

Steur, Andreas J.; Fritzsche, Fabian; Seiter, Mischa

Article — Published Version

It's all about the text: An experimental investigation of inconsistent reviews on restaurant booking platforms

Electronic Markets

Provided in Cooperation with:

Springer Nature

Suggested Citation: Steur, Andreas J.; Fritzsche, Fabian; Seiter, Mischa (2022) : It's all about the text: An experimental investigation of inconsistent reviews on restaurant booking platforms, Electronic Markets, ISSN 1422-8890, Springer, Berlin, Heidelberg, Vol. 32, Iss. 3, pp. 1187-1220, <https://doi.org/10.1007/s12525-022-00525-3>

This Version is available at:

<https://hdl.handle.net/10419/311907>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



It's all about the text: An experimental investigation of inconsistent reviews on restaurant booking platforms

Andreas J. Steur¹ · Fabian Fritzsche¹ · Mischa Seiter¹

Received: 22 March 2021 / Accepted: 10 January 2022 / Published online: 18 March 2022
© The Author(s) 2022

Abstract

Consumer-generated reviews play a decisive role in creating trust and facilitating transactions on digital platforms. However, prior research shows various problems, e.g., only a small number of consumers providing reviews, fake reviews, and inconsistent reviews. We use an experiment in the context of a restaurant booking platform to examine the impact of inconsistent reviews on the duration of consumers' transaction decisions. In a second experiment, we investigate the relative importance of the review components in the case of inconsistent reviews. Drawing on the dual-process theory and media richness theory, we predict that inconsistent reviews result in a longer time required for consumers' transaction decisions (H1) and lead to users' transaction decisions being predominantly based on the qualitative component (H2). Although we do not find general support that inconsistent restaurant reviews negatively determine the duration of transaction decisions, we find evidence that in the case of inconsistent restaurant reviews, the polarity of the qualitative component is crucial for both the duration of the transaction decision and the decision itself.

Keywords Consumer reviews · Online reviews · Inconsistent reviews · Reputation · Digital platform · Consumer decision-making

JEL Classification M31

Introduction

Electronic markets such as digital platforms have evolved into a central business sector that has a major impact on the modern economy (Alt & Zimmermann, 2014). However, due to the lack of personal interaction on platforms, they have to incentivize trustworthiness (Bolton et al., 2013). Therefore, many platforms use consumer reviews, which provide additional information about sellers and previous transactions and thus contribute to building reputation and creating trust

in sellers (Hesse & Teubner, 2020). On the one hand, online reviews lead to positive economic outcomes for sellers as they increase prices (Ba & Pavlou, 2002) and transaction volume (Bolton et al., 2004). On the other hand, reviews are important for consumers because they increasingly rely on reviews to make transaction decisions (Bae & Lee, 2011). In particular, informative consumer reviews are crucial for selecting restaurants, hotels, and other services (Ruiz-Mafe et al., 2018), as they help to differentiate sellers and minimize uncertainty (Nazlan et al., 2018).

Reviews are consumer-generated online information that allows consumers to share information about previous transactions and make informed purchase decisions (Vallurupalli & Bose, 2020). They typically consist of two components: quantitative (e.g., stars) and qualitative (e.g., text) (Gutt et al., 2019). Quantitative components are ratings on a pre-defined scale, with both the scale and the symbol of the rating varying across platforms (Steuer & Seiter, 2021). Qualitative components mainly consist of text and complement the quantitative component, allowing consumers to provide detailed information, as is the case with Amazon and Yelp.

Responsible Editor: Markus Bick.

✉ Andreas J. Steur
andreas.steur@uni-ulm.de
Fabian Fritzsche
fabian.fritzsche@uni-ulm.de
Mischa Seiter
mischa.seiter@uni-ulm.de

¹ Institute of Business Analytics, Ulm University,
Helmholtzstraße 22, 89081 Ulm, Germany

However, the two components are not necessarily aligned within a review. Inconsistency between the quantitative and qualitative components of products, services, or sellers has been identified for several platforms. For instance, 30% of Amazon's product reviews (Mudambi et al., 2014) were inconsistent. Also, multiple seller reviews on TripAdvisor (6%) (Fazzolari et al., 2017) and the German physician rating platforms jameda and DocInsider (more than 3%) (Geierhos et al., 2015) were inconsistent.

There are several reasons for inconsistent reviews. Users can make mistakes in the submission or intentionally write inconsistent reviews. Errors can arise when users do not rate items carefully, either forgetting to rate certain parts of the transaction performance or mistakenly giving inaccurate ratings. Moreover, inconsistent reviews can be intentional when reviewers want to provide a more differentiated view. Compared to the quantitative component, more detailed information can be conveyed within the qualitative component, which can also be observed in the frequent positive and negative text sequences within reviews (Ruiz-Mafe et al., 2018). Such inconsistent reviews are particularly relevant in consumers' decision-making, as consumers use reviews for information search and seller evaluation (Mudambi and Schuff 2010).

On digital platforms, the decision-making process typically consists of five steps (Darley et al., 2010). Beginning with the problem recognition, consumers become aware of the discrepancies between the actual and the desired state (Del Hawkins & Mothersbaugh, 2010), such as finding a suitable restaurant. Subsequently, consumers search through a large set of relevant sellers based on their search terms and constraints to identify a subset of the most promising restaurants (Xiao & Benbasat, 2007). In the next step, consumers process information and evaluate the subset of restaurants, thus seeking further information about the restaurants related to their search criteria. Therefore, consumers use the information provided by the restaurants and the individual consumer reviews. This step ends with consumers making their restaurant decision (Darley et al., 2010). Consumers complete the transaction by purchasing the selected alternative (Del Hawkins & Mothersbaugh, 2010). Finally, the decision process ends with the purchase outcomes (Darley et al., 2010). In our case, consumers experience the quality of the chosen restaurant.

Consumer reviews typically facilitate both information search and seller evaluation (Bae & Lee, 2011). Inconsistent reviews send conflicting information and signal different levels of quality to consumers which could have several consequences, such as increased cognitive processing costs, suboptimal purchase decisions, and lower overall utility of the platform (Mudambi et al., 2014). As these reviews can be confusing for consumers (Geierhos et al., 2015), inconsistency in reviews could negatively affect the review

helpfulness (Aghakhani et al., 2020). Hence, the conflicting information could affect the effort involved in the information processing and the evaluation of a subset of restaurants within the decision-making process. This information processing is usually measured by the duration of consumers' transaction decisions and the extent of consumers' search (Xiao & Benbasat, 2007). The longer decision duration could hamper the well-known positive effects of review systems on sellers' prices and transaction volume for sellers on digital platforms (e.g., Ba & Pavlou, 2002; Bajari & Hortaçsu, 2003; Bolton et al., 2004, 2013; Melnik & Alm, 2002; Resnick et al., 2006). Thus, consumers could either use other platforms to make further transaction decisions or leave without completing the decision-making process.

Prior research on consumer reviews has shown that the two components are frequently not aligned.¹ However, only limited attention has been given to studying the effects of inconsistent reviews on the consumer decision-making process (Aghakhani et al., 2020). Tsang and Prendergast (2009) analyzed the effects of inconsistent product critiques on consumer decision-making. In particular, they found that inconsistency in movie critiques hampers their trustworthiness. Aghakhani et al. (2020) showed that consistent reviews positively affect review helpfulness as a proxy for review quality.

Mudambi et al. (2014) presumed that inconsistent reviews increase consumers' cognitive processing costs, resulting in consumers taking more time to make transaction decisions, making suboptimal purchase decisions, and lowering the overall utility of the platform. In the case of inconsistent product critiques, the interplay of both components negatively affects trustworthiness (Tsang & Prendergast, 2009). Moreover, it is not yet fully understood which review component determines the decision. Thus, we propose the following two research questions:

- (1) *How do inconsistent reviews affect the duration of transaction decisions?*
- (2) *Which review component – quantitative or qualitative – determines the transaction decision in the case of inconsistent reviews?*

Drawing on dual-process theory and media richness theory, we conducted two experiments. We used a 2 × 2 within-subjects design for the first experiment, in which 442 participants chose one of four restaurants after being exposed to either consistent or inconsistent reviews. In our second

¹ In particular, previous research focused on two types of review inconsistencies: individual and collective inconsistencies. Individual inconsistencies refer to a misalignment within a review (i.e., inconsistencies in the quantitative and qualitative review component), whereas collective inconsistencies refer to differing reviews across raters. In this study, we focus on individual inconsistencies.

experiment, we applied a similar design, in which 233 participants decided between two restaurant options: one with positive quantitative and negative qualitative reviews vs. one with negative quantitative and positive qualitative reviews.

