Association for Information Systems

AIS Electronic Library (AISeL)

ICIS 2019 Proceedings

Crowds, Social Media and Digital Collaborations

Head over Feels? Differences in Online Rating Behavior for Utilitarian and Hedonic Service Aspects

Janina Seutter
University, janina.seutter@wiwi.uni-paderborn.de

Jürgen Neumann
Paderborn University, juergen.neumann@wiwi.uni-paderborn.de

ICIS 2019 Munich
Head over Feels? Differences in Online Rating Behavior for Utilitarian and Hedonic Service Aspects

Short Paper

Janina Seutter
Paderborn University,
Warburger Str. 100
33098 Paderborn, Germany
Janina.Seutter@uni-paderborn.de

Jürgen Neumann[1]
Paderborn University,
Warburger Str. 100
33098 Paderborn, Germany
Jürgen.Neumann@uni-paderborn.de

Abstract

Online reviews play a considerable role in reducing the information asymmetry between sellers and potential consumers. Despite the rich body of literature on online reviews, little is known about how the chosen content of reviews influences the rating behavior. As products or services offer more than one possible evaluation characteristic, different reviews on a product or service refer to different characteristics. In our research-in-progress we investigate to what extent the valence of online ratings differs depending on whether the rating refers to utilitarian characteristic or hedonic characteristics. To answer this question, we crawled 55,601 customer reviews on Google Maps of visits to 149 German theaters and classified each review as being primarily utilitarian, hedonic, or ambiguous. For our dataset we can determine that reviews with hedonic content are on average 0.48 stars higher rated than utilitarian reviews. Our results carry substantial managerial implications for designers of review platforms and customers.

Keywords: Online reviews, review content, hedonic, utilitarian

Introduction

Online reviews have become a popular source of information which reduces the information asymmetry between sellers and potential consumers. Accordingly, online ratings help the consumer decide whether to purchase a particular good or service (Wu et al. 2015). To compare products with each other, consumers rely on reviewers to give an overall rating on all the features of a product. It has been shown that the valence of online ratings (e.g., in the form of an average rating) causally influences sales (Chevalier and Mayzlin 2006). However, reviewers often focus on individual characteristics in their reviews. For instance, they could leave a low rating for a restaurant just because of a long waiting time. If many reviews of *restaurant A* focus on long waiting times, whereas reviews of *restaurant B* focus on the quality of the food, it makes it more difficult to compare the two restaurants. The fact that review systems often apply multidimensional rating systems to ensure that reviewers give individual ratings on several product or service features amplifies this problem (Chen et al. 2018). Reviewers are nevertheless free to decide on the content of their review. Even in the presence of multidimensional ratings, reviewers may predominantly choose to write about long waiting times, for instance. Consequently, reviewers self-select a specific review content which determines the overall reputation of a product or service. The self-selection of consumers has been shown

[1] This work has been partially supported by the German Research Foundation (DFG) within the Collaborative Research Center 901 “On-The-Fly Computing” under the project number 160364472-SFB901.

to be a driver of online ratings (i.e., the overall average rating) which potentially affects the effectiveness of online reviews (Hu et al. 2017; Li and Hitt 2008). Therefore, understanding the reasons for potential drivers behind self-selection and knowing how to mitigate them has received substantial interest from both academia and industry (Gutt et al. 2019).

In this paper we aim to analyze potential drivers of online ratings arising from reviewers choosing different review contents. While one reviewer could concentrate on the atmosphere they have experienced during their last restaurant visit, another could point out the long time they had to wait for their meal. In this case, the former considers rather hedonic characteristics of the service whereas the latter provides information on rather utilitarian aspects. The literature has identified various motives that drive reviewing behavior such as venting negative feelings, concern for other consumers, extraversion, positive self-enhancement, or social benefit (Hennig-Thurau et al. 2004). Thus, when consumers decide on the content of their reviews, they are also driven by certain motives. To this end, we aim to investigate the relationship between choice of review content and rating behavior by answering the following research question:

Do online ratings differ for reviews with primarily hedonistic characteristics compared to reviews with primarily utilitarian characteristics?

To answer our research question, we derive a hypothesis from theory of self-enhancement applied to the behavior of consumers (Berger and Iyengar 2013). We argue that reviews featuring predominantly hedonic characteristics have a higher rating than reviews featuring primarily on utilitarian characteristics because consumers are driven by positive self-enhancement when reviewing hedonic aspects. To test our hypothesis, we examine a comprehensive data set consisting of 55,601 Google Maps reviews for theaters in Germany. Using a text-mining approach we distinguish between reviews that comment on the play itself (i.e., hedonic characteristics) and those that consider the theater's additional offers such as cloakroom or parking facilities (i.e., utilitarian characteristics). We find support for our hypothesis from an ordinary least squares regression with theater-year-level fixed effects. In particular, our preliminary results indicate that reviews focusing on hedonic characteristics are associated with an increase in the corresponding rating by 0.48 stars compared to reviews focusing on utilitarian characteristics.

