Decision Making Based on Bimodal Rating Summary Statistics-An Eye-Tracking Study of Hotels

Coba, Ludovik; Zanker, Markus ; Rook, Laurens

DOI
10.1007/978-3-030-05940-8_4

Publication date
2019

Document Version
Final published version

Published in
Information and Communication Technologies in Tourism 2019

Decision Making Based on Bimodal Rating Summary Statistics - An Eye-Tracking Study of Hotels

Ludovik Coba[1(✉)], Markus Zanker[1], and Laurens Rook[2]

[1] Free University of Bozen - Bolzano, Piazza Domenicani, 3, Bolzano, Italy
{lucoba,mzanker}@unibz.com
[2] TU Delft, Jaffalaan 5, Delft, The Netherlands
L.Rook@tudelft.nl

Abstract. Rating-based summary statistics have become ubiquitous, and of key relevance to compare offers on booking platforms. Largely left unexplored, however, is the issue to what extent the descriptives of rating distributions influence the decision making of online consumers. In this work a conjoint experiment was eye-tracked to explore how different attributes of these rating summarisations, such as the mean rating value, the bimodality of the ratings distribution as well as the overall number of ratings impact users' decision making. Furthermore, participants' maximising behavioural tendencies were analysed. Depending on their scores on *Decision Difficulty*, participants were guided by different patterns in their assessment of the characteristics of rating summarisations, and in the intensity of their exploration of different choice options.

Keywords: e-Tourism · Rating summaries · Conjoint analysis Explanations · Recommender systems

1 Introduction

This research targets the first layer of the framework classifying current research challenges at the intersection of IT and tourism [1], namely the interaction of individuals with web platforms such as online travel agencies (OTAs). Travel planning via OTAs represents a big, but saturated, market [2]. Thus, research on evaluating and comparing the information offers on tourism platforms has a long tradition, and surveys like [3] demonstrate the maturity of this sub-field of e-Tourism research. With the transformation of the traditional web into a participatory and social one, electronic Word-of-Mouth (eWOM) has quickly been recognised as strong influencer of online decision making [4], and as an important determinant of business performance [5]. Consequently, the analysis of the contents of online review platforms [6] and research analysing how different aspects

J. Pesonen and J. Neidhardt (Eds.): *Information and Communication Technologies in Tourism 2019*, pp. 40–51, 2019.
https://doi.org/10.1007/978-3-030-05940-8_4

of online reviews influence decisions like [7] are widespread. The present work, however, focuses on a very particular and, so far, largely unexplored aspect, namely how the statistical characteristics of rating summarisations influence users' decision making. The study participants had to rank different tourism offers (i.e., hotels) that, in their perception, only differed in their rating summarisations – that is, in the total number of ratings, the mean rating value, and the bimodality of the rating distribution on a discrete scale. The actual content of reviews or descriptions of the accommodation services et cetera were not shown. A rank-based conjoint experiment was conducted, supplemented with eye-tracking in order to measure the focus of participants' attention, and to understand how they value the displayed differences in their decision making process.

In addition, since decision making strategies vary from person to person [8], it was hypothesised that users high on dispositional maximisation would behave differently from those low on dispositional maximising (so-called satisficers) in such an experiment. This study shows that these two groups trade-off the characteristics of the aggregated rating distributions in different ways and behave – under the lens of the eye-tracker – according to the initial hypothesis grounded in decision making theory.

After shortly referring to related work on decision making strategies and conjoint studies in the context of recommendations in Sect. 2, a detailed description of the conjoint and eye-tracking methodologies is provided in Sect. 3. Section 4 presents obtained results, and Sect. 5, discusses and outlines the implications of this work.

2 Related Work

Constructing consumers' decision making process has been a focal point for many years [9,10]. The literature on recommender systems acknowledges that the decision making strategies differ from person to person. For example, people employing a non-compensatory strategy, like the *Satisficing* principle, would differ in the extent to which they search for "the best possible item", and/or settle for "a good enough alternative given the circumstances" [8,11–13].

The Satisficing principle stems from the seminal work by Herbert Simon [11], which describes the satisficing nature of human decision making. Simon argued that, in a "rational" model, a person would explore a set of multi-attribute items until she finds an item that exceeds some minimum acceptance level. This strategy falls into the category of non-compensatory approaches. In contrast, in a compensatory approach, a consumer would make a trade-off between high values on one characteristic and low values on another when determining the overall utility [9].

