



Choosing between hotels: impact of bimodal rating summary statistics and maximizing behavioral tendency

Ludovik Coba¹ · Laurens Rook² · Markus Zanker¹

Received: 28 June 2019 / Revised: 7 October 2019 / Accepted: 21 October 2019 /
Published online: 6 November 2019
© The Author(s) 2019

Abstract

Rating summary statistics are basic aggregations that reflect users' assessments of experienced products and services in numerical form. Thus far, scholars primarily investigated textual reviews, but dedicated considerably less time and effort exploring the potential impact of plain rating summary statistics on people's choice behavior. Notwithstanding their fundamental nature, however, rating summary statistics also are relevant to electronic commerce in general, and to e-tourism in particular. In this work, we attempted to fill this void, by exploring the effects of different types of rating attributes (the mean rating value, the overall number of ratings, and the bimodality of rating distributions) on hotel choice behavior. We also investigated whether individual differences in the cause of people's maximizing behavioral tendency moderated the effect of rating summary statistics on hotel choice behavior. Results of an eye-tracked conjoint experiment show that people's high or low on decision difficulty as the cause of maximization determined whether and how rating summary statistics have an impact on the choice between hotels. Implications for the tourism and hospitality domain are addressed.

Keywords e-Tourism · Rating summaries · Maximization · Conjoint analysis · Eye-tracking · Recommender systems · eWOM

✉ Markus Zanker
mzanker@unibz.it

Ludovik Coba
lucoba@unibz.it

Laurens Rook
l.rook@tudelft.nl

¹ Free University of Bozen-Bolzano, Piazza Domenicani 3, 39100 Bolzano, Italy

² Delft University of Technology, Jaffalaan 5, Delft, The Netherlands

1 Introduction

Online customer ratings and reviews were coined 'novel components of the marketing communication mix' about a decade ago (Chen and Xie 2008). Nowadays, due to the omnipresence of Web 2.0-based information technologies and applications in everyday life, user-produced and customer-shared evaluations of products and services have, indeed, become an essential ingredient of electronic commerce. Considerable evidence has been generated for their impact on sales and customer's purchase behavior for a wide range of products and services (De Maeyer 2012; Zhu and Zhang 2010). Customer ratings and reviews also are important factors in today's tourism and hospitality landscape—especially on the web platforms of online travel agencies (OTAs), cf. Gavilan et al. (2018). Competition on such web platforms is severe, due to the saturation of the market (Xiang et al. 2015). Moreover, tourism products are so-called 'experience goods' that are simultaneously delivered and experienced. This unique feature renders them highly variable in terms of experienced quality, and complicated to assess for the interested traveller prior to purchase (Chevalier et al. 2018). This has led to considerable research attention in the e-tourism community on the content, valence, and influence of online user reviews on hotel choice behavior (Gretzel and Yoo 2008; Xiang et al. 2017; Xie et al. 2011, 2014), while largely neglecting the impact of numerical user rating summary statistics on hotel selection.

The aim of the present study is to explore the impact of basic rating summary statistics on the hotels people choose from a choice set. Prior work in the marketing domain (i.e., in the setting of online booksellers) identified several metrics for the analysis of online product reviews, including basic product-related rating summary statistics like number of ratings and mean rating (Chevalier and Mayzlin 2006). For the movie domain, the volume and average movie ratings published online—together with a range of industry-specific factors—have been used to develop accurate sales forecasting models, cf. Dellarocas et al. (2007) and Liu (2006). In spite of this sparse evidence, basic rating summary statistics are usually included in empirical research as control variables only, cf. Chevalier et al. (2018). Hardly ever—if at all—do they take the center stage in research endeavors. However, the authors of a single modeling study once posited that simple ratings (such as number and average) could serve as a means for potential customers to reduce the complex task of choosing between product alternatives (Chen and Xie 2008). The very same idea was recently also put forward in an e-tourism study on the effect of ratings and written user reviews on hotel selection at booking platforms (Gavilan et al. 2018). This begs into question, if basic hotel rating summary statistics *as such* constitute any universal informative significance for online customers while making choices between hotels. Moreover, are we to assume that each specific rating summary statistic is of equal importance to the online customer during selection process? To the best of our knowledge, to date, hardly any research has dealt with these issues.

Interestingly, in his seminal work on bounded rationality, Herbert Simon specifically discussed the numerical mean as an example of a simple rule of thumb (or: heuristic) people could rely on in choice behavior. In fact, reliance on such heuristics

illustrated the principle of satisficing, in which someone settled for a reasonable choice alternative rather than maximized on the best possible option (Simon 1959). In recent years, scholars in the behavioral domain have begun to treat these principles of maximizing vs. satisficing in choice behavior as distinct personality traits. People high on so-called maximizing behavioral tendency display radically different choice behaviors than low scorers (satisficers), cf. Schwartz (2000, 2004) and Schwartz et al. (2002). Similarly, early marketing research recognized that maximizers in complicated choice settings often painstakingly scrutinize available choice options, whereas non-maximizers use faster and simpler routines (Payne 1976). In recent years, scholars have started to disentangle the fundamental, trait-based, causes of maximization from the more specific goals and strategies (Cheek and Schwartz 2016). Therefore, a person's sensitivity towards hotel rating summary statistics could very well be moderated by one's high or low trait-based cause of maximization—i.e., how simple or difficult the making of choices from a set of alternatives feels to the decision maker. It is important to better understand how the statistical characteristics of ratings and personality traits impact people's choice behavior, as this might have repercussions on the need to further develop and personalize present and future decision support systems in tourism and hospitality.

The remainder of the paper is organized as follows: Sect. 2 will provide an overview of classic and contemporary behavioral and marketing research into choice behavior. This section will discuss the general underpinnings and consequences of engaging in complicated choice settings. Hypotheses will be formulated for the impact of rating summary statistics and trait-based causes of maximization on user preferences for hotels. Like most studies on user preferences in choice behavior, a conjoint methodology will be used to put the hypotheses to the test. Section 3 outlines this methodology in greater detail. Given that the conjoint study was supplemented with an eye-tracking part, this section will also specify the procedures and metrics used in the eye-tracking sequence of the study. Section 4 will present the results of the behavioral data on choice behavior as well as of the implicit (i.e., eye-tracked) measurements. Finally, Sect. 5 will offer a detailed discussion of the scientific and practical implications of the findings, and—from the perspective of the recommender systems research domain—provide suggestions for future work on choice behavior in the e-tourism and hospitality domain.

2 Related work

How people make product and service-oriented consumer choices is an important topic in marketing research (Bettman et al. 1998; Payne 1976). Provision with a large variety of choice alternatives is traditionally considered beneficial to the customer—as this increases the likelihood that the desired product or service offering will eventually be selected. In recent years, however, scholars have begun to realize that too much choice might be overwhelming and harmful (Iyengar and Lepper 1999; Schwartz 2000). Exposure to extensive choice could even lead to choice overload, a situation in which someone experiences demotivation, dissatisfaction, or

feelings of regret (Iyengar and Lepper 2000); for reviews, see Chernev et al. (2015) and Scheibehenne et al. (2010).