We find that inconsistent restaurant reviews did not necessarily result in a longer duration for consumers' transaction decisions. However, our experimental results show that in the case of inconsistent restaurant reviews, reviews with positive texts resulted in faster transaction decisions. Moreover, in the case of inconsistent restaurant reviews, positive qualitative review components determine consumers' decision-making.

Our study makes two contributions to the literature on online reviews. First, our work advances the literature on inconsistent online reviews by showing the effect of inconsistent reviews on the duration of user decision-making. Second, our research extends the literature on the relative importance of review components, as the polarity of review texts is crucial for the duration of the transaction decision and the decision itself.

The remainder of the paper is organized as follows. In Sect. 2, we provide a brief overview of the related literature, and we develop our hypotheses in Sect. 3. Section 4 covers the experimental studies, including the method and the results. In Sect. 5, we discuss our findings and their theoretical and managerial implications. We summarize the work and highlight further research directions in Sect. 6.

Literature on inconsistent reviews

Inconsistent reviews have been addressed in several studies. While most studies have focused on the occurrence of inconsistent reviews, the analysis of their effects on consumer decision-making is lacking.

Fu et al. (2013) analyzed individual inconsistencies on the Google Play Store and found 1% of the reviews to be inconsistent. Mudambi et al. (2014) discussed individual inconsistencies using different Amazon products (e.g., books and cameras) and showed that 30% of the reviews were inconsistent. In addition, Mudambi et al. (2014) showed that individual inconsistencies were more common for high star ratings. Moreover, their results showed that inconsistencies were more common for experience goods (41%) than for search goods (16%). Search goods (e.g., toasters) have attributes that can be objectively compared, while experience goods (e.g., books) have attributes that are subjective to the user (Nelson, 1970). However, the study by Mudambi et al. (2014) was based on a small data set of only 1,734 reviews on 23 products on Amazon.

Fazzolari et al. (2017) focused on individual inconsistent reviews on TripAdvisor using sentiment analysis. Their findings showed that 6% of 164,300 hotel reviews

on TripAdvisor were inconsistent. Specifically, they found an asymmetrical occurrence of inconsistencies. While 5% of the reviews with positive quantitative components were inconsistent, 12% of negative ones were inconsistent.

Geierhos et al. (2015) developed a method to identify inconsistencies for reviews with multiple quantitative criteria. Focusing on the two German physician rating platforms, jameda and DocInsider, they found varying inconsistencies within the categories, from 3% for "time" (time taken for the treatment) to 12% for "responsiveness" (accessibility and waiting time) (Geierhos et al., 2015).

Shan et al. (2018) focused on the occurrence of individual inconsistencies within authentic and fake reviews. They found that for authentic reviews, quantitative components and the sentiment value of the qualitative components have a higher positive correlation than fake reviews.

In recent literature, inconsistent reviews have been acknowledged, but their effects on consumers and their transaction decisions remain unclear (Mudambi et al., 2014). Aghakhani et al. (2020) found that consistent reviews positively affect review helpfulness (Aghakhani et al., 2020). While review helpfulness is a measure of review quality (Aghakhani et al., 2020; Mudambi and Schuff 2010) and thus provides first insights into the effects of inconsistent reviews, knowledge on the effects on consumer decision-making is still lacking.

Tsang and Prendergast (2009) were the first to investigate how the interplay between quantitative and qualitative components affects consumers' decision-making processes. Using the example of movie critiques, they particularly emphasized the trustworthiness of consistent and inconsistent critiques and the related effects on purchase intention. Their results showed that the qualitative component has a much stronger influence on purchase intention and trustworthiness. Additionally, they found that positive reviews produced a higher purchase intention than inconsistent or negative reviews. Moreover, inconsistent critiques did not produce a higher purchase intention than negative critiques (Tsang & Prendergast, 2009). However, Tsang and Prendergast (2009) focused on purchase intention, which does not correlate perfectly with actual purchases (Morwitz, 1997). Moreover, their study included only a single product critique (consisting of a quantitative and a qualitative review component) and analyzed its interestingness, trustworthiness, and purchase intention.

Table 1 provides an overview of the related literature on inconsistent reviews and presents our contribution against this background.

Prior research focused on the occurrence of individual inconsistencies. However, prior studies lack an examination of whether inconsistent reviews affect actual consumer decision-making. We aim to close this research gap by analyzing two effects of inconsistent restaurant reviews on consumer

Table 1 Literature on inconsistent reviews

Authors	Findings	Product/Service	Data
Fu et al. (2013)	1% of the reviews were inconsistent	App reviews	Observational (Google Play Store)
Mudambi et al. (2014)	<ul style="list-style-type: none"> ■ Search goods 16% ■ Experience goods 41% 	<ul style="list-style-type: none"> ■ Search goods: Cameras, coffee makers, grills, and toasters ■ Experience goods: books, music CDs, MP3 players, and diapers 	Observational (Amazon)
Fazzolari et al. (2017)	<ul style="list-style-type: none"> ■ Reviews with positive quantitative components 5% ■ Reviews with negative quantitative components 12% 	Hotel reviews	Observational (Booking.com and TripAdvisor)
Geierhos et al. (2015)	<ul style="list-style-type: none"> ■ Assurance 5% ■ Reliability 11% ■ Responsiveness 12% ■ Tangibility 8% ■ Time 3% 	Physician reviews	Observational (jameda and DocInsider)
Shan et al. (2018)	Inconsistency is more common with fake reviews than with authentic reviews	Restaurant reviews	Observational (Yelp)
Aghakhani et al. (2020)	Negative effect of inconsistencies on review helpfulness	Restaurant reviews	Observational (Yelp)
Tsang and Prendergast (2009)	Negative effect of inconsistencies on trustworthiness	Movie critiques	Experimental

decision-making. First, we examine whether inconsistent reviews affect the decision duration. Second, we analyze the relative importance of quantitative and qualitative review components for consumer decision-making.

Theoretical background and hypotheses development

Effects of inconsistent reviews on the duration of transaction decisions

When evaluating a subset of restaurants with inconsistent reviews that send conflicting signals, consumers have to process and evaluate this conflicting information. Thus, the cognitive processing could increase (Mudambi et al., 2014) the time needed for consumer decision-making. Therefore, inconsistent reviews could increase search costs within the decision-making process and negatively affect conversion rates.

According to dual-process theory, information search is based on heuristic processing (System 1) and systemic processing (System 2) (Chaiken, 1980). In particular, System 1 generates impressions, feelings, and inclinations which System 2 can then check and either accept, modify, or override (Kahneman, 2013; Kahneman & Frederick, 2002). The processes of System 1 run automatically, quickly, and in parallel, in an associative form without deliberate control (Kahneman, 2013). Conversely, System 2 involves deliberate

and conscious abstract and hypothetical thinking; it is slow and sequential and occupies the capacity of the central working memory system (Evans, 2003). For this purpose, System 2 uses the working memory (Evans, 2003) and is therefore used rather sparingly (Kahneman, 2013). The control of the impressions by System 2 is thus relatively lenient (Kahneman & Frederick, 2002). Based on the least effort and sufficiency principles of System 2 (Baek et al., 2012), we argue that in cases of consistent reviews, consumers' transaction decisions are made relatively quickly based on the impressions of System 1, which System 2 merely accepts.

Conversely, both systems are fundamental for consumers' decision-making in the cases of conflicting information within textual reviews (Ruiz-Mafe et al., 2018). Specifically, Ruiz-Mafe et al. (2018) analyzed conflicting qualitative review components. They found that in cases where positive text sequences follow negative sequences, consumers evaluate the argument quality, and System 2 modifies the emotions and impressions of System 1. Moreover, Aghakhani et al. (2020) used the dual-process theory to explain the effects of inconsistent reviews on review helpfulness. Thus, we assume that in cases of inconsistent reviews, System 1 continues to provide impressions to System 2. If consumers of digital platforms using System 1 detect inconsistent reviews, System 2 refuses to simply accept the impressions of System 1 (Kahneman, 2013; Kahneman & Frederick, 2002). Instead, we assume that System 2 uses working memory to check, modify, or override these impressions.