To the best of our knowledge, our study is the first to empirically scrutinize the impact of specific review content, identifying whether the review primarily considers a hedonic or a utilitarian feature of a product or service, and to compare their respective ratings. This research in progress provides managerial implications for designers of online review systems and consumers. Consumers making a purchase decision need to be aware that depending on the review content, ratings might be positively biased due to the self-enhancement behavior of consumers. System designers can mitigate potential biases resulting from self-enhancement behavior of reviewers by aggregating reviews with a hedonic content and those with a utilitarian content separately.

Related Literature and Theoretical Background

This study is part of a series of studies exploring drivers of online review valence (Gutt et al. 2019). This stream of literature analyzes factors that influence the valence (i.e., rating or average rating) of online reviews. For instance, consumer expectations are seen to drive the magnitude of online ratings because ratings are often biased if expectations are not met (Ho et al. 2017). Similarly, consumer preferences affect ratings because consumers with a higher preference for a product tend to buy it and subsequently provide a high rating (Li and Hitt 2008). The valence of prior ratings can also influence reviewing behavior. If prior ratings are positive and the reviewer plans to provide a negative rating, they tend to adjust their own rating to align with the existing ones (Muchnik et al. 2013). Finally, consumers are more likely to provide reviews if they have extreme experiences, which leads to an increase in the number of very high and very low ratings (Hu et al. 2017). Consequently, a series of drivers of online review valence have been studied in the literature. However, to the best of our knowledge, no study so far has investigated the effect of how a reviewer's choice of review content might affect their reviewing behavior.

Based on our review of the literature we posit that the reviewer's self-enhancement behavior explains how a review content focused on utilitarian or hedonic characteristics affects the review's rating. Therefore, our study is also related to research on the relationship between self-enhancement behavior as a motivation to write reviews. Research suggests that a consumer's need for signaling competence and superiority (i.e., self-enhancement) is a driver for providing reviews (Hennig-Thurau et al. 2004). Thus, self-enhancement is a driver of the number of reviews (i.e., volume). To the best of our knowledge, no research has as yet analyzed the role played by self-enhancement for determining the valence of reviews.

Self-Enhancement has been identified as one of the main motivations for writing an online review (Hennig-Thurau et al. 2004). The idea behind this notion is that, by writing a recommendation, a person gains by attracting attention to themselves, thereby showing connoisseurship, giving the impression of possessing insider information, and asserting superiority and status (Hennig-Thurau et al. 2004). Literature supports the idea that positive product reviews signal the competence of the reviewer and therefore enhance the reviewer in the eye of the reader (Chen and Lurie 2013; Mizerski 1982). As one can control the decision over which products to buy, giving negative feedback implies this reviewer's ineptitude as a consumer to make good purchase decisions (Angelis et al. 2012). Goals of self-enhancement behavior are the positive representation of the self in social interaction and positive recognition by others (Angelis et al. 2012). Previous studies have considered whether the evaluation behavior changes when consumers evaluate something they have experienced or when they transmit the evaluation of a third person. It has been observed that self-enhancement can lead consumers to generate positive evaluations of their own experiences but transmit negative evaluations of the experiences of others (Angelis et al. 2012). One explanation for transmitting of negative experiences of others is that, the more negative the performance and experiences of others are, the better people feel about themselves (Tesser 1988). Prior research states that the consumption of a hedonic good raises a need for justification (Okada 2005). Thus, self-enhancement behavior might be amplified by the characteristic (hedonic or utilitarian) of the reviewed good or service. The consumption of a hedonic good is associated with a sense of guilt and the difficulty to quantify the benefits of the good when comparing it to a utilitarian good (Okada 2005). Hedonic goods are characterized by an esthetic, intangible and subjective aspect of consumption (Hirschman and Holbrook 1982). In contrast to this, the consumption of a utilitarian good is more cognitively driven, instrumental, goal oriented, and accomplishes a functional task (Strahilevitz and Myers 1998). Even though goods are often categorized to be either hedonic or utilitarian this is a simplification, as goods or services are never entirely hedonic or utilitarian but contain both hedonic and utilitarian features (see Figure 1), e. g. a car has the utilitarian feature *gas mileage* and the hedonic feature *design* (Dhar and Wertenbroch 2000). When evaluating a car, gas mileage can be objectively measured, potential readers can more easily agree on a negative review regarding a high gas mileage, whereas it is arguably harder to agree on a negative review on the car's design.