The research on choice behavior is grounded in the seminal work of Herbert Simon from the 1950s (Simon 1955, 1956, 1959). Simon criticized the assumptions underlying the rational choice paradigm, which were based on an understanding of decision making as characterized by utility maximization. Maximization under exposure to a set of choices would lead to selection of the best—optimal—choice from a range of inferior alternatives. Maximizing choice, however, would require certainty regarding expected outcomes. Given the frequent occurrence of unexpected events and consequences (i.e., incomplete information and uncertainty) in everyday life, Simon regarded the rational choice paradigm an unrealistic oversimplification of human choice settings. Moreover, even in simple choice experiments, participants often did not behave as utility maximizers (Simon 1955). He, therefore, stated that “most real-life choices [...] lie beyond the reach of maximizing techniques—unless the situations are heroically simplified by drastic approximations” (Simon 1959), p 259.

Simon argued that human decision making in everyday life was more adequately described in terms of a satisficing principle (Simon 1955, 1956, 1959). Satisficing is a different form of choice behavior, in which people explore their options within a certain bandwidth, and in accordance with their individual drives and needs (or ‘aspiration levels’). Specifically, people search for—and eventually select—a choice that is regarded a satisfactory pick relative to a larger set of available alternatives. In other words, people tend to settle for a decent alternative that comes close to their aspiration levels—i.e., not the best possible item from a larger list. This is ‘approximate’ (Simon 1956) or ‘boundedly rational’ choice behavior “in a way that is procedurally reasonable in the light of the available knowledge and means of computation” (Simon 1986), p S211.

Schwartz et al. (2002) drew on the work of Herbert Simon on the satisficing principle in their development of a dispositional measure of choice behavior. The authors translated maximization and satisficing into two overarching but separate behavioral tendencies or personality traits, and developed a Maximization Scale to observe the differences. Consistent with Simon’s classic work, maximizers were defined as people with a general tendency to aim for the best choice, while satisficers were thought to be people who settle sooner for a “reasonable” alternative. Substantial evidence now exists that maximizers are often “doing better but feeling worse” than satisficers (Iyengar et al. 2006). They invest more time and energy in making a choice (Misuraca and Teuscher 2013), but commit less to those choices made (Sparks et al. 2012). Also, maximizers experience more demotivation, dissatisfaction and regret than satisficers, while being less happy with their choices (Dar-Nimrod et al. 2009; Schwartz et al. 2002); for review, see Cheek and Schwartz (2016) and Schwartz (2004).

Recommender systems are automated decision support tools that enable people to overcome information overload in complicated choice settings (Jannach et al. 2011; Ricci et al. 2015). From this angle, scholars have begun to explore the impact of system users’ maximizing behavioral tendencies on recommender systems—with mixed results. Several studies failed to report significant differences between

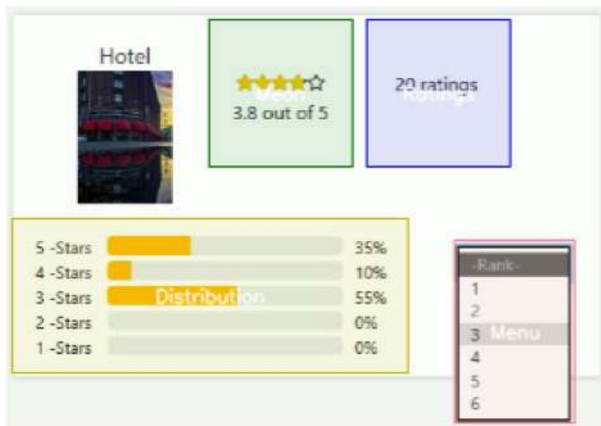


Fig. 1 Profile snapshot

maximizers and satisficers once exposed to recommendations (Jugovac et al. 2018; Knijnenburg et al. 2011). Knijnenburg et al. (2011) argued that this might have had something to do with the difference between choice process and choice outcome. This observation resonates with the growing insight that individual differences in overall maximizing behavioral tendency are better understood in terms of separate goals, strategies, and causes (Nenkov et al. 2008; Schwartz et al. 2016). Focusing on the latter dimension (i.e., the trait-based causes underlying maximization), recent studies by Coba et al. (2019) showed that maximizers respond differently to collaborative explanations of choice alternatives than satisficers, and that this effect especially holds for the individual difference items that tap into causes of maximization (labelled 'decision difficulty' in the literature, cf. Cheek and Schwartz 2016; Nenkov et al. 2008; Schwartz et al. 2002).

3 Methodology

Conjoint analysis was developed in the marketing domain with the purpose to identify the most important attributes of a product in relation with a person's preferences and (intended and/or actual) buying behavior (Zwerina et al. 1996). From a formal point of view, conjoint designs always comprise the following building blocks: items (also known as *profiles*, see example in Fig. 1), which are composed of sets of categorical or quantitative *attributes*, which are further refined into distinct (usually: high and low) *levels*, cf. Rao (2008). In the present study, the rank-based conjoint method was selected. In this specific version of the conjoint methodology, people are asked to rank items within a given choice set in terms of their (high to low) appreciation for those items. This also is common practice on web platforms of OTAs in tourism, on which users are often given the opportunity to rank lists of services according to their preferences Chung and Rao (2012).



Fig. 2 Items drawn from the TripAdvisor dataset (bimodality coefficients > 0.7)

3.1 Attribute selection and experimental design

Rating summary statistics are often depicted as frequency distributions on the class of discrete ratings values (such as one to five stars). This particular characteristic was used to select ecologically valid attributes, and to develop the stimuli (i.e., the profiles/items) accordingly, as will be discussed in greater detail below.

Attribute selection In prior work (Coba et al. 2018a, b), the mean rating value was found to be the strongest predictor of choice behavior, while the number of ratings had only a weak influence. This was primarily the case when the number of ratings was considered relatively high (i.e. in the three digits and above). When the number of ratings were in the double digits, participants put more weight on them during decision making. People, apparently, prefer a slightly lower mean rating value over a higher number of overall ratings—making the mean rating value appear more reliable. Marginal and null effects were observed, when additional characteristics of rating distributions, such as the variance or skewness were also taken into consideration (Coba et al. 2019). However, rating distributions might actually exhibit an asymmetric bimodal (J-shaped) distribution (Hu et al. 2009). This J-shaped distribution represents a *purchasing bias* (i.e., one tends to buy what one likes) as well as an *under-reporting bias* (i.e., polarized opinions are more likely to be reported). This potentially renders the mean rating value a biased measure for product quality (de Langhe et al. 2016). Therefore, it was hypothesized that even though an item might have a high overall score, an additional “minor” peak on low rating values would actually discourage users to choose such an item. As recommended by Pfister et al. (2013), the range of these “peaks” was measured using the bimodality coefficient, which was computed as:

$$BC = \frac{m_3^2 + 1}{m_4 + 3 \frac{(n-1)^2}{(n-2)(n-3)}} \quad (1)$$

where m_3 is skewness, m_4 kurtosis and n the sample size of the distribution. The bimodality coefficient varies from 0 to 1, in which a low value indicates an unimodal bell-shaped distribution. The value of 0.55 is considered a threshold, where a bimodal distribution is recognized as such. Values above this threshold clearly exhibit a bimodal distribution (see examples in Fig. 2).