Within inconsistent reviews, consumers will primarily use System 2 since inconsistent reviews require logical thinking and reasoning. Therefore, we assume that decisions based on inconsistent reviews (assuming that consumers recognize these inconsistencies) are slower than decisions based on consistent reviews because of the greater involvement of the slower System 2 (Kahneman, 2013). Specifically, we hypothesize that inconsistencies are processed in the third step of the consumer decision-making process, where consumers evaluate alternatives. In this stage, consumers process the information on the restaurant options, such as the individual reviews, and evaluate these options according to their pre-defined decision criteria (Del Hawkins & Mothersbaugh, 2010). Thus, following the dual-process theory, inconsistent reviews should affect the information processing and evaluation such that the duration of the decision-making process increases. Hence, we propose:

H1: Inconsistent reviews result in a longer time required for consumers' transaction decisions.

The Relative importance of quantitative and qualitative review components

Confronted with inconsistent reviews, consumers have to emphasize either the quantitative or the qualitative component of such reviews when assessing important seller characteristics during the decision-making process. Following prior research on online reviews (e.g., Xu et al., 2015; Zinko et al., 2020), we use media richness theory (Daft & Lengel, 1984), which initially described communication behavior and media choice within organizations (Daft & Lengel, 1986), to better understand the consumer decision-making process. According to media richness theory, the ability of a medium to convey and promote mutual understanding depends on its information richness. In the case of high message ambiguity, a high-richness medium is appropriate (Daft et al., 1987). In contrast, low richness suffices in repetitive, routine communication settings (Trevino et al., 1987).

Consumer reviews including quantitative and qualitative components (Gutt et al., 2019) help users of digital platforms to receive information on previous transactions. The different review components have different levels of information richness (Daft & Lengel, 1986). In the case of inconsistent reviews, quantitative and qualitative review components send conflicting quality signals. Hence, inconsistent reviews are associated with higher ambiguity compared to consistent reviews. Therefore, consumers noticing inconsistencies have to decide which component is more trustworthy.

Due to the unique nature of inconsistent reviews, the processing of inconsistent reviews on digital platforms can generally be considered to be a non-routine task. For information with high uncertainty in a non-routine setting, a medium

of greater richness is more appropriate (Trevino et al., 1987). However, the media (i.e., written and numeric media), through which reviews are shared on digital platforms are constant and cannot be changed to a medium with a greater richness (e.g., face-to-face) (Daft & Lengel, 1984). Therefore, consumers must decide on the trustworthiness of both components and their individual information richness due to the different quality signals.

The information richness of both review components varies in terms of the quantitative and qualitative components. The quantitative review component ("numeric language") has a lower level of information richness than the qualitative component ("natural language") (Daft & Lengel, 1984). The depth of the qualitative component enhances the relevance of the qualitative component as it is known for the review helpfulness (Mudambi and Schuff 2010).

Prior research has shown that higher information richness is associated with higher trustworthiness (Lu et al., 2014) and that the information quality of online reviews affects purchase intention (Zinko et al., 2020). Thus, given that recipients of reviews benefit from using a richer medium in situations of ambiguity, inconsistent reviews should result in more weight being placed on the qualitative review components. The findings of Tsang and Prendergast (2009) support this argumentation. In situations of inconsistent product critiques, they showed that textual reviews have a higher significance in terms of purchase intention and trustworthiness of reviews. Hence, we assume that the relative importance of the qualitative component is higher within the decision-making process, and we propose:

H2: In the case of inconsistent reviews, users' transaction decisions are predominantly based on the qualitative component.

Experimental studies

Experiment 1: Duration of the transaction decision

Basic setup and treatments

We conducted a controlled experiment in the context of restaurant reviews to test our hypothesis regarding whether inconsistent reviews affect the duration of transaction decisions. Similar to Schneider et al. (2021), we developed a fictitious restaurant-visit scenario where participants had to choose a restaurant option based on previous consumer reviews. In experiment 1, we used a 2×2 within-subjects design (Table 2). In each treatment, the participants chose one of four restaurants, for which they were shown reviews with quantitative and qualitative components. The first two treatments included consistent reviews; that is, both the

Table 2 Treatments of experiment 1

Quantitative components		Qualitative components	
		Positive	Negative
		Treatment 1 (<i>pos-pos</i>)	Treatment 2 (<i>pos-neg</i>)
	Positive		
	Negative	Treatment 3 (<i>neg-pos</i>)	Treatment 4 (<i>neg-neg</i>)

quantitative and qualitative components were positive in treatment 1 (*pos-pos*) and negative in treatment 4 (*neg-neg*). In the other two treatments, the reviews were inconsistent. Specifically, treatment 2 consisted of positive quantitative and negative qualitative components (*pos-neg*), and treatment 3 included negative quantitative and positive qualitative components (*neg-pos*).

Within the four treatments, all participants received five reviews for each of the four restaurants. Following Fazzolari et al. (2017), we only focus on strong inconsistent reviews in our research. Hence, the quantitative component included only negative reviews (one or two out of five stars) and positive reviews (four or five out of five stars). As in Ruiz-Mafe et al. (2018), the qualitative component consisted of real textual reviews about Italian restaurants posted on Yelp. We only selected reviews that did not contain any people's names, geographical regions, or restaurant names, to avoid biases within the experiment due to familiarity (Ruiz-Mafe et al., 2018). Moreover, we ensured that none of the reviews included information about the reviewer.

Similar to prior studies (e.g., Geierhos et al., 2015; Mudambi et al., 2014), a sentiment analysis helped classify the qualitative review components regarding their polarity. In particular, we used restaurant reviews of the publicly available Yelp data set (Yelp, 2020). We used TextBlob for our sentiment analysis, as it is a widely used library (e.g., Kühl et al., 2020; Mousavi et al., 2020) that offers high accuracy. TextBlob provides a polarity score for each piece of a review ranging from -1 for extreme negative sentiment to 1 for extreme positive sentiment of the text. We only selected texts with extreme sentiment values to ensure that participants noticed the inconsistencies and avoid distortion due to different sentiment values (polarity of -1, or 1). We used extreme polarities to ensure a uniform design across the treatments. To prevent a false classification based on the previous sentiment analysis (false positive or false negative), we used five independent raters who received the qualitative components in random order to conduct the sentiment classification for further analysis. All classified reviews were consistent with the corresponding polarities of the sentiment analysis.

Procedure

We conducted our experiment in July 2020 using Amazon Mechanical Turk (MTurk). We chose MTurk for three reasons. First, as a digital platform, MTurk has digitally competent users with characteristics similar to those of users of other digital platforms (Vazquez, 2021). Second, recent studies have shown that MTurk experiments provide similar data quality and results to traditional methods such as laboratory studies (Horton et al., 2011). Third, MTurk was used in other studies concerning online reviews (e.g., Garnefeld et al., 2020; Zinko et al., 2020). The experiment was performed using the software oTree (Chen et al., 2016).

Participants were freelancers on MTurk who earned \$1.50 for participating in the experiment. A total of 1,405 participants attended the experiment. However, 308 participants did not finish the experiment. We conducted attention checks (Abbey & Meloy, 2017), one after the instructions (see Appendix 20) and one after all decision rounds (see Appendix 26), to verify that users carefully read the reviews provided for each restaurant. These included questions about the business category, the scale level of the quantitative component, and the type of cuisine used within the experiment. Also, we asked the participants how often inconsistent reviews occurred within the experiment to control whether the participants recognized the inconsistent reviews. A total of 635 participants failed this test. Additionally, we excluded another 20 participants that deviated in their decision duration from the mean by more than three standard deviations. The final sample consisted of 442 participants (Table 3).

Each participant took part in all four treatments. However, the individual treatments were ordered randomly to avoid order effects (e.g., learning and framing) (Charness et al. 2012). Initially, the participants were instructed to imagine a situation in which they would go out for dinner to a restaurant with friends, and their attentiveness was checked for the first time (see Appendix 20). Prior to the first selection of a restaurant, all participants could practice the selection in two practice rounds.