Barry Schwartz and colleagues [12] developed a self-reporting scale to assess personal differences in people's maximising behaviour. People scoring high on the maximisation scale manifest a tendency towards determining the best choices for themselves, while people scoring low on this scale, so-called satisficers, settle sooner for a "good enough" alternative, and, opposed to maximisers, are less likely to experience regret and low choice satisfaction after making a decision.

In the field of recommender systems, relatively few studies consider the maximising decision making process. Bart Knijnenburg and his colleagues [8], in a between-subjects experiment, controlled the interfaces to provoke different decision making strategies. However, contrary to Schwartz's theory, they observed that maximisers were more rather than less satisfied with their choices. Jugovac et al. [13] even reported null effects in the presence of recommendations.

Conjoint analysis is a widely appreciated methodological tool from marketing and consumer research, which is particularly applicable to the study of user preferences and trade-offs in the decision making process [14]. In the field of recommender systems and online decision support, Zanker and Schoberegger [15] employed a ranking-based conjoint experiment to understand the persuasive power of different explanation styles over users' preferences that also included a product category from the field of tourism and hospitality.

In his inspiring work, John Payne investigated conjunctive models by explicitly collecting respondents' verbal commands [9]. Latent preferences of the user can be captured with conjoint experiments [14]. Glaholt et al. [16] and Orquin and Loose [17] suggested that eye movement and gazes reflect the screening and evaluation of choices.

By complementing conjoint analysis with eye-tracking, the authors shed more light on how information is actually processed, and observe how the study participants build their decisions.

3 Methodology

The authors conducted a within-subjects study to better understand the differences in the decision making strategies between maximisers and satisficers. The analysis was based on the Rank-Based Conjoint (RBC) methodology, supported by an eye-tracking device. In conjoint designs, items (a.k.a., *profiles*, see example in Fig. 1) are modelled by sets of categorical or quantitative *attributes*, which have different *levels*, cf. [18]. The RBC experiment is designed such that participants rank items from the most to the least preferred one. This design feature nicely matches real-world settings, where users are confronted with lists of tourism services [19].

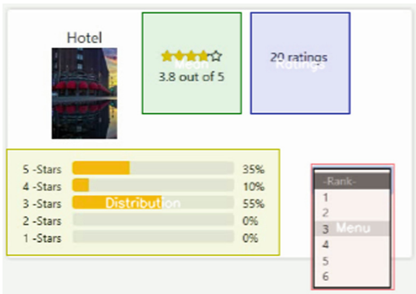


Fig. 1. Profile snapshot.

3.1 Attribute Selection and Study Design

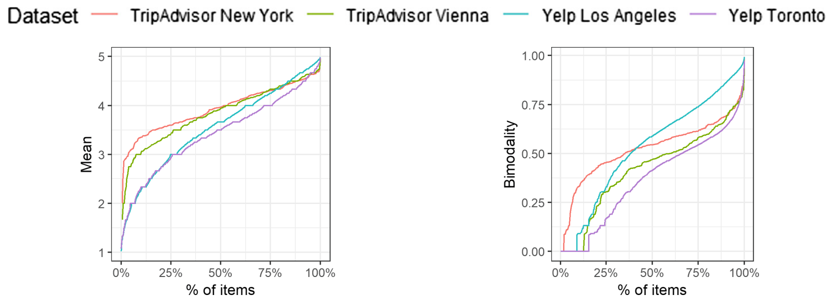
Earlier study results [20,21] showed that the mean rating value is actually the most important predictor, and that the number of ratings has only a mediocre influence on the user's choices. However, the latter is only true if the number of ratings is perceived as relatively high (i.e. in the three digits and above), while, when the number of ratings are in the double digits, users tend to put more weight on this aspect when making choices. Thus, users are willing to trade in a slightly lower mean rating value for a higher number of overall ratings, which obviously makes the mean rating value appear more reliable. Only minor or no effects were observed, when the research was extended to additional characteristics of rating distributions, such as the variance or skewness [22]. However, Hu et al. [23] observed that rating distributions actually exhibit an asymmetric bimodal (J-shaped) distribution. This J-shaped distribution can be explained by the *purchasing bias* (i.e., one tends to buy what one likes) and an *under-reporting bias* (i.e., polarised opinions are more likely to be reported). Thus, the mean rating value has been considered as a biased measure for product quality [24]. Therefore, it was hypothesised 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. The range of these "peaks" was measured using the bimodality coefficient, as recommended by [25]. The bimodality coefficient is computed as:

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

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

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 recognised as such. Values above this threshold clearly exhibit a bimodal distribution (see examples in Fig. 2).