In order to develop ecologically valid levels for the three attributes in the experiment (i.e., number of ratings, mean rating, bimodality), they were aligned with real

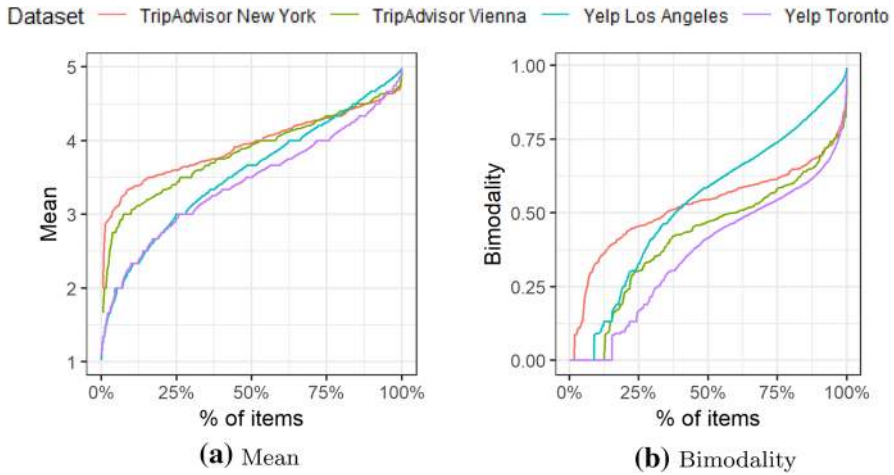


Fig. 3 Statistical descriptives of datasets

Table 1 Characteristics of datasets

Dataset	# Ratings	# Users	# Items	Density (%)	Source/reference
TripAdvisor					
New York	11,061	9597	258	0.45	Fuchs and Zanker (2012)
Paris	10,725	9500	1008	0.11	
Yelp					
Las Vegas	1,605,396	506,278	26,809	< 0.01	Yelp.com ^a
Toronto	431,047	93,117	17,211	< 0.01	

industry data from the tourism and hospitality domain crawled from TripAdvisor (Fuchs and Zanker 2012; Jannach et al. 2014), and a public data set from Yelp¹ (for a general overview of these data sets, see Table 1).

Specifically, examination of the mean rating values (Fig. 3a) in these data sets revealed that the ratings in the industry were skewed towards higher values. Likewise, the bimodality coefficient (Fig. 3b) was present in all the data sets. This confirmed that J-shaped rating distributions do indeed occur in tourism data. The number of ratings was set at 20 and 80, since prior work reported that participants clearly notice the difference between these levels (Coba et al. 2018a). Mean rating values between 3.6 and 4.0 were found for many rated items in our real-life Tripadvisor data sets (see Fig. 3a). The bimodality coefficient was set at 0.3 (no noticeable second peak present) and 0.7 (clear unanimity of reviewers). Table 2 shows the high and low levels derived from the industry data, and how these ecologically valid

¹ <https://www.yelp.com/dataset/challenge>.

Table 2 Attributes and levels

Attribute	Level	Value
A1: number of ratings	L1	20
	L2	80
A2: mean rating	L1	3.6
	L2	3.8
	L3	4
A3: bimodality	L1	0.3
	L2	0.5
	L3	0.7

levels enabled us to put the three attributes of rating summary statistics of hotels in our study under experimental control.

Experimental design A full-factorial design (Zwerina et al. 1996) was built on the following three attributes: number of ratings, mean rating, and bimodality. Specifically, the full factorial contained 3 attributes (2 levels \times 3 levels \times 3 levels), which resulted in 18 different profiles that were put to test. Importantly, all items represented statistically feasible level combinations. Note that the profiles were blocked into three subsets in order to lower the cognitive load for respondents to feasible levels. That is to say, they had to rank 3×6 alternatives. This simulated complicated choice as on existing OTA web platforms (Werthner et al. 2015), but prevented the occurrence of choice overload, which could cause the choice process to come to a stand still (Iyengar and Lepper 2000).

3.2 Procedure

Participants took part in a controlled experiment on a terminal, which had been configured in advance to host an eye-tracking study. Specifically, the stimuli were presented on a 22" display, and gazes were recorded with a static remote eye-tracking system equipped with a 150 Hz research-grade machine-vision camera. On arrival in the laboratory, participants were briefly introduced to the purpose of the study, and asked to give informed consent to have their data used for research purposes. If they did, participants were next provided with a pre-survey, which contained a test measure on maximizing behavioral tendency (described in detail below). The eye-tracking experiment was started from a remote console, once the participant had filled out this short personality test. They were asked to work on a series of ranking exercises framed as a representative e-tourism problem as follows:

“You need to rank hotels on a booking platform for your holiday stay. All hotels are equally preferable for you with respect to cost, location, facilities, services, etc. Other users’ ratings of this hotel are aggregated and summarized by their number of ratings, the mean of their ratings, and their distribution over the different rating values. Given the above, which of the hotels below would

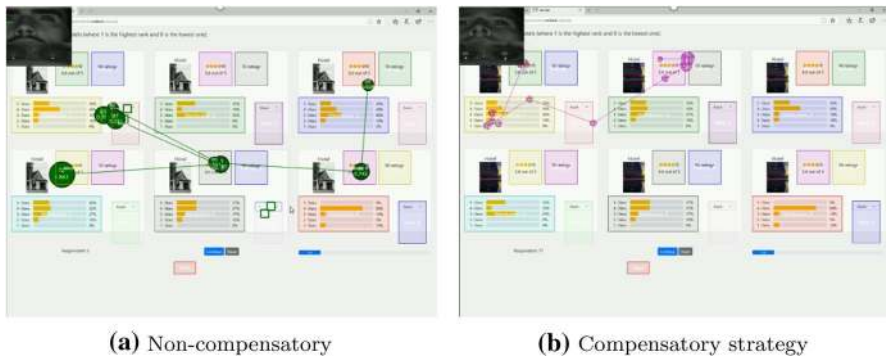


Fig. 4 Screenshot highlighting the scan paths on the areas of interest

you prefer, when you were to solely consider the ratings for the displayed accommodations?”

