *129*, 860–877. https://doi.org/10.1016/j.jbusres.2020.11.007