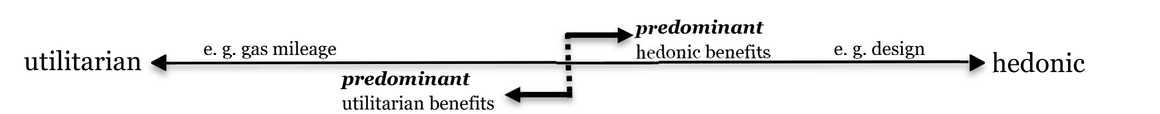

Figure 1. Continuum of hedonic and utilitarian characteristics

A consumption object is placed on both a utilitarian dimension and a hedonic dimension and these benefits in combination lead to an *overall goodness of a consumer good* (Batra and Ahtola 1991). The multidimensionality of the hedonic dimension of a good – i.e. the aspects of consumption which primarily fulfill a hedonic purpose – often leaves even experts to rely on their intuition when evaluating the good (Holbrook and Schindler 1994). Therefore, the rating of a rather hedonic good will be more subjective than the rating of a rather utilitarian good which satisfies well-defined needs (Holbrook and Schindler 1994). Accordingly, the evaluation of a product may address either (mostly) hedonic, (mostly) utilitarian, or a mixture of both features. In line with this, based on the heightened desire for justification after the consumption of a hedonic good and the lack of objective quality criteria, we argue that reviews concerning a hedonic feature of a good or service are more affected by self-enhancement behavior than reviews on a utilitarian feature. Accordingly, we formulate our hypothesis as follows.

Hypothesis: Ratings of reviews which comment predominantly on hedonistic aspects of a good or service are more positive than ratings for reviews which comment more on the utilitarian aspects of a good or service.

Empirical Analysis

Research Environment

To provide a first analysis regarding the research question, we analyze reviews for German theaters posted on Google Maps. There are two main reasons as to why Google Maps data on theaters is very well suited for dealing with the research question of this study. First, to this day there is no established platform in Germany that allows to write online reviews for theater plays.[2] As a result, the reviews on Google Maps cover different aspects of the services provided by the theater e.g. one review might focus on the play while another focuses on the cloakroom facility. Second, the service provided by a theater is complex and includes many supplementary services, which can also be included in the evaluation of a theater visit. Naturally, a theater visit is of predominantly hedonic nature, which might suggest a disadvantage when compared to other domains, but it also bears an advantage for classifying review texts. Using restaurant or laptop reviews, for instance, makes the classification into rather utilitarian or rather hedonic reviews more complex because it is harder to determine whether a product service feature is hedonic or utilitarian. For example, food temperature might be seen as hedonic by some consumers but as utilitarian by others. This would require supplementary studies to identify which features are rather utilitarian or rather hedonic for the restaurant domain. For theater visits, this is arguably easier. Since they are inherently hedonic, the play or selection of plays as a central component of this visit should be hedonic as well, whereas additional services like the availability of cup holders for example should be rather utilitarian (Figure 2).

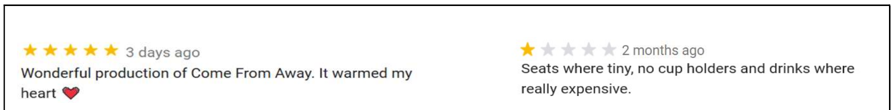

Figure 2. Example of Hedonic and Utilitarian Reviews

Lacking an established platform to review plays, the reviewers choose to focus either on a hedonic or on a utilitarian aspect of the theater visit. Following the definition of hedonic goods, for example, the theater play is mainly hedonic, as it contains an esthetic, intangible and subjective aspect of consumption (Charters 2006). A review of the play means that the author places her emphasis on a hedonic aspect of the service, while reviews concerned with, for example, the ticket sale, cloakroom, toilets, car parking or the seating are focused on a utilitarian service aspect. Even though, other features apart from the play of the theater could also be seen as hedonic, we argue that the play is a highly subjective aspect and evaluations of the play should be influenced by self-enhancing and justifying behavior (Holbrook and Schindler 1994) and thus have higher ratings. Thus, the reviews can be assigned to one of two categories: primarily hedonic or primarily utilitarian.