Following this introduction, the participant went through a sequence of 3 ranking exercises. In each of these exercises, a participant had to rank 6 hotels based on specific summary statistics. These hotels were taken from the full factorial set of combinations of attribute levels (consisting of 18 distinct hotel summary statistics described above). Attributes levels were equally distributed and balanced between the exercises, so as to allow for proper test of main effects. Presented profiles were item-agnostic, which meant that a participant could rely on nothing but the three variables put under experimental control (i.e., number of ratings, mean, and bimodality) during ranking the profiles. Note that the order in which the six hotel profiles were presented on the 22" screen was perfectly randomized. In other words, each participant received and ranked these six hotels in a unique configuration. A screenshot of the ranking exercise is presented in Fig. 4.²

After finishing, participants were asked to indicate in a brief post-survey, which characteristics of rating summaries in their opinion had guided their choice behavior most. Also, they were asked to provide demographic information, after which they were debriefed, thanked, and sent away.

3.3 Measures

Eye-tracking: areas of interests Areas of Interest (AOIs) (Holmqvist et al. 2011) are regions defined in the stimulus so as to extract data specifically for those areas. The three AOIs per item identified in the present study were associated with the three attributes in the conjoint design (see Fig. 1). The dwell or *gaze* refers to a focal visit of an AOI, from entry to exit, while a gaze cluster constitutes a *fixation*. A hit on

² The implementation of the platform is downloadable at: <https://github.com/ludovikcoba/rankBasedConjoint.git>.

an AOI is typically operationalized as a lock of the gaze into a specific area for as long as minimally required to cognitively process the information within that AOI (Holmqvist et al. 2011). A transition is the movement of the gaze from one AOI to another, while a *revisit* is the transition back to an AOI already visited.

Eye-tracking: fixation time Fixation time was measured so as to assess the amount of time a participant spent on arriving at a choice. Because of the modest sample size in the present study (see also Sect. 4), the geometrical mean and the log-transformation of the confidence interval were used in order to best estimate the average time spent on the experimental task (Sauro and Lewis 2010). In addition, revisits were used as a proxy for the manner in which a participant examined alternatives, searched for what amount of information, and for the comparisons made during the task (Payne 1976).

Maximization: decision difficulty Several scales exist to assess a person's maximizing behavioral tendency; see Cheek and Schwartz (2016) for a review. The unique selling point of the shortened Maximization Scale of Nenkov et al. (2008) lies in its distinction between three sub-dimensions: having high standards, alternative search, and decision difficulty. As we were interested in the causes of maximization (Cheek and Schwartz 2016), in the present study, we looked into the decision difficulty sub-scale. It comprises the following items: "I often find it difficult to shop for a gift for a friend", and "Booking a hotel is really difficult. I'm always struggling to pick the best one".³ Each item was measured on a 7-point scale ranging from 1 (*completely disagree*) to 7 (*completely agree*). The decision difficulty sub-scale was reliable with a Cronbach's α within the range outlined by Nenkov et al. (2008).

4 Results

A total of 42 volunteers participated in our conjoint experiment, and made a series of rank-based hotel choices, while being eye-tracked. The general demographic characteristics of this sample are summarized in Table 3. Our first prediction was that high (vs. low) hotel rating summary statistics (i.e., number of ratings, mean rating, and bimodality) positively impact someone's rating and choice behavior. Second, we predicted that decision difficulty (due to behavioral maximizing tendency) moderates someone's rating and choice behavior. Therefore, the sample was divided (via median split) into sub-samples based on a participant's high or low score on decision difficulty. This enabled us to probe the differences between and within choice behaviors for each of these groups.

Table 4 outlines the parameter estimates of our multinomial test. Analysis of the behavioral data output revealed that the choice behavior of people high on decision difficulty was especially contingent on a higher mean ($\beta = 1.35, p < 0.001$), while they were less inclined to consider a hotel with a higher number of ratings ($\beta = 0.76, p < 0.001$). The choice behavior of people

³ Note, that the original, outdated, phrase *Renting a video[...]* in the scale of Nenkov et al. (2008) in the present research was replaced with *Booking a hotel[...]*.

Table 3 Participants' demographic details

Personal feature	Category				Total
Gender	Male	Female	No answer		
#	24	18	0		42
%	57.1	42.9	0		100
Age	18–24	25–30	31–40	40 +	
#	30	7	4	1	42
%	71.4	16.7	9.5	2.4	100
Country	Italy	Albania	Germany	Other	
#	16	8	8	10	42
%	38.1	21.4	19.0	23.8	100

Table 4 Parameter estimates for all respondents, and grouped by the median split on the decision difficulty sub-scale

Attribute	Level	Estimate (β)	
		Below median	Above median
# Ratings	80	0.90 (0.09)***	0.76 (0.11)***
	20	–	–
Mean	4	1.11 (0.12)***	1.35 (0.14)***
	3.8	–0.04 (0.09)	0.05 (0.14)
	3.6	–	–
Bim.	0.7	–0.08 (0.12)	0.26 (0.14)
	0.5	–0.02 (0.12)	0.07 (0.14)
	0.3	–	–

Dashes (–) are the baseline levels. The estimated coefficients are the change in odds of choosing a particular mode rather than the baseline category. The values in parentheses are estimated standard errors

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

low on decision difficulty, on the other hand, was almost equally influenced by presence of a high mean or a high number of ratings in a hotel summary. Apparently, people low on decision difficulty could more confidently trade-in the different attribute levels of these two rating summary statistics against each other in their choice behavior. These results were in line with predictions. However, no effects were observed for the bimodality of distributions as a function of a person's high or low decision difficulty. That is, even for hotel rating distributions with a stronger J-shape, people would always rely on hotels with a higher mean value, also in the presence of more low hotel rating values. This unexpected behavioral effect was observed for participants independent of their maximizing behavioral tendency (decision difficulty).

Early research on consumer choice behavior suggested that maximizers are more prone to comparing (many) alternatives prior to making (and committing themselves to) a choice. That is, maximizers tend to adopt a compensatory (rather than non-compensatory) information processing strategy, which is characterized by (a) a thorough and (b) more time-consuming process of making comparisons between multiple item attributes in a choice set (Payne 1976). So as to explore the presence of this particular—compensatory—behavior in our data, we first looked into the average number of revisits. The average number of revisits is a proxy for the frequency, with which a person switches the focus of attention from an AOI of one item to the AOI of another item, and vice versa. Figure 5 shows that people high on decision difficulty made more revisits to certain hotel offerings than those low on decision difficulty. This offered first evidence that high (vs. low) maximizers in our sample had displayed a more thorough, compensatory, choice behavior.