Rating summarisations are typically depicted as frequency distributions on the class of discrete ratings values (such as one to five stars). This study focused on three distinct attributes: *number of ratings*, *mean rating*, and *bimodality*.



(a) Mean (b) Bimodality

Fig. 3. Statistical descriptives of datasets.

These three attributes were used to develop the stimuli (i.e., the profiles/items) of rating summaries, and kept variance and skewness constant.

In order to validate the plausibility of the selected levels for the three attributes, they were aligned with real data from the tourism and hospitality domain crawled from TripAdvisor [26,27] and a public dataset from Yelp[1].

Comparison of the mean rating values (Fig. 3a) among the four datasets, as expected, revealed that ratings tend to be skewed towards higher rating values. Likewise, as hypothesised, the bimodality coefficient (Fig. 3b) is nicely spread within all datasets – i.e.,the J-shaped rating distributions occur in tourism data. The number of ratings was set at 20 and 80, since it was observed in an earlier study that participants clearly notice the difference between these levels [20]. Mean rating values from 3.6 to 4.0 are representative for many rated items on Tripadvisor (see Fig. 3a). The bimodality coefficient was varied from 0.3, which means there is no noticeable second peak present, to 0.7 which clearly indicates the unanimity of reviewers. Table 1 summarises the selected levels for the three attributes of rating summarisations in the study. Based on the identified attribute levels, a *full-factorial design* [14] was built that included all possible combinations of attributes and levels – that is, a design, which consisted of 3 attributes (2 levels $\times$ 3 levels $\times$ 3 levels), and resulted in 18 different profiles that were put to test. Importantly, all items represented statistically feasible level combinations. The profiles were blocked into three subsets in order to lower the cognitive load for respondents to feasible levels. In other words, they had to rank 3×6 alternatives.

Table 1. 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

[1] https://www.yelp.com/dataset/challenge.

3.2 Metrics and Analysis

Conjoint analysis. One basic assumption is the *additive utility model.* It assumes that the different attributes and characteristics of an item/profile contribute, independently of each other, to the overall utility. Individual utilities were estimated for each item characteristic [14], i.e., the respondent's preferences could be modelled via a utility function $u(\boldsymbol{x_i})$, Formula 2, representing the overall utility, which a respondent assigns to an item.

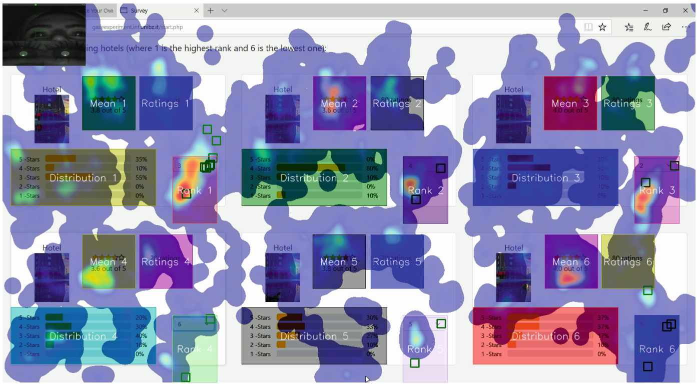


Fig. 4. Screenshot of one task that also highlights the Areas of Interest.

$$u(\boldsymbol{x_i}) = \boldsymbol{x_i}\boldsymbol{\beta} + \epsilon \tag{2}$$

where $\boldsymbol{x_i}$ is a vector characterising a profile i, $\boldsymbol{\beta}$ constitutes the vector with the preference values for each attribute level, and ϵ is the residual error. The utility $u(\boldsymbol{x_i})$ of an item $\boldsymbol{x_i}$ is the sum of the partial utilities for each attribute.