Data

We used a customized web crawler on March 13th 2019 to obtain a comprehensive data set from Google Maps consisting of all available reviews of visits to 149 German theaters. Every Google Maps user is able to provide online reviews, photos and further basic information about any location. Each review contains a star rating from 1 to 5 and an optional textual component (i.e., comment). The data set consists of 55,601 reviews that had been posted between January 2011 and the middle of March 2019.

Ratings without a textual component were excluded from further consideration, because a content analysis of reviews can only be performed for reviews with text. To compute our main variable of interest, the remaining comments were classified into the categories *hedonic*, *utilitarian,* and *ambiguous*. The classification was done with the support of Latent Dirichlet Allocation (LDA). LDA is a widely-used unsupervised machine learning method that can identify topics in large collections of documents with written text such as online reviews. The key idea behind LDA is that the reviewing authors compose online reviews by first deciding about a discrete distribution of topics T to write about, relying on words from a

[2] For instance, show-score.com is an evaluation platform which offers expert and consumer reviews for theatrical performances and shows for Greater New York City.

discrete distribution of words that are typical for the chosen topic. Put differently, a document is defined by a probability distribution over a fixed set of topics and each topic is defined by a probability distribution over a limited set of words (Debortoli et al. 2016). For each topic from the fixed set of topics, the LDA assigns a probability between 0 and 1 to each document, indicating how likely it is that this particular document belongs to a certain topic. Before the LDA was carried out, a preprocessing was performed, which included the removal of standard stop words, stemming, lemmatizing, the removal of HTML tags and the removal of numbers. We set the algorithm to create T=20[3] different topics from all comments. Subsequently, we assigned the topics to one of the three categories described above on the basis of the most probable words associated with the topic. Two of the authors first categorized the topics independently, before deciding on a final categorization. Of all 20 topics, five topics were assigned to the category hedonic, six to the category utilitarian and nine to ambiguous. For example, the most probable words for Topic 2 are *play*, *child*, *theatrical performance* and *production*. As these words indicate that the comments refer to the theater play, Topic 2 was assigned to the category of *hedonic reviews*. As the text mining tool gives a probability for each comment to belong to one of the 20 topics, it is possible to generate groups of comments related by content. For each review, we observe the star-rating it has received ($COMMENT\_RATING$). Each review is assigned to one of the 20 topics based on the highest probability given by the LDA. The dummy variables $COMMENT\_HEDONIC$ and $COMMENT\_UTILITARIAN$ take on the value of 1 if the review's topic was previously identified as either hedonic or utilitarian. We are aware of the limitations associated with the use of the LDA analysis that are pointed out by R2. To increase the credibility of the LDA classification we conducted a human coding and calculated the interrater reliability. The human coding sample included 50 randomly selected online reviews. The comments were classified to be either hedonic, utilitarian or ambiguous by two human coders and the unsupervised LDA classification algorithm. The results suggest that there is a good interrater reliability for hedonic classification (α_h = .857) and acceptable interrater reliability for utilitarian classification (α_u = .681).

Table 1 provides descriptive statistics for the variables of our dataset on a review level.

Table 1. Descriptive Statistics

Variable	Num.	Mean	Std. Dev.	Min.	Max.
$COMMENT\_RATING$	22,923	4.585962	0.8334722	1	5
$COMMENT\_HEDONIC$	7,994	1	0	0	1
$COMMENT\_UTILITARIAN$	4,386	1	0	0	1
$REVIEW\_HELPFUL$	22,923	0.1564368	0.5154035	0	16
$REVIEW\_PHOTO$	22,923	0.061772	0.2421348	0	1
$REVIEW\_MANAGEMENT$	22,923	0.0155739	0.1238225	0	1
$REVIEW\_LENGTH$	22,923	115.3195	180.0788	1	6,785
$AUTHOR\_PHOTOS$	22,923	204.4928	1409.094	0	7,8145
$AUTHOR\_REVIEWS$	22,923	72.21247	166.0558	1	5,920

$COMMENT\_HEDONIC$ and $COMMENT\_UTILITARIAN$ are our main variables and capture whether the review focuses on aspects that are more prone to self-enhancement and justification. In total there are 22,923 reviews with text, of which 7,994 were identified as *hedonic,* 4,386 as *utilitarian* and 10,543 as *ambiguous*. Reviews that are not classified as hedonic or utilitarian by the LDA algorithm fall into the category ambiguous. The variables REVIEW_HELPFUL, REVIEW_PHOTO, REVIEW_MANAGEMENT and REVIEW_LENGTH contain information on the individual review while the variables AUTHOR_PHOTOS and AUTHOR_REVIEWS contain information on the author that has written a certain