Second, we examined how much time people high (vs. low) on maximizing behavioral tendency (decision difficulty) had actually spent on each hotel offering in their choice set. Figure 6 presents the geometrical mean of the fixation times on the hotels in the choice sets based on the ranking positions they eventually

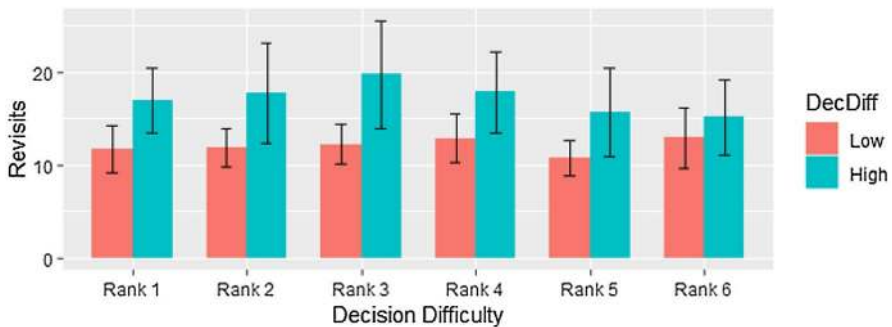


Fig. 5 Mean number of revisits per item, median split on the decision difficulty sub-scale

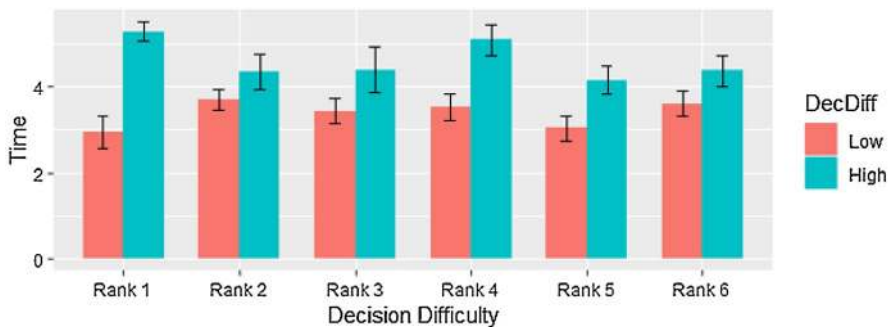


Fig. 6 Geometrical mean of the time spent on item (confidence level of 95%), median split on decision difficulty sub-scale

received. These fixation times were computed only for the first ranking task (so as to capture the learning effect), and grouped according to the participant’s high or low decision difficulty. In line with predictions, people high on decision difficulty invested more time into studying choice alternatives than those low on decision difficulty. We used the Gini coefficient, which is a classic statistical measure of inequality from the domain of economics (Dorfman 1979) to test for inequality between high and low maximizers on their fixation frequencies and time spent on choice behavior. The Gini coefficient ranges from 0 to 1, and a higher value stands for a higher degree of inequality. Figure 7 shows that the degree of inequality (in fixations and time spent) is much lower for people high than low on decision difficulty—i.e., the eye-tracking activities of maximizers were much more focused than those of satisficers; for significance test, see Table 5. This offered further evidence for the early observation of Payne (1976) that maximizers apply (focused) compensatory information processing strategies in their consumer choice behavior.

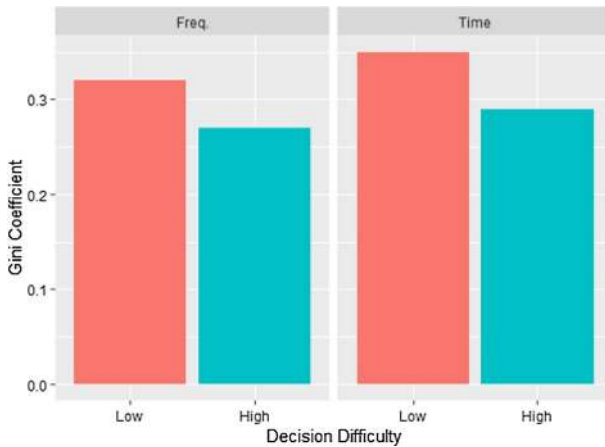


Fig. 7 Gini

Table 5 Gini coefficient for the time and the frequencies per median split on the scale

Scale	Gini coefficient	
	Time	Freq.
Maxim.		
Low	0.33	0.30
High	0.32	0.29
Dec. diff.*		
Low	0.35	0.32
High	0.29	0.27

* $p < 0.05$

Summarizing, analysis of the behavioral hotel choices as well as of the implicit (i.e., eye-tracking) measurements offered support for our prediction that high (but not low) hotel rating summary statistics influence ranking and choice behavior. Second, with the exception of summary statistics for bimodality, maximizing behavioral tendency produced different rating and choice behavior under exposure to hotel rating summary statistics.

5 Discussion and conclusion

The present study investigated how rating summary statistics influence people's choice behavior of hotels. Three commonly found online numerical summaries of customer preferences—total number of ratings, mean rating, and bimodality of the distribution—were put under full experimental control, and tested for differences in interaction with a person's self-reported maximizing behavioral tendency. The impact of hotel rating summary statistics on choice behavior clearly depended on the decision difficulty a person high (vs. low) on maximization experienced. Implicit measurement of dynamic search processes via eye-tracking confirmed that such people spent more time weighting the pros and cons of choice alternatives.

In general terms, the scientific contribution of this work lies in the introduction of an established line of behavioral research into maximization, set in motion by Schwartz (2000, 2004) and Schwartz et al. (2002), but inspired by the seminal work of Nobel laureate Herbert Simon on satisficing (Simon 1955, 1956, 1959), to the tourism community. Thus far, the role of behavioral maximizing tendency on complex choice behavior had been largely overlooked in research on tourism and hospitality. The present study shows that people high on maximization experience difficulty while making choices between hotels, which is a novel insight for the research community that might invite further research into maximization and satisficing in tourism and hospitality.

More specifically, the present work fits into a longer academic tradition of studying consumer choice processes of any complexity with the help of eye-tracking equipment (Bettman et al. 1998; Glaholt and Reingold 1985; Orquin and Mueller Loose 2013; Payne 1976; Reutskaja et al. 2011). For our study, this offered a rich analysis of a person's gazes, revisits, and fixations on a set of hotel choices, and brought to light that it took people high on decision difficulty longer to settle on a choice. Still, they primarily relied on high mean rating values. In comparison, people low on maximization moved more freely between the various choice alternatives, and needed less time to choose. This is a fascinating finding, given that reliance on the arithmetic mean has always been regarded a simple rule of thumb in games of choice, cf. Simon (1959). Our work offers a caveat to such an all too straightforward interpretation. Implicit measurement of the underlying search and decision dynamics indicated that maximizers actually had invested more time and effort in the choice process, and had continued comparing alternatives before they—finally—settled on the highest mean value. This outcome is consistent with recent theorizing on maximization (Cheek and Schwartz 2016), and illustrates the viability of eye-tracking for theory development on choice behavior, and beyond.