Eye-tracking: Areas of Interests. Areas of Interest (AOIs) [28] are regions defined in the stimulus in order to extract data specifically for those areas. Three AOIs were observed per item (see Fig. 1) specifying the three attributes. The dwell or *gaze* refers to a focal visit of an AOI, from entry to exit, while a gaze cluster constitutes a *fixation.* It would be considered a hit on an AOI, if the participant locked his gaze into a specific area, spending the minimum time that it takes to cognitively process the information therein [28]. 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. Fixation time was observed, since maximisers supposedly spend more time assessing their choices than satisficers [12]. As is explained in Sect. 4, sample sizes were small (below 25 per group of maximisers and satisficers). In order to present the best estimate of the average task time, therefore, the geometrical mean and the log-transformation of the confidence interval was used [29]. In addition, revisits indicate how decision makers examine alternatives, the amount of information searched, and the comparisons performed in order to complete the task [9].

Decision Difficulty. Several scales exist to assess individual differences in maximising versus satisficing behavioural tendencies, ranging from the 13-item Maximisation Scale [30] to several shorter forms. In the present study, the shortened 6-item scale proposed by [31] was used. It comprises the following items: "When I am in the car listening to the radio, I often check other stations to see if something better is playing, even if I am relatively satisfied with what I'm listening to", "No matter how satisfied I am with my job, it's only right for me to be on the lookout for better opportunities", "I often find it difficult to shop for a gift for a friend", "Booking a hotel is really difficult. I'm always struggling to pick the best one"[2], "No matter what I do, I have the highest standards for myself", and "I never settle for second best". Each item was measured on a 7-point scale ranging from 1 (*completely disagree*) to 7 (*completely agree*). The sum of these six items yielded an overall maximisation score [30,31]. The overall scale as well as each of the three sub-scales individually are reliable with Cronbach's α within the ranges outlined by [31]. Schwartz et al. [30] proposed to use the appropriate sub-scales, depending on the purpose of the analysis. Here the participants had no possibility to include other solutions or tourism offers that they could search for, but were instructed to rank exactly six items that only differed based on their rating summaries. The analysis was thus focused on the decision difficulty sub-scale that distinguishes maximisers, who frequently experience difficulties when making decisions, due to their attitude to always go for the best choices. Satisficers, in contrast, seem to settle quicker for a solution. It was therefore hypothesised that participants experiencing more decision difficulty would need longer, as they would compare the differences between the different offers more intensely.

3.3 Study Procedure

Participants volunteered to take part in a controlled lab experiment on a preconfigured terminal. The stimuli were presented on a 22" display, and the gazes were recorded with a static remote eye-tracking system utilising a 150Hz research-grade machine-vision camera. Participants (having given informed consent to have their data used for research purposes) were asked to fill in the shortened Maximisation Scale described above. Next, the experiment was started from a remote console, where participants were asked to consider the following tourism-inspired decision making and ranking task:

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

[2] Note, that the original, outdated, phrase *Renting a video[..]* in the scale of [31] in the present research was replaced with *Booking a hotel[..]*.

Following this introduction, the participant went through 3 consecutive ranking tasks. In each task the respondents had to rank 6 items out of 18 from the full factorial combinations of the attribute levels. Each task was created such that attributes levels were equally distributed and balanced between tasks to allow proper measurement of main effects. A screenshot of a task is presented in Fig. 4[3]. Presented profiles were item-agnostic, thus users could base their choice solely on the three study variables (i.e., number of ratings, mean, and bimodality). The order of the six displayed items (i.e., the profiles) on a single screen was randomised for each respondent. Additional feedback (on which characteristics of rating summaries guided decisions most) and demographic information were included in a post-questionnaire.

4 Results

In June 2018, *42 participants* took part in the eye-tracking experiment. Table 2 presents the demographics of the participants in the sample. Based on the participants' scores on the *Decision Difficulty* sub-scale, a median split of the sample was performed to contrast the decisions of the two groups. Table 3 shows that participants experiencing more decision difficulty (i.e. above median) had a tendency to strongly rely on the higher mean ($\beta = 1.35, p < .001$); they would less likely select an alternative with a higher number of ratings ($\beta = 0.76, p < .001$) rather than a slightly lower mean. In contrast, respondents that scored low on *Decision Difficulty* seemed almost equally likely to choose the high mean or the high number of ratings alternative – that is, they could more confidently trade-in different attribute levels of these two characteristics of rating summary statistics against each other in their decision making strategies. However, in contrast to the initial hypothesis, the bimodality of distributions did not noticeably influence the decisions for both groups – i.e., having the choice between a mean rating

Table 2. Participants' demographic details.