After the training phase, the participants were explicitly asked to begin the selection phase. Within the selection phase, the participants chose one of four restaurants (see Appendix 21, Appendix 22, Appendix 23, Appendix 24). For each restaurant, the participants were shown five different reviews consisting of qualitative and quantitative components. Additionally, an overall rating was displayed, which represented an average value of all quantitative reviews. We showed this overall rating because most platforms do so, and therefore we assume a high external validity. Within each treatment, the duration from the display of the restaurant reviews to confirmation of the decision was measured.

Finally, the participants answered the second attention check (see Appendix 26) and responded to a questionnaire

Table 3 Characteristics of the participants of experiment 1

Age			
18–29	30–49	50 or older	
133 (30.1%)	259 (58.6%)	50 (11.3%)	
Gender			
Male	Female		
289 (65.4%)	153 (34.6%)		
Education level			
No formal education	Primary education	Secondary education	Tertiary education
1 (0.2%)	28 (6.3%)	160 (36.2%)	253 (57.2%)

including several controls. Since previous studies (e.g., Hennig-Thurau & Walsh, 2003; von Helversen et al., 2018) found differences in decision-making in relation to age, gender, highest completed level of education, and risk preferences, we used these variables as controls (see Appendix 28). For these controls, we used established constructs (Hennig-Thurau & Walsh, 2003; Ludwig et al., 2017; von Helversen et al., 2018). We also asked the participants about their frequency of using reviews in general, using reviews to choose a restaurant, and visiting restaurants. Additionally, we elicited information from the participants about their choice of individual restaurants.

Results of the duration of transaction decisions

Table 4 shows the average decision-making time required by the participants within each treatment. In the case of inconsistent *pos-neg* reviews, the participants needed significantly more time for their decision compared to all other treatments.

Table 5 shows the result of *t*-tests between the different variants.

Table 4 Decision duration

	Average duration required for decision	σ	Std err
Treatment 1 (<i>pos-pos</i>)	38.57 s	32.30 s	1.54 s
Treatment 2 (<i>pos-neg</i>)	52.00 s	42.31 s	2.01 s
Treatment 3 (<i>neg-pos</i>)	38.35 s	27.67 s	1.32 s
Treatment 4 (<i>neg-neg</i>)	43.59 s	34.71 s	1.63 s

Table 5 Results of two-tailed *t*-tests

	<i>Pos-pos</i>	<i>Pos-neg</i>	<i>Neg-pos</i>	<i>Neg-neg</i>
<i>Pos-pos</i>	-	-8.0567*	0.1700	-3.5546*
<i>Pos-neg</i>	8.0567*	-	8.2649*	4.9662*
<i>Neg-pos</i>	-0.1700	-8.2649	-	-3.8375*
<i>Neg-neg</i>	3.5546*	-4.9662*	3.8375*	-

* indicates < 0.05 significance.

The two-tailed *t*-test for two dependent samples showed that the participants decided more quickly with *pos-pos* reviews than with *neg-neg* reviews. The decisions with *pos-pos* reviews were also faster than with *pos-neg* reviews. However, there was no significant difference in decision duration between *pos-pos* reviews and *neg-pos* reviews. With *neg-neg* reviews, decisions were faster than with *pos-neg* reviews. In contrast, the decision took longer with *neg-neg* reviews than with *neg-pos* reviews. Finally, decisions took significantly longer with *pos-neg* reviews than with *neg-pos* reviews.² Thus, we did not find support for H1.

Experiment 2: The relative importance of quantitative and qualitative components

Basic setup and treatment

In our second experiment, we used a design similar to the first one. However, the second experiment focused on the relative importance of the review components. Therefore, the experiment consisted of only one treatment (Table 6). The participants were asked to decide between two restaurants to examine which component dominated the transaction decisions in the case of inconsistent reviews (see Appendix 25). For one of the two restaurants, reviews included positive quantitative (four or five out of five stars) and negative qualitative components (polarity of -1). In contrast, the reviews of the other restaurant consisted of negative quantitative (one or two out of five stars) and positive qualitative components (polarity of 1).

Table 6 Options within experiment 2

	Option 1 (<i>pos-neg</i>)	Option 2 (<i>neg-pos</i>)
Quantitative components	Positive	Negative
Qualitative components	Negative	Positive

² The order in which participants worked through the experiment had no systematic influence on the results of the *t*-tests. Moreover, the results are robust to the inclusion of additional control variables.

Procedure

We conducted our second experiment in July 2020. The participants who completed the experiment received \$0.70. A total of 713 participants took part in the experiment, but 190 participants stopped the experiment prematurely. As in the first experiment, we identified random clickers using attention checks (see Appendix 20 and Appendix 27). A total of 290 participants failed at least one of the attention checks and were thus excluded from the analysis. The final data set for experiment 2 consisted of 233 participants (Table 7).

After participants were shown the instructions, we conducted a first attention check to filter random clickers and bots (see Appendix 20). Then, the participants practiced within two training rounds (as in experiment 1). Following the training rounds, the participants started the selection round, in which they chose between a restaurant with *pos-neg* reviews and a restaurant with *neg-pos* reviews. The order of the two restaurants was random. Further, both restaurants were displayed in the same way as in experiment 1.

Finally, the participants' attention was checked in another attention check (see Appendix 27), similar to the first experiment. This attention check helped to ensure that the participants noticed the inconsistencies within the reviews. The participants concluded by answering the final questionnaire using the same controls as in the first experiment (see Appendix 29). Since the content within the qualitative review component can affect information processing and decision-making, we controlled for argument quality. In line with prior research, we measured argument quality by perceived informativeness and perceived persuasiveness (Zhang et al., 2014).

Results on the relative importance of quantitative and qualitative components

We conducted a one-tailed binomial test ($p = 0.5$) to analyze whether the participants based their decisions on the

Table 7 Characteristics of the participants of experiment 2

Age		
18–29	30–49	50 or older
75 (32.2%)	141 (60.5%)	17 (7.3%)
Gender		
Male	Female	Other
160 (68.7%)	72 (30.9%)	1 (0.4%)
Education level		
Primary education	Secondary education	Tertiary education
15 (6.4%)	68 (29.2%)	150 (64.4%)

Table 8 Logit regression results

$N = 232$	Coef	Std err
<i>Const</i>	2.2883*	0.853
<i>Age</i>	-0.0177	0.017
<i>Female</i>	-0.0183	0.343
<i>Education</i>	1.0907*	0.344
<i>Review use in general</i>	-0.5377	0.327
<i>Review use to choose a restaurant</i>	0.0374	0.364
<i>Restaurant use</i>	-0.2282	0.344
<i>Risk</i>	-0.1906*	0.066
<i>Perceived informativeness</i>	-0.1644	0.150
<i>Perceived persuasiveness</i>	-0.0078	0.106
<i>Df Model</i>		9
<i>Pseudo R^2</i>		0.08206
<i>Log-Likelihood</i>		-134.69
<i>LL-Null</i>		-146.73
<i>LLR p value</i>		0.004173

* indicates < 0.05 significance.

qualitative component. The one-tailed binomial test result was significant ($\alpha = 0.05$).³ Thus, H2 was supported.

We conducted a logit regression (Table 8) to gain more insight into the decision-making process. In particular, a low level of *perceived informativeness* and *perceived persuasiveness* resulted in a lower probability that users made their decisions based on qualitative components. However, these results were not significant. Moreover, *education* and *risk* controls showed significant effects. Specifically, participants with higher education levels were more likely to choose the restaurant with positive qualitative components. Likewise, the higher was the willingness to take risks, the lower was the probability that decisions were based on qualitative components. Other control variables did not show any significant results.

Summary of the findings

Our analysis of experiment 1 showed that only *pos-neg* reviews resulted in slower transaction decisions compared to consistent reviews. Since decisions did not take longer for *neg-pos* reviews than consistent reviews, H1 cannot be supported. The results of experiment 2 showed that participants predominantly chose the restaurant with the positive qualitative component, supporting H2. The results of H1 and H2 are shown in Table 9.

³ The order in which the participants worked through the experiment had no systematic influence on the result of the binomial test.