Choice complexity is a major problem in tourism and hospitality. Most businesses within the industry tend to offer large varieties of (online and offline) products and services, and allow for high levels of customization of already complicated (multiple events and/or multidestinatons) product-service bundles (Park and Jang 2013; Thai and Yuksel 2017). Sometimes, the customer's choice behavior in such complicated settings is still understood in terms of rational utility maximization (de Oliveira Santos et al. 2011). Our work—in general terms—shows the feasibility for the industry of taking a 'boundedly rational' stance to maximization as determined by individual differences, instead. Clearly, maximizers respond differently to tourism-related product and service offerings than satisficers, and it might pay-off for the service provider to take away some of the maximizer's underlying decision difficulty in the various stages of the customer-seller relationship. This is a practical contribution of our work, which applies to traditional players within the industry as well as to e-businesses with a strong online presence.

The most important practical implication of the present work has to do with e-businesses in tourism and hospitality. These days, summary statistics of each and every product and service are available on the platforms of online travel agents. However minor these numerical values, they summarize the experiences of prior customers, and are very informative to the next traveller. Together with more unstructured written customer reviews, ratings offer valuable pointers to the intended traveller as to whether to visit a hotel facility or not. Considerable research within tourism and hospitality shows that such online information might trigger (positive and negative) word-of-mouth with financial consequences for the e-business; cf. Xie et al. (2014). The advantage of rating summary statistics over customer reviews, however, is that they do not require tedious text mining protocols (Gavilan et al. 2018; Xu and Li 2016; Zhao et al. 2019)—with the inherent danger to put a wrong interpretation on the online customer review (Antioco and Coussement 2018). What is more, they can even be implemented in ranking algorithms in rather straightforward fashion. The present study was the first to identify, which rating attributes might have the strongest impact on tourism-related choice behavior, and in what way. The practical relevance for the industry lies in using these insights for the development of personalized prediction models (see also our discussion on future research below).

The limitations of this work are the following: first, the results of the present study derived from a sample of 42 participants. The recommended sample sizes for conjoint studies in marketing research tend to be much larger than this—especially, when customer segmentation is the main research objective. Nonetheless, a smaller sample size (i.e., fluctuating between thirty and sixty participants) is not regarded problematic for exploratory conjoint experiments, and research aimed at hypothesis testing (Orme 2010). Second, and in addition to the previous observation, a sample size of the magnitude used in the present research is more than reasonable for an eye-tracking study conducted in the laboratory, cf. Reutskaja et al. (2011). In the specific case of choice behavior, eye-tracking studies tend to be designed to measure the dynamic search processes prior to decision making—rendering the many metrics underlying those choice processes (i.e., the gazes, fixations, and revisits to AOIs) the actual focal points of the analysis; see Reutskaja et al. (2011) for a similar observation. Third, participants voluntarily took part in

the study, and might (or might not) have been conveniently available to the experimenter. It should be emphasized, however, that our eye-tracking results correspond to the pattern of results reported in a previous choice-based conjoint experiment, derived from a larger randomized sample of 200 respondents with diverse international backgrounds and demographic characteristics (Coba et al. 2019). This, in our opinion, offers support to the notion that the eye-tracked results for maximizers and satisficers in this study are not, in fact, confounded by convenience or quota sampling. Fourth, sample sizes such as in the present study are incomparable with the larger samples collected with the help of commercial eye-tracking companies using real products on site (Chandon et al. 2009). For sure, eye-tracking studies with real consumers contemplating their next tourism-related product or service choice on one of the larger e-tourism booking platforms would be richer in external validity, and yield additional insights.

This brings us to ways, in which our work might inspire future research on the topic of tourism-related choice behavior (Zanker et al. 2019). First, and foremost, our current research endeavors are focused on the further improvement of choice and ranking-based recommendation algorithms in the setting of e-tourism—as the development of such algorithms has clear practical relevance for the industry (see also above). Specifically, we are working on adjustment of such algorithms based on a user's overall maximizing behavioral tendency. We hope that, in the near future, this will lead us to develop a practical IT tool to offer personalized recommendations of matching product and service offerings to maximizers and satisficers in the tourism domain. As also mentioned above, it is important to overcome—and, ideally, to avoid the occurrence of—structural dissatisfaction of the maximizing customer with a certain tourism product or service offering. Algorithms that are personalized on the customer's disposition for maximization allow for the targeted offering of practical support in this case, and might complement existing IT-driven customer recovery and retention programs.

Second, the insights from our conjoint experiment could be used to more explicitly test in future work the notion of choice overload in e-tourism. Due to the rise of online travel agents, customers now have unprecedented access to enormous amounts of complicated (i.e., bundled) travel products (Tanford et al. 2012). Even though richness of choice might be good from a corporate perspective, the online offerings in hospitality and tourism clearly lead to situations, in which someone interested in travel might be offered too much choice. Scholars have begun to explore the phenomenon of choice overload and its boundary conditions in such settings, and discovered that exposure to 22 or more choices might cause a potential traveller to refrain from making a decision (Park and Jang 2013). It is an intriguing question, whether this number of 22 choices would serve as a universal critical threshold for e-tourism. The groundbreaking work of Iyengar and colleagues in the behavioral and marketing domain indicates that choice overload tends to occur between the range of 24 and 30 or more choices (Iyengar and Lepper 1999, 2000). In the light of our results, it even makes sense to posit that such a critical threshold probably varies for people depending on their high or low behavioral tendency to maximize (but see Iyengar and Lepper (2000), Study 3 for an alternative position). This, for sure, is an interesting issue to explore in future work on the topic. It would

be particularly interesting to do so with the help of the eye-tracking procedure that was developed for the present study.

Third, the studies referred to above typically think about choice behavior in tourism in connection with notions of dissatisfaction, anger and perceived regret, cf. Park and Jang (2013, 2018). This is in line with a larger investigation in research on e-tourism of the antecedents of online customer switching behavior. Prior findings for instance show that hotel visitors more likely switch to other service providers when they experienced anger and regret after service failure. Disappointing treatment of complaints increases the likelihood that people engage in negative WOM regarding the hotel (Sánchez-García and Currás-Pérez 2011), which might ruin a company's online reputation; see also He and Harris (2014). Theoretically, the constructs of regret and maximization are significantly correlated with each other (Cheek and Schwartz 2016; Dar-Nimrod et al. 2009). In fact, Schwartz et al. developed their Maximization Scale in parallel with a Regret Scale, and postulated that "concern about potential regret [...] influences some people to *be* maximizers" (Schwartz et al. 2002), p. 1179. It would be worthwhile to include this insight into future research on online switching behavior after service failures in tourism and hospitality, and study maximization and regret in tandem.