Personal feature	Category				Total
Gender	**Male**	**Female**	**No answer**		
#	24	18	0		**42**
%	57.1%	42.9%	0%		**100%**
Age	**18–4**	**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%**

[3] The implementation of the platform is downloadable at: https://github.com/ludovikcoba/rankBasedConjoint.git.

value of 4.0 or 3.8, where the rating distribution with a mean of 4.0 would also show a stronger J-shape, participants would still confidently go for the higher mean, although there were more low rating values present. Furthermore, results presented in Table 3 are perfectly in line with the observation made in a previous study with a sample of 200 respondents [22]. Figure 5 shows the geometrical mean of the fixation times on the items ordered by the ranking positions they received from the respective participant, and grouped according to the median split on *Decision Difficulty*. Note that the time was computed only for the first ranking task in order to account for the learning effect.

Table 3. 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	–		–	

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

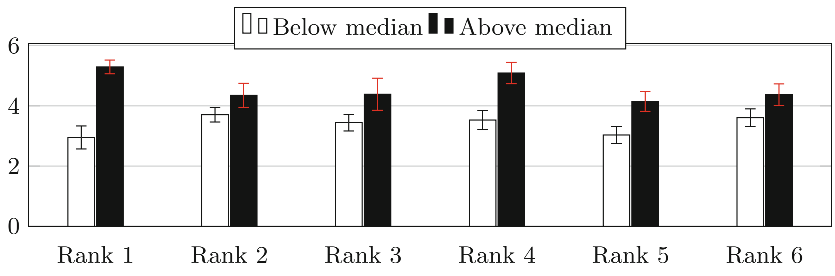


Fig. 5. Geometrical mean of the time spent on item (confidence level of 95%), median split on decision difficulty sub-scale.

As hypothesised, results suggest that respondents experiencing more difficulties in making decisions also spent more time in inspecting alternatives.

Payne [9] suggested that maximisers would compare the different alternatives more frequently before making a decision. Consequently, Fig. 6 shows the mean number of revisits – i.e., how frequently the participant's gazes switched between an AOI of one item and the AOI of another item, forth and back – grouped by median split and ordered based on the actual rank assigned by the participant. Thus, although the confidence levels are large due to sample sizes, Fig. 6 supports the hypothesis.

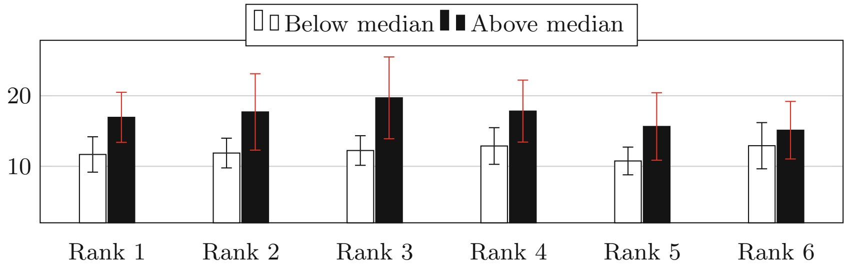


Fig. 6. Mean number of revisits per item, median split on the decision difficulty subscale.

5 Discussion and Conclusions

This paper presented a study that eye-tracked a rank-based conjoint (RBC) experiment. The purpose was to explore how different configurations of rating summary statistics, which take the total number of ratings, the mean rating value and the bimodality of the distribution into account, steer users' decision making under a *ceteris paribus* assumption. While the authors in an earlier study observed a trade-off between the number of ratings and the mean rating value, the results of the present study indicate that the bimodality of distributions has only a minor effect on user choices. Furthermore, participants' behaviour was investigated under the lens of different decision styles, with a specific focus on decision difficulty as a result of a maximising behaviour. A detailed analysis of users' gazes showed that maximisers, and in particular those high on the decision difficulty sub-dimension underlying maximisation, spent more time comparing the different alternatives, but still primarily relied on high mean rating values. In contrast, their counterparts with less decision difficulty were more free to weight in against each other different characteristics of rating distributions, and needed less time to decide. This outcome is remarkable, since one would expect a predominant reliance on higher mean rating values to be a simplistic decision heuristic. Nevertheless, the detailed analysis of gazes disclosed that these participants had actually spent more gaze time and compared more by jumping back and forth. A limitation of this study is that the estimated effects are based on the median split of a sample of 42 participants. Although this sample size is clearly commensurate for eye-tracking studies, it should also be acknowledged that it reduced the power of the conjoint analysis.