Table 9 Results of the hypotheses

Hypothesis	Test	Supported
H1: Inconsistent reviews result in a longer time required for consumers' transaction decisions	Two-tailed <i>t</i> -test	No
H2: In the case of inconsistent reviews, users' transaction decisions are predominantly based on the qualitative component	One-tailed binomial test	Yes

Discussion

Theoretical implications

We contribute to the literature on inconsistent online reviews in two ways. First, our results show that inconsistent restaurant reviews do not generally lead to longer transaction decisions. Whereas Ruiz-Mafe et al. (2018) showed that System 2 intervenes when positive sequences follow negative sequences, we did not find that *neg-pos* reviews led to longer transaction decisions. This result could indicate that the effects of inconsistent reviews on the duration of consumer decisions are not as decisive as previously expected. However, since reviews and their effects on consumer decision-making are highly complex (Xiao & Benbasat, 2007), and since inconsistent reviews negatively affect review helpfulness (Aghakhani et al., 2020), inconsistencies could still hamper the effectiveness of reviews.

We found no significant time differences in decision-making between *neg-pos* and *pos-pos* reviews. In contrast, consumers took longer to make decisions based on *neg-neg* or *pos-neg* reviews, indicating that positive polarities of the qualitative components are crucial for the decision duration and that quantitative components have no effect on the decision duration. One possible explanation for the text's importance is the context of restaurant reviews. The perceived quality of restaurants is very subjective, which is why many users might focus on the qualitative components to get a more detailed overview of the reviewed restaurant. The importance of qualitative review components might also help to explain the lack of significant differences in the time required for consumers to make a transaction decision based on *pos-pos* reviews and *neg-pos* reviews. These findings are in line with previous literature showing that sufficient positive information within a text is required in the information processing and the evaluation of a subset of restaurants within the decision-making process (Tsang & Prendergast, 2009). Thus, the polarity of the review texts is crucial for the decision duration, which might explain why we did not find support for H1. In the case of reviews with positive qualitative components, we found that the participants made decisions faster regardless of the quantitative component.

Second, we showed that in the case of inconsistent restaurant reviews, consumers decided based on the qualitative

rather than the quantitative component. These results indicate that in the case of inconsistent restaurant reviews, consumers are motivated to exert extra effort to minimize uncertainties due to the intangible nature of restaurant service (Nazlan et al., 2018). These findings are in line with our expectations based on media richness theory and the findings of Tsang and Prendergast (2009) in the context of product critiques. Whereas Tsang and Prendergast (2009) focused on both components' importance using purchase intention as a proxy, we extended the literature by focusing on actual decisions.

Managerial implications

Based on our findings that showed the importance of the qualitative component, we propose three managerial implications to highlight the qualitative review component. First, we recommend platforms to focus on textual reviews to simplify the review system and facilitate consumers' decision-making. For instance, rating platforms could use the polarity score drawn from sentiment analysis to replace the quantitative review component. This polarity score of the individual reviews could then be aggregated to an overall rating. Thus, the aggregated polarity score could strengthen the textual component when screening through restaurants within the decision-making process.

Second, we suggest highlighting keywords that enable consumers to filter individual reviews. Therefore, platforms could use text mining to identify relevant topics within the qualitative review components. These topics could facilitate the evaluation of the restaurant subset within consumers' decision-making process. For instance, Amazon highlights keywords within their product reviews. However, currently, keywords are not a common feature within review systems on restaurant platforms, such as Yelp or TripAdvisor. In addition, platforms could offer further sorting and filtering options to promote the text as an essential review component. Thus, users could easily find reviews that fit their search requests. As a result, new peer evaluation methods of the qualitative component, such as "helpfulness" as a kind of perceived value for decision-making, could be of high importance in the design of review systems.

Third, we recommend platforms to split the qualitative review component into multiple parts. Platforms could

introduce multiple criteria, such as service quality, food quality, or ambiance, on which users can submit a qualitative review. Another approach could be to split the qualitative component into positive and negative parts. This adjusted design could offer a more detailed and differentiated view of the restaurant visit. Moreover, this design could help to indicate that the two components of restaurant reviews are not necessarily consistent.

Conclusion

Consumer reviews have become increasingly popular in electronic markets, especially on digital platforms. As inconsistent reviews frequently occur in consumer reviews on digital platforms, this paper was motivated by the lack of knowledge about the effects of such reviews on consumer decision-making. Based on the first experiment, we show the effects on the duration of transaction decisions. Subsequently, we also show in the second experiment that the decisions are predominantly based on the qualitative review components. However, this paper is subject to several limitations.

First, we assume that inconsistent reviews are related to a switch from System 1 to System 2. However, with our online experiment, we could not directly measure the switch between these systems in consumer decision-making. Second, our analyses focused on strongly inconsistent reviews. Although we could not generally show that inconsistent

reviews negatively determine the duration of transaction decisions, future studies might look at different degrees of inconsistency with lower differences in quantitative and qualitative review components to examine their effects on consumer decision-making. Moreover, within the experiments, all reviews of a restaurant were either consistent or inconsistent. However, as consumer reviews typically consist of both consistent and inconsistent reviews, both types of reviews could be examined in further studies. Third, our experiments focused on decisions among a few restaurants with a limited number of reviews. While this choice is a good approximation of reality, it does not fully represent the complex structure of consumer decision-making. Moreover, the relative importance of the textual review component could result from the uniqueness of restaurant reviews, as they are subjective in terms of the quality perceived by the consumer. Therefore, the relative importance could be different for other products or services. Future studies could consider different price levels and types of consumption (e.g., accommodations, books, cameras, or cars). Finally, we conducted our experiments on Amazon MTurk without incentivizing the decision quality, as participants received a flat payment.

Appendix A

Selection Content (Experiment 1 and 2).

Before you start the practice phase, please answer the following question based on the instructions you read on the previous page.

Which business category will be used in this survey?

- ☐ Active Life
- ☐ Automotive
- ☐ Beauty & Spas
- ☐ Home Services
- ☐ Nightlife
- ☐ Restaurants
- ☐ Shopping

Appendix B

Decision Round (*Pos-Pos*, Experiment 1).

Please choose the restaurant that pleases you the most.

Restaurant 1 

Review 1 

Owner is extremely sweet and cares about her customers satisfaction!!!!
Closest Chicago style pizza to my home! Best pizza I've ever had.

Review 2 

we ordered a specialty pizza yesterday. and it was one of the best i have ever had!

Review 3 

This place has great pizza! We took a coworker there for her birthday lunch.
Perfect place for a celebration. They even sang happy birthday and made her a dessert pizza. Can't beat it!!

Review 4 

Excellent pesto pizza with a perfectly chewy crust. I wish this place was in my neighborhood.

Review 5 

I had their signature pizza. This is some of the best pizza I have ever had!!
The flavors went so well together. The crust was delicious. Everything delicious!

Restaurant 2



Review 1



*Delicious food! I went there for my birthday. The service was great.
I recommend the brick oven pizza!*

Review 2



*Best freaking "thicker" pizza on the west side! Must get the 8corner pizza!
The crust is the buttery and crispy with burnt cheese on the corners!*

Review 3



*The best Chicago style pizza in town! Being from the east coast, that should
say a lot*

Review 4



*Their pizza needs to be in your mouth. BUTTERY crust. Like, butter buttery.
And it's superb reheated at home in the toaster oven.*

Review 5



*5 stars it is! Best pizza in town! The house made mozzarella is the best and
the service is always excellent.*

Restaurant 3**Review 1**

Excellent pizza! From the generous toppings and tasty cheese all the way to the bottom of the wonderful crust, this is some of the best pizza I have ever tasted. Try it for yourself- you won't regret it!

Review 2

Probably some of the best pizza I have had in awhile

Review 3

*You have to try their pizza- quality meat and the crust isn't over powering.
Perfect ratio of crust to toppings*

Review 4

This pizza is so yummy. The updated look is awesome and the environment is so chill. A place you will always want to go eat!

Review 5

This was some of the best pizza I've had. The crust was perfect. We will be ordering takeout again from here. Delicious pizza and calzone.

Restaurant 4



Review 1



Excellent! We just happened to stop in and got delicious food and delightful service.

Review 2



awesome tasty pizza....going there again tonight

Review 3



The best pizza that i ever eat!! The service its amazing!!!! I always go there with my friends they give me a feeling likei make pizza at home.