In conclusion, the present study explored the impact of online hotel rating summary statistics on choice behavior, and found that they depend on a person's high (vs. low) behavioral maximizing tendency. People high on maximization found it more difficult to make a choice, and spent much more time exploring alternatives. Theoretically, this finding might be linked with tourism research on the (dis)satisfaction, anger, regret, and switching behavior of travellers. It shows how important it is to account for the individual differences of potential customers that browse, search, and evaluate various alternative offers on online hotel booking platforms prior to making a purchase.

Acknowledgements The authors would like to acknowledge Gabriela Boyadjijyska for supporting the eye-tracking experimentation as part of her thesis project.

Funding This work was supported by the Open Access Publishing Fund provided by the Free University of Bozen-Bolzano.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Antiocho M, Coussement K (2018) Misreading of consumer dissatisfaction in online product reviews: writing style as a cause for bias. *Int J Inf Manag* 38(1):301–310
- Bettman JR, Luce MF, Payne JW (1998) Constructive consumer choice processes. *Sour J Consum Res* 25(3):187–217. <https://doi.org/10.1086/209535>

- Chandon P, Hutchinson JW, Bradlow ET, Young SH (2009) Does in-store marketing work? effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase. *J Mark* 73(6):1–17
- Cheek NN, Schwartz B (2016) On the meaning and measurement of maximization. *Judgm Decis Mak* 11(2):126
- Chen Y, Xie J (2008) Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Manag Sci* 54(3):477–491
- Chernev A, Böckenholt U, Goodman J (2015) Choice overload: a conceptual review and meta-analysis. *J Consum Psychol* 25(2):333–358
- Chevalier JA, Mayzlin D (2006) The effect of word of mouth on sales: online book reviews. *J Mark Res* 43(3):345–354
- Chevalier JA, Dover Y, Mayzlin D (2018) Channels of impact: User reviews when quality is dynamic and managers respond. *Mark Sci* 37(5):688–709
- Chung J, Rao VR (2012) A general consumer preference model for experience products: application to internet recommendation services. *J Mark Res* 49(3):289–305
- Coba L, Zanker M, Rook L, Symeonidis P (2018a) Exploring users' perception of collaborative explanation styles. In: 2018 IEEE 20th conference on business informatics (CBI), IEEE, vol 1, pp 70–78. <https://doi.org/10.1109/CBI.2018.00017>. <http://arxiv.org/abs/1805.00977>
- Coba L, Zanker M, Rook L, Symeonidis P (2018b) Exploring users' perception of rating summary statistics. In: Proceedings of the 26th conference on user modeling, adaptation and personalization, UMAP '18. ACM, New York, pp 353–354. <https://doi.org/10.1145/3209219.3209256>
- Coba L, Rook L, Zanker M, Symeonidis P (2019) Decision making strategies differ in the presence of collaborative explanations: two conjoint studies. In Proceedings of the 24th international conference on intelligent user interfaces. ACM, pp 291–302
- Dar-Nimrod I, Rawn CD, Lehman DR, Schwartz B (2009) The maximization paradox: the costs of seeking alternatives. *Personal Individ Differ* 46(5–6):631–635
- de Langhe B, Fernbach PM, Lichtenstein DR (2016) Navigating by the stars: investigating the actual and perceived validity of online user ratings. *J Consum Res* 42(6):817–833. <https://doi.org/10.1093/jcr/ucv047>
- De Maeyer P (2012) Impact of online consumer reviews on sales and price strategies: a review and directions for future research. *J Prod Brand Manag* 21(2):132–139
- de Oliveira Santos GE, Ramos V, Rey-Maqueira J (2011) A microeconomic model of multidestination tourism trips. *Tour Econ* 17(3):509–529
- Dellarocas C, Zhang XM, Awad NF (2007) Exploring the value of online product reviews in forecasting sales: the case of motion pictures. *J Interact Mark* 21(4):23–45
- Dorfman R (1979) A formula for the Gini coefficient. *Rev Econ Stat* 61(1):146–149
- Fuchs M, Zanker M (2012) Multi-criteria ratings for recommender systems: an empirical analysis in the tourism domain. In: International conference on electronic commerce and web technologies. Springer, Berlin, Heidelberg, pp 100–111
- Gavilan D, Avello M, Martinez-Navarro G (2018) The influence of online ratings and reviews on hotel booking consideration. *Tour Manag* 66:53–61
- Glaholt MG, Reingold EM (1985) Eye movement monitoring as a process tracing methodology in decision making research 4(2):125–146. <https://doi.org/10.1037/a0020692>. <http://www.psych.utoronto.ca/users/reingold/publications/PDFs/2011.Glaholt.Reingold.EyeMovementMonitoringasaProcessTracingMethodology.pdf>
- Gretzel U, Yoo KH (2008) Use and impact of online travel reviews. In: O'Connor P, Höpken W, Gretzel U (eds) Information and communication technologies in tourism 2008. Springer, Vienna, pp 35–46
- He H, Harris L (2014) Moral disengagement of hotel guest negative wom: moral identity centrality, moral awareness, and anger. *Ann Tour Res* 45:132–151
- Holmqvist K, Nyström M, Andersson R, Dewhurst R, Jarodzka H, Van De Weijer J (2011) Eye tracking. A comprehensive guide to methods and measures. Oxford University Press. <https://global.oup.com/academic/product/eye-tracking-9780199697083?cc=us&lang=en&>
- Hu N, Zhang J, Pavlou PA (2009) Overcoming the J-shaped distribution of product reviews. *Commun ACM* 52(10):144. <https://doi.org/10.1145/1562764.1562800>
- Iyengar SS, Lepper MR (1999) Rethinking the value of choice: a cultural perspective on intrinsic motivation. *J Personal Soc Psychol* 76(3):349
- Iyengar SS, Lepper MR (2000) When choice is demotivating: can one desire too much of a good thing? *J Personal Soc Psychol* 79(6):995