The practical implications of the present work lie in the area of personalisation and ranking algorithms. Clearly, the differences of users, when assessing and evaluating alternative offers on a booking platform, must be taken into account. Needless to say, future work will include further developments in order to tune recommendation algorithms according to users' presumed decision making styles.

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

References

1. 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
2. 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
3. Law R, Qi S, Buhalis D (2010) Progress in tourism management: a review of website evaluation in tourism research. Tour Manag 31(3):297–313
4. 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, pp 35–46, Vienna, Springer Vienna
5. 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
6. Xiang Z, Qianzhou D, 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
7. Xie H (Jimmy), Miao L, Kuo P-J, Lee B-Y (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
8. Knijnenburg BP, 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, pp 141–148
9. 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
10. Bettman JR, Luce MF, Payne JW (1998) Processes, constructive consumer choice, source. J Consum Res 25(3):187–217
11. Simon HA (1955) A behavioral model of rational choice. Q J Econ 69(1):99–118
12. Schwartz B, Ward A, Lyubomirsky S, Monterosso J, White K, Lehman DR (2002) Maximizing versus satisficing: happiness is a matter of choice. J Personal Soc Psychol 83(5):1178–1197
13. Jugovac M, Nunes I, Jannach D (2018) Investigating the decision-making behavior of maximizers and satisficers in the presence of recommendations. In: 26th conference on user modeling, adaptation and personalization (UMAP'18)

14. Zwerina K, Huber J, Kuhfeld WW (1966) A general method for constructing efficient choice designs. Durham, NC: Fuqua School of Business, Duke Univesrsity, 39–59
15. Zanker M, Schoberegger M (2014) An empirical study on the persuasiveness of fact-based explanations for recommender systems. CEUR Workshop Proc 1253:33–36
16. Glaholt MG, Reingold EM (1985) Eye movement monitoring as a process tracing methodology in decision making research 4(2):125–146 (1985)
17. Orquin JL, Loose SM (2013) Attention and choice: a review on eye movements in decision making. Acta Psychologica 144(1), 190–206:9
18. Rao VR (2008) Developments in conjoint analysis. In: Handbook of marketing decision models, pp 23–53. Springer
19. Chung J, Rao VR (2012) A general consumer preference model for experience products: application to internet recommendation services. J Market Res 49(3):289–305
20. Coba L, Zanker M, Rook L, Symeonidis P (2018) Exploring users' perception of collaborative explanation styles. In: 2018 IEEE 20th conference on business informatics (CBI), pp 70–78
21. Coba L, Zanker M, Rook L, Symeonidis P (2018) Exploring users' perception of rating summary statistics. In: Proceedings of the 26th conference on user modeling, adaptation and personalization, UMAP 18, 353–354. ACM, New York, NY, USA (2018)
22. Coba L, Zanker M, Rook L, Symeonidis P (2018) Decision making of maximizers and satisficers based on collaborative explanations. arXiv preprint arXiv:1805.11537
23. Hu N, Zhang J, Pavlou PA (2009) Overcoming the J-shaped distribution of product reviews. Commun ACM 52(10):144
24. 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
25. 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
26. Fuchs M, Zanker M (2012) Multi-criteria ratings for recommender systems: an empirical analysis in the tourism domain. In: Lecture notes in business information processing, vol 123 LNBIP, pp 100–111
27. 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
28. Holmqvist K, Nyström M, Andersson R, Dewhurst R, Jarodzka H, Van De Weijer J (2011) Eye tracking. Oxford University Press, A comprehensive guide to methods and measures
29. Sauro J, Lewis JR (2010) Average task times in usability tests. In: Proceedings of the 28th international conference on Human factors in computing systems - CHI '10, p 2347, New York, New York, USA, 2010. ACM Press
30. Schwartz B, Ward A, Monterosso J, Lyubomirsky S, White K, Lehman DR (2002) Maximizing versus satisficing: happiness is a matter of choice. J Perso Soc Psychol 83(5):1178
31. 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