Review 4



Best pizza ive ever had ! My family have dinner here often. The waitress are great remembers are drinks and food every time. I recommend this place to everyone!

Review 5



Pepperoni and garlic toppings is all you need with this delicious sauce and cheese on the pizza!

Which restaurant do you choose?

- ☐ Restaurant 1
- ☐ Restaurant 2
- ☐ Restaurant 3
- ☐ Restaurant 4

Appendix C

Decision Round (*Pos-Neg*, Experiment 1).

Please choose the restaurant that pleases you the most.

Restaurant 1

Review 1

Awful pizza. I've been twice and will not return.

Review 2

The worst pizza ever the crust was doughy. The cheese was sogreasy. Will never go here ever again

Review 3

This is a horrible place for pizza. They have kids work there that to rip you off. Beware!!

Review 4

Worst service ever! Sat at 6:30, ignored for 20 mins, ordered food and never arrived...we had to leave for a show at 7:45.

Review 5

This place is disgusting I can't believe I ate this pizza look at the state of the restroom

 Springer

Restaurant 2



Review 1



This is a horrible place , you go in here and none of the staff look at you let alone say anything. They do not greet or acknowledge at all. No customer service and the pizza food and service .

Review 2



Horrible service. Waited 15min after being seated and not one server came and took our order.

Review 3



terrible service. terrible workers. don't go here at all.

Review 4



This was the worst food I have ever eaten. The pizza tasted like it was from a bowling alley and undercooked. The sandwich was like rubber. They came to the door with no change and took my money and said he had to go. This place is overpriced and horrible!!! Go somewhere else!!!

Review 5



Terrible! Took my money and didn't deliver what I ordered. Fraudulent pizza company! Don't pay till you check your order.

Restaurant 3



Review 1



Pasta was terrible, pizza was horrible. This is a place that uses tourists ignorance and charges an arm and leg. Stay away.

Review 2



Awful...took an hour to get our pizza! Worst service ever and the manager took ZERO responsibility! Yelled at me as a result

Review 3



It's terrible. All four of us had stomach after visiting this store. Not recommended.

Review 4



Don't eat here, food was disgusting and service was horrible. This was the worst food I have ever had. Over priced

Review 5



Sicilian pizza ? Tastes like plastic and cardboard. This has to be one of the worst pizzas I've had in quite some time.

Restaurant 4**Review 1**

Pizza taste like cardboard and games are horrible. The prizes are just crappy as they get.

Review 2

Awful food service and way to expensive !!!! Would not return!

Review 3

I promise from the bottom of my heart this place is the worst place to order anything from. The people are HORRIBLE to customers and I just....nevermind...it's upsetting me even writing about them.

Review 4

Horrible customer service. Took almost 2 hours to get food after they told me 4560 minutes. And pizza was nasty at this location.

Review 5

The worst service ever. If there was another pizza place nearby this one would be out of business.

Which restaurant do you choose?

- ☐ Restaurant 1
- ☐ Restaurant 2
- ☐ Restaurant 3
- ☐ Restaurant 4

Appendix D

Decision Round (Neg-Pos, Experiment 1).

Please choose the restaurant that pleases you the most.

Restaurant 1

Review 1

Yes, yes, yes. Perfect in almost every way.

Review 2

Best pizza I've had in a looong time. Reminds me of the pizza I had in Italy. I've been craving it every day since. Delicious!

Review 3

The friendliest people , they remember our order and the very best sausage and mushroom pizza ever !!

Review 4

With out a doubt the best pizza in this area I've been ordering from the restaurant from years you have to try them !

Review 5

Got their meat lovers pizza. Awesome.

 Springer

Restaurant 2**Review 1**

Best pizza in my neighborhood. I could eat here every day. Yum.

Review 2

Great service and prices!

Review 3

*This is the best pizza I have had in the valley yet, plenty of cheese sauce and toppings.
I will be returning.*

Review 4

*My family loves this place the food has always been delicious. No complaints from us
and the staff is awesome!*

Review 5

*The food was very delicious. I had a clam pasta and a prosciutto pizza. Both were
delicious. I'll have to come here again!*

Restaurant 3



Review 1



Best pizza and drinks around! Thank you for re-opening and look forward to seeing you in February.

Review 2



Great job! Get the lunch size macaroni and cheese with a pizza. Great stuff!

Review 3



Amazing pizza!!! Eat in, carry out, or have them delivery it to your door! Mega meat pizza is our go-to!! Yummy!!!

Review 4



Best pizza and best costumes service and they sell a variety of food that's just as delicious a must try

Review 5



All I can say is YUM. Best pizza in the city without a doubt. And the art in the place is out of control!

Restaurant 4



Review 1



Excellent pizza, service and beer selection. I look forward to another visit!

Review 2



Great Chicago style pizza! A+ in my book.

Review 3



*Been going here for years! Love the owner the people who work there and the pizza!
YUMMY!!!*

Review 4



*Great food!! Ended up ordering one of the pizzas mmmm cubed pepperoni Awesome
staff and the scones are to die for!!!*

Review 5



Best pizza in the city. Period.

Which restaurant do you choose?


- ☐ Restaurant 1
- ☐ Restaurant 2
- ☐ Restaurant 3
- ☐ Restaurant 4

Appendix E


Decision Round (Neg-Neg, Experiment 1).

Please choose the restaurant that pleases you the most.

Restaurant 1




Review 1




Terrible service, horrible prices for the type of food. Ghost town and once we were served, I can understand why.

Review 2




It is the worst Pizza in had in 10 years it has no flavor the tomatoes aren't cook the feda cheese is not salty the bacon has no flavor it's awful.. !

Review 3




Waited 2 hours for pizza, owners are liars. Terrible customer service, take your business elsewhere. Terrible, terrible, terrible...

Review 4



Very horrible quality in food I would not recommend them to anyone I don't see how there in business and it smells like cigaret smoke

Review 5



This pizza was probably one of the worst pizzas I've ever had, no flavor in dough, pizza sauce, even the pepperoni!! Would not recommend to pizza lovers. hope this helps someone

Restaurant 2**Review 1**

They didn't even cut my pizza..... Awful customer service as soon as I walked in the door.

Review 2

Terrible service terrible food so to sum up terrible!!

Review 3

Horrible customer service. I don't understand how this location is still in business. Do not order pizza here. Go anywhere else.

Review 4

Worst service ever! This place has the slowest service ever. We waited for about 50 mins for a pizza and pasta. You'll lose appetite when the food arrives.

Review 5

Horrible pizza. The shop at that location looks out of place. Not recommended, even if you are on a budget.

Restaurant 3



Review 1



This was the worst pizza that I've ever had. My stomach hurt from the doughy crust. It was terrible. I don't know what the ingredients are in the dough, but I won't eat it again.

Review 2



Horrible service I don't recommend this place at all save your money and time get food somewhere else

Review 3



Don't order from this place!! The manager is horrible and pizza is equally disgusting.

Review 4



Food is terrible. Service is lacking but may be because understaffed. Either way horrible experience. Would not recommend

Review 5



Worst service, pizza was Burnt. Bunch of kids that have no idea what they are doing. Terrible experience!

Restaurant 4**Review 1**

This place was horrible. They don't even deserve a one-star. The pizza was soggy and disgusting. Bad customer service!!

Review 2

Horrible service . It's like a bother to ask them for a drink . Never coming again !!!!!

Review 3

*Awful pizza. Ordered a handful of times and everything was awful all the time.
Probably won't ever order from there again.*

Review 4

Terrible service. I've gotten a pissed off cashier all three times I've ordered from this location.

Review 5

This place is awful . The pizza is awful . Don't understand how people are return customers .

Which restaurant do you choose?


- ☐ Restaurant 1
- ☐ Restaurant 2
- ☐ Restaurant 3
- ☐ Restaurant 4

Appendix F


Decision Round (Experiment 2).

Please choose the restaurant that pleases you the most.

Restaurant 1



Review 1




Best pizza I've ever had.

Review 2




Excellent. Service, food, beverages. Everything. We order a pizza and spaghetti meatballs.

Review 3




I have tried almost everything on the menu and it is all delicious! The place is beautiful! I am hoping to meet the interior designer!