- Iyengar SS, Wells RE, Schwartz B (2006) Doing better but feeling worse: looking for the “best” job undermines satisfaction. *Psychol Sci* 17(2):143–150
- Jannach D, Zanker M, Fuchs M (2014) Leveraging multi-criteria customer feedback for satisfaction analysis and improved recommendations. *Inf Technol Tour* 14(2):119–149
- Jannach D, Zanker M, Felfernig A, Friedrich G (2011) Recommender systems: an introduction. Cambridge University Press, Cambridge
- Jugovac M, Nunes I, Jannach D (2018) Investigating the decision-making behavior of maximizers and satisficers in the presence of recommendations. In: Proceedings of the 26th conference on user modeling, adaptation and personalization. ACM, pp 279–283
- Knijnenburg BP, Reijmer NJ, Willemsen MC (2011) Each to his own: how different users call for different interaction methods in recommender systems. In: Proceedings of the 5th ACM conference on Recommender systems-RecSys '11, ACM, pp 141–148. <https://doi.org/10.1145/2043932.2043960>
- Liu Y (2006) Word of mouth for movies: its dynamics and impact on box office revenue. *J Mark* 70(3):74–89
- Misuraca R, Teuscher U (2013) Time flies when you maximize—maximizers and satisficers perceive time differently when making decisions. *Acta Psychol* 143(2):176–180
- Nenkov GY, Morrin M, Ward A, Schwartz B, Hulland J (2008) A short form of the maximization scale: factor structure, reliability and validity studies. *Judgm Decis Mak* 3(5):371–388
- Orme B (2010) Getting started with conjoint analysis: strategies for product design and pricing research, 4th edn. Research Publishers LLC, Madison, Wis
- Orquin JL, Mueller Loose S (2013) Attention and choice: a review on eye movements in decision making. *Acta Psychol* 144(1):190–206. <https://doi.org/10.1016/J.ACTPSY.2013.06.003>
- Park JY, Jang SS (2013) Confused by too many choices? Choice overload in tourism. *Tour Manag* 35:1–12
- Park JY, Jang S (2018) Did i get the best discount? Counterfactual thinking of tourism products. *J Travel Res* 57(1):17–30
- Payne JW (1976) Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organ Behav Hum Perform* 16(2):366–387. [https://doi.org/10.1016/0030-5073\(76\)90022-2](https://doi.org/10.1016/0030-5073(76)90022-2). <https://pdfs.semanticscholar.org/460e/9406c71561db85a5052f8068a4b0a4a84d39.pdf>
- Pfister R, Schwarz KA, Janczyk M, Dale R, Freeman JB (2013) Good things peak in pairs: a note on the bimodality coefficient. *Front Psychol* 4:700. <https://doi.org/10.3389/fpsyg.2013.00700>. <http://www.ncbi.nlm.nih.gov/pubmed/24109465>www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3791391
- Rao VR (2008) Developments in conjoint analysis. In: Wierenga B (eds) Handbook of marketing decision models. Springer, Boston, MA, pp 23–53
- Reutskaja E, Nagel R, Camerer CF, Rangel A (2011) Search dynamics in consumer choice under time pressure: an eye-tracking study. *Am Econ Rev* 101(2):900–926
- Ricci F, Rokach L, Shapira B (2015) Recommender systems: introduction and challenges. In: Ricci F, Rokach L, Shapira B (eds) Recommender systems handbook. Springer, Boston, MA, pp 1–34
- Sánchez-García I, Currás-Pérez R (2011) Effects of dissatisfaction in tourist services: the role of anger and regret. *Tour Manag* 32(6):1397–1406
- Sauro J, Lewis JR (2010) Average task times in usability tests: what to report? In Proceedings of the SIGCHI conference on human factors in computing systems, pp 2347–2350. ACM. <https://doi.org/10.1145/1753326.1753679>
- Scheibehenne B, Greifeneder R, Todd PM (2010) Can there ever be too many options? A meta-analytic review of choice overload. *J Consum Res* 37(3):409–425
- Schwartz B (2000) Self-determination: the tyranny of freedom. *Am Psychol* 55(1):79
- Schwartz B (2004) The paradox of choice: why more is less. Ecco, New York
- Schwartz B et al (2016) On the meaning and measurement of maximization. *Judgm Decis Mak* 11(2):126
- Schwartz B, Ward A, Monterosso J, Lyubomirsky S, White K, Lehman DR (2002) Maximizing versus satisficing: happiness is a matter of choice. *J Pers Soc Psychol* 83(5):1178–1197. <https://doi.org/10.1037//0022-3514.83.5.1178>
- Simon HA (1986) Rationality in psychology and economics. *J Bus* S209–S224
- Simon HA (1955) A behavioral model of rational choice. *Q J Econ* 69(1):99–118
- Simon HA (1956) Rational choice and the structure of the environment. *Psychol Rev* 63(2):129
- Simon HA (1959) Theories of decision-making in economics and behavioral science. *Am Econ Rev* 49:253–283

- Sparks EA, Ehrlinger J, Eibach RP (2012) Failing to commit: maximizers avoid commitment in a way that contributes to reduced satisfaction. *Personal Individ Differ* 52(1):72–77
- Tanford S, Baloglu S, Erdem M (2012) Travel packaging on the internet: the impact of pricing information and perceived value on consumer choice. *J Travel Res* 51(1):68–80
- Thai NT, Yuksel U (2017) Choice overload in holiday destination choices. *Int J Cult Tour Hosp Res* 11(1):53–66
- Werthner H, Alzua-Sorzabal A, Cantoni L, Dickinger A, Gretzel U, Jannach D, Neidhardt J, Pröll B, Ricci F, Scaglione M, Stangl B, Stock O, Zanker M (2015) Future research issues in it and tourism. *Inf Technol Tour* 15(1):1–15
- Xiang Z, Magnini VP, Fesenmaier DR (2015) Information technology and consumer behavior in travel and tourism: insights from travel planning using the internet. *J Retail Consum Serv* 22:244–249
- Xiang Z, Du Q, Ma Y, Fan W (2017) A comparative analysis of major online review platforms: implications for social media analytics in hospitality and tourism. *Tour Manag* 58:51–65. <https://doi.org/10.1016/j.tourman.2016.10.001>
- Xie HJ, Miao L, Kuo PJ, Lee BY (2011) Consumers' responses to ambivalent online hotel reviews: the role of perceived source credibility and pre-decisional disposition. *Int J Hosp Manag* 30(1):178–183. <https://doi.org/10.1016/j.ijhm.2010.04.008>
- Xie KL, Zhang Z, Zhang Z (2014) The business value of online consumer reviews and management response to hotel performance. *Int J Hosp Manag* 43:1–12. <https://doi.org/10.1016/j.ijhm.2014.07.007>
- Xu X, Li Y (2016) The antecedents of customer satisfaction and dissatisfaction toward various types of hotels: a text mining approach. *Int J Hosp Manag* 55:57–69
- Zanker M, Rook L, Jannach D (2019) Measuring the impact of online personalisation: past, present and future. *Int J Human-Comput Stud* 131:160–168
- Zhao Y, Xu X, Wang M (2019) Predicting overall customer satisfaction: big data evidence from hotel online textual reviews. *Int J Hosp Manag* 76:111–121
- Zhu F, Zhang X (2010) Impact of online consumer reviews on sales: the moderating role of product and consumer characteristics. *J Mark* 74(2):133–148
- Zwerina K, Huber J, Kuhfeld WF (1996) A general method for constructing efficient choice designs. Fuqua School of Business, Duke University, Durham, NC, pp 39–59. <http://support.sas.com/techsup/technote/mr2010e.pdf>. Accessed Sept 2019

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