Review 4



Everything about this place is awesome... the ambiance, the wine selection, the food, and service... if u r looking for a perfect date....or trying to impress... brunch, lunch, or dinner... this is the spot

Review 5



The Best Pizza In Town! The best tasting pizza I've ever had and they have great deals!

Restaurant 2



Review 1



Horrible service! Crappy food! We waited 45 min at our table before we even got any type of food.

Review 2



Worst pizza I've had. They don't make it like they used to. Everything looked like it was slapped on with out care no pride what so ever. no parmesan cheese & peppers

Review 3



The worst pizza, I have ever had. Over priced and just not edible. I wouldn't recommend wasting your time or money here.

Review 4



Service is horrible had to ask for water 4 times. Waitress has no clue about the drinks or menu.

Review 5



Terrible. Some of the worst service and worst pizza I've ever had.

Which restaurant do you choose?

- ☐ Restaurant 1
☐ Restaurant 2

Appendix G

Feedback Content (Experiment 1).

Please answer the following questions corresponding to the previous tasks.

How many stars were shown within each review?

- ☐ stars from 1 to 3
- ☐ stars from 1 to 4
- ☐ stars from 1 to 5
- ☐ stars from 1 to 6
- ☐ stars from 1 to 7
- ☐ stars from 1 to 10
- ☐ stars from 1 to 20

Which cuisine was offered in the restaurants of the previous study?

- ☐ Chinese
- ☐ Indian
- ☐ Italian
- ☐ Greek
- ☐ Mexican
- ☐ Spanish
- ☐ Thai

How often did the star ratings and the textual ratings differ within the individual customer reviews in the decision rounds?

- ☐ Never
- ☐ Sometimes
- ☐ Always

Appendix H

Feedback Content (Experiment 2).

Please answer the following questions corresponding to the previous tasks.

How many stars were shown within each review?

- ☐ stars from 1 to 3
- ☐ stars from 1 to 4
- ☐ stars from 1 to 5
- ☐ stars from 1 to 6
- ☐ stars from 1 to 7
- ☐ stars from 1 to 10
- ☐ stars from 1 to 20

Which cuisine was offered in the restaurants of the previous study?

- ☐ Chinese
- ☐ Indian
- ☐ Italian
- ☐ Greek
- ☐ Mexican
- ☐ Spanish
- ☐ Thai

Did the star rating and the text differ within the individual customer reviews in the decision round?

- ☐ Yes
- ☐ No

Appendix I

Final Questions (Experiment 1).

What is your age?

What is your gender?

☐ Male
☐ Female
☐ Other

What is the highest level of education you have completed?

☐ No formal education
☐ Primary education
☐ Secondary education
☐ Tertiary education

Please indicate on a scale ranging from 0 to 10 whether you are willing to take risks (or try to avoid risks). 0 represents a very weak willingness to take risks, while 10 represents a strong willingness to take risks.

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10

How often do you use customer reviews in electronic markets?

☐ Never
☐ Less than once a month
☐ 1-3 times a month
☐ Once a week
☐ 2-3 times a week
☐ More than 5 times a week
☐ Other

How often do you use customer reviews to choose a restaurant?

☐ Never
☐ Less than once a month
☐ 1-3 times a month
☐ Once a week
☐ 2-3 times a week
☐ More than 5 times a week
☐ Other

How often do you eat out?

☐ Never
☐ Less than once a month
☐ 1-3 times a month
☐ Once a week
☐ 2-3 times a week
☐ More than 5 times a week
☐ Other

Based on which criteria did you choose the restaurants?

☐ Stars
☐ Text
☐ Both

Appendix J

Final Questions (Experiment 2).

What is your age?

What is your gender?

☐ Male

☐ Female

☐ Other

What is the highest level of education you have completed?

☐ No formal education

☐ Primary education

☐ Secondary education

☐ Tertiary education

Please indicate on a scale ranging from 0 to 10 whether you are willing to take risks (or try to avoid risks). 0 represents a very weak willingness to take risks, while 10 represents a strong willingness to take risks.

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10

How often do you use customer reviews in electronic markets?

☐ Never

☐ Less than once a month

☐ 1-3 times a month

☐ Once a week

☐ 2-3 times a week

☐ More than 5 times a week

☐ Other

How often do you use customer reviews to choose a restaurant?

☐ Never

☐ Less than once a month

☐ 1-3 times a month

☐ Once a week

☐ 2-3 times a week

☐ More than 5 times a week

☐ Other

How often do you eat out?

☐ Never

☐ Less than once a month

☐ 1-3 times a month

☐ Once a week

☐ 2-3 times a week

☐ More than 5 times a week

☐ Other

Based on which criteria did you choose the restaurants?

☐ Stars

☐ Text

☐ Both

How would you rate the following statements in relation to your choice of restaurant?

		Strongly agree		Strongly disagree
These reviews provided relevant information about the restaurant.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
These reviews provided complete information about the restaurant.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The arguments of these reviews were convincing.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The arguments of these reviews were persuasive.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The arguments of these reviews were strong.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The arguments of these reviews were good.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Funding Open Access funding enabled and organized by Projekt DEAL. This research was supported by the Péter Horváth Foundation.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abbey, J. D., & Meloy, M. G. (2017). Attention by design: Using attention checks to detect inattentive respondents and improve data quality. *Journal of Operations Management*, 53–56(1), 63–70. <https://doi.org/10.1016/j.jom.2017.06.001>
- Aghakhani, N., Oh, O., Gregg, D. G., & Karimi, J. (2020). Online review consistency matters: An elaboration likelihood model perspective. *Information Systems Frontiers*. <https://doi.org/10.1007/s10796-020-10030-7>
- Alt, R., & Zimmermann, H.-D. (2014). Editorial 24/3: Electronic markets and general research. *Electronic Markets*, 24(3), 161–164. <https://doi.org/10.1007/s12525-014-0163-9>
- Ba, S., & Pavlou, P. A. (2002). Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior. *MIS Quarterly*, 26(3), 243–268.
- Bae, S., & Lee, T. (2011). Product type and consumers' perception of online consumer reviews. *Electronic Markets*, 21(4), 255–266. <https://doi.org/10.1007/s12525-011-0072-0>
- Baek, H., Ahn, J., & Choi, Y. (2012). Helpfulness of online consumer reviews: Readers' objectives and review cues. *International Journal of Electronic Commerce*, 17(2), 99–126. <https://doi.org/10.2753/JEC1086-4415170204>
- Bajari, P., & Hortacsu, A. (2003). The winner's curse, reserve prices, and endogenous entry: Empirical insights from eBay auctions. *The RAND Journal of Economics*, 34(2), 329–355. <https://doi.org/10.2307/1593721>
- Bolton, G. E., Katok, E., & Ockenfels, A. (2004). How effective are electronic reputation mechanisms? An Experimental Investigation. *Management Science*, 50(11), 1587–1602. <https://doi.org/10.1287/mnsc.1030.0199>
- Bolton, G., Greiner, B., & Ockenfels, A. (2013). Engineering trust: Reciprocity in the production of reputation information. *Management Science*, 59(2), 265–285. <https://doi.org/10.1287/mnsc.1120.1609>
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39(5), 752–766. <https://doi.org/10.1037/0022-3514.39.5.752>
- Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree: An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9, 88–97. <https://doi.org/10.1016/j.jbef.2015.12.001>
- Daft, R. L., & Lengel, R. H. (1984). Information richness: A new approach to managerial behavior and organizational design. *Research in Organizational Behavior*, 6, 191–233.
- Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science*, 32(5), 554–571. <https://doi.org/10.1287/mnsc.32.5.554>
- Daft, R. L., Lengel, R. H., & Trevino, L. K. (1987). Message equivocality, media selection, and manager performance: Implications for information systems. *MIS Quarterly*, 11(3), 355–366. <https://doi.org/10.2307/248682>
- Darley, W. K., Blankson, C., & Luethge, D. J. (2010). Toward an integrated framework for online consumer behavior and decision making process: A review. *Psychology & Marketing*, 27(2), 94–116. <https://doi.org/10.1002/mar.2032>
- Del Hawkins, I., & Mothersbaugh, D. L. (2010). *Consumer behavior: Building marketing strategy* (11th ed.). McGraw-Hill Irwin.
- Evans, J. S. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, 7(10), 454–459. <https://doi.org/10.1016/j.tics.2003.08.012>
- Fazzolari, M., Cozza, V., Petrocchi, M., & Spognardi, A. (2017). A study on text-score disagreement in online reviews. *Cognitive Computation*, 9(5), 689–701. <https://doi.org/10.1007/s12559-017-9496-y>
- Fu, B., Lin, J., Li, L., Faloutsos, C., Hong, J., & Sadeh, N. (2013). Why people hate your app: making sense of user feedback in a mobile app store. *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge discovery and data mining*, (pp. 1276–1284). <https://doi.org/10.1145/2487575.2488202>
- Garnefeld, I., Helm, S., & Grötschel, A.-K. (2020). May we buy your love? Psychological effects of incentives on writing likelihood and valence of online product reviews. *Electronic Markets*, 30(4), 805–820. <https://doi.org/10.1007/s12525-020-00425-4>
- Geierhos, M., Bäumer, F., Schulze, S., & Stuß, V. (2015). “I grade what I get but write what I think.” Inconsistency Analysis in Patients' Reviews. *Proceedings of the 23rd European Conference on Information Systems (ECIS)* (pp. 1–15). <https://doi.org/10.18151/7217324>
- Gutt, D., Neumann, J., Zimmermann, S., Kundisch, D., & Chen, J. (2019). Design of review systems: A strategic instrument to shape online reviewing behavior and economic outcomes. *The Journal of Strategic Information Systems*, 28(2), 104–117. <https://doi.org/10.1016/j.jsis.2019.01.004>
- Hennig-Thurau, T., & Walsh, G. (2003). Electronic word-of-mouth: Motives for and consequences of reading customer articulations on the Internet. *International Journal of Electronic Commerce*, 8(2), 51–74. <https://doi.org/10.1080/10864415.2003.11044293>
- Hesse, M., & Teubner, T. (2020). Reputation portability: Quo vadis? *Electronic Markets*, 30(2), 331–349. <https://doi.org/10.1007/s12525-019-00367-6>
- Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, 14(3), 399–425. <https://doi.org/10.1007/s10683-011-9273-9>
- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich (Ed.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). Cambridge: Cambridge Univ. Press.
- Kahneman, D. (2013). *Thinking, fast and slow* (1st paperback ed.). New York: Farrar Straus and Giroux.
- Kühl, N., Mühlthaler, M., & Goutier, M. (2020). Supporting customer-oriented marketing with artificial intelligence: Automatically quantifying customer needs from social media. *Electronic Markets*, 30(2), 351–367. <https://doi.org/10.1007/s12525-019-00351-0>
- Lu, Y., Kim, Y., Dou, X., & Kumar, S. (2014). Promote physical activity among college students: Using media richness and interactivity in web design. *Computers in Human Behavior*, 41, 40–50. <https://doi.org/10.1016/j.chb.2014.08.012>
- Ludwig, S., Fellner-Röhlhling, G., & Thoma, C. (2017). Do women have more shame than men? An experiment on self-assessment and the shame of overestimating oneself. *European Economic Review*, 92, 31–46. <https://doi.org/10.1016/j.eurocorev.2016.11.007>

- Melnik, M. I., & Alm, J. (2002). Does a seller's ecommerce reputation matter? Evidence from eBay auctions. *The Journal of Industrial Economics*, 50(3), 337–349. <https://doi.org/10.1111/1467-6451.00180>
- Morwitz, V. G. (1997). It seems like only yesterday: The nature and consequences of telescoping errors in marketing research. *Journal of Consumer Psychology*, 6(1), 1–29. https://doi.org/10.1207/s15327663jcp0601_01
- Mousavi, R., Raghu, T. S., & Frey, K. (2020). Harnessing artificial intelligence to improve the quality of answers in online question-answering health forums. *Journal of Management Information Systems*, 37(4), 1073–1098. <https://doi.org/10.1080/07421222.2020.1831775>
- Mudambi, S. M., & Schuff, D. (2010). What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Quarterly*, 34(1), 185–200. <https://doi.org/10.2307/20721420>
- Mudambi, S. M., Schuff, D., & Zhang, Z. (2014). Why aren't the stars aligned? An analysis of online review content and star ratings. *47th Hawaii International Conference on System Sciences* pp. 3139–3147.
- Nazlan, N. H., Tanford, S., & Montgomery, R. (2018). The effect of availability heuristics in online consumer reviews. *Journal of Consumer Behaviour*, 17(5), 449–460. <https://doi.org/10.1002/cb.1731>
- Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy*, 78(2), 311–329. <https://doi.org/10.1086/259630>
- Resnick, P., Zeckhauser, R., Swanson, J., & Lockwood, K. (2006). The value of reputation on eBay: A controlled experiment. *Experimental Economics*, 9(2), 79–101. <https://doi.org/10.1007/s10683-006-4309-2>
- Ruiz-Mafe, C., Chatzipanagiotou, K., & Curras-Perez, R. (2018). The role of emotions and conflicting online reviews on consumers' purchase intentions. *Journal of Business Research*, 89, 336–344. <https://doi.org/10.1016/j.jbusres.2018.01.027>
- Schneider, C., Weinmann, M., Mohr, P. N. C., & vom Brocke, J. (2021). When the stars shine too bright: The influence of multidimensional ratings on online consumer ratings. *Management Science*, 67(6), 3871–3898. <https://doi.org/10.1287/mnsc.2020.3654>
- Shan, G., Zhang, D., Zhou, L., Suo, L., Lim, J., & Shi, C. (2018). Inconsistency investigation between online review content and ratings. *24th Americas Conference on Information Systems* pp. 2–11.
- Steur, A. J., & Seiter, M. (2021). Properties of feedback mechanisms on digital platforms: An exploratory study. *Journal of Business Economics*, 91(4), 479–526. <https://doi.org/10.1007/s11573-020-01009-6>
- Trevino, L. K., Lengel, R. H., & Daft, R. L. (1987). Media symbolism, media richness, and media choice in organizations. *Communication Research*, 14(5), 553–574. <https://doi.org/10.1177/009365087014005006>
- Tsang, A. S. L., & Prendergast, G. (2009). Is a “star” worth a thousand words? *European Journal of Marketing*, 43(11/12), 1269–1280. <https://doi.org/10.1108/03090560910989876>
- Vallurupalli, V., & Bose, I. (2020). Exploring thematic composition of online reviews: A topic modeling approach. *Electronic Markets*, 30(4), 791–804. <https://doi.org/10.1007/s12525-020-00397-5>
- Vazquez, E. E. (2021). Effect of an e-retailer's product category and social media platform selection on perceived quality of e-retail products. *Electronic Markets*, 31(1). <https://doi.org/10.1007/s12525-020-00394-8>
- von Helversen, B., Abramczuk, K., Kopeć, W., & Nielek, R. (2018). Influence of consumer reviews on online purchasing decisions in older and younger adults. *Decision Support Systems*, 113(3), 1–10. <https://doi.org/10.1016/j.dss.2018.05.006>
- Xiao, B., & Benbasat, I. (2007). E-Commerce product recommendation agents: Use, characteristics, and impact. *MIS Quarterly*, 31(1), 137–209. <https://doi.org/10.2307/25148784>
- Xu, P., Chen, L., & Santhanam, R. (2015). Will video be the next generation of e-commerce product reviews? Presentation format and the role of product type. *Decision Support Systems*, 73, 85–96. <https://doi.org/10.1016/j.dss.2015.03.001>
- Yelp. (2020). *Yelp open dataset: An all-purpose dataset for learning*. Retrieved from <https://www.yelp.com/dataset>
- Zhang, K. Z. K., Zhao, S. J., Cheung, C. M. K., & Lee, M. K. O. (2014). Examining the influence of online reviews on consumers' decision-making: A heuristic-systematic model. *Decision Support Systems*, 67, 78–89. <https://doi.org/10.1016/j.dss.2014.08.005>
- Zinko, R., Stolk, P., Furner, Z., & Almond, B. (2020). A picture is worth a thousand words: How images influence information quality and information load in online reviews. *Electronic Markets*, 30(4), 775–789. <https://doi.org/10.1007/s12525-019-00345-y>

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