[3] For future versions of this work, we plan to determine the number of topics computationally, following Tirunillai and Tellis (2014).

review. We incorporate this additional available information in our analysis as it has been shown that it influences online review behavior. In prior research the helpfulness score (REVIEW_HELPFUL) (Chen and Lurie 2013), the attachment of a photo (REVIEW_PHOTO) (Karimi and Wang 2017), the existence of a management response to the review (REVIEW_MANAGEMENT) (Proserpio and Zervas 2017) and review length (REVIEW_LENGTH) (Kuan et al. 2015) has been identified to be influential on the interpretation of online reviews. Also, the experience of a reviewer has an effect on the review behavior and whether a review is trusted or not. In our analysis the experience of a reviewer is measured by the number of photos (AUTHOR_PHOTOS) and the number of reviews posted on Google Maps (AUTHOR_REVIEWS) (Kuan et al. 2015). In our subsequent analysis, we restricted our sample to only those reviews that were identified as either hedonic or utilitarian, which left us with 12,380 reviews for our main analysis. The histogram in Figure 3 displays the star-rating distribution of *hedonic* and *utilitarian* reviews. This descriptive figure reveals that *hedonic* reviews have a higher occurrence of five-star ratings than *utilitarian* reviews.

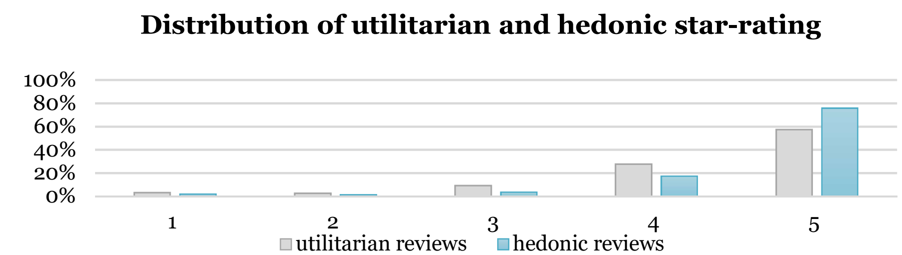

Figure 3. Star-rating distribution

Empirical Model and Preliminary Results

To test our hypothesis, we estimate the empirical model in Equation (1).

$$\begin{aligned} Y_{ijt} = {} & \beta_0 + \beta_1 COMMENT\_HEDONIC + \beta_2 REVIEW\_HELPFUL + \beta_3 REVIEW\_PHOTO \\ & + \beta_4 REVIEW\_MANAGEMENT + \beta_5 REVIEW\_LENGTH + \beta_6 AUTHOR\_REVIEWS \\ & + \beta_7 AUTHOR\_PHOTOS + \delta_{jt} + \epsilon_{ij} \end{aligned} \tag{1}$$

In this Ordinary Least Squares (OLS) regression Y_{ij} represents the outcome variable for the review written by reviewer i for theater j in year t. We examine COMMENT_RATING as the outcome variable. y_{ij} is a vector that contains review-specific control variables as presented in Table 1. Reviewing certainly depends on the theater that is being rated. If reviewers rate a theater with generally high quality e. g. a large selection of plays, modern equipment or an attractive interior, they will probably give higher ratings than for low-quality theaters with a more limited program or outdated interior. Additionally, the theaters differ in their location, e. g. the size of the city or the number and range of alternative offers for cultural goods in the vicinity. To control for these unobservable time-constant theater characteristics, we introduce theater-year-fixed effects δ_{jt}. ϵ_{ij} denotes the random unobserved error term.

Table 2 presents our preliminary results. Column (1) shows the results of our basic regression model using Equation 1. The coefficient $\beta 1$ is statistically significant and positive, suggesting that a hedonic review is positive by 0.26 stars than a more utilitarian review. Hence, we find support for our hypothesis.

In order to analyze our results in more detail and test their robustness, three further modifications to our empirical model were estimated. The model in Column (2) contains the condition of a minimum probability distance. In order to select comments that are identified as rather hedonic or rather utilitarian as clearly as possible, only comments that show a distance between the topic with the highest probability and the topic with the second highest probability of at least 30% were considered. Again, $\beta 1$ is statistically significant and positive, and the difference in the star rating between more hedonic and more utilitarian reviews rises to 0.39. In Column (3) a review must contain at least 30 characters to be considered in the model. Some reviews contain only smileys or meaningless reviews like "great theater". However, in some cases LDA will assign these comments to either the hedonic or the utilitarian topic category. To avoid this potentially distorting effect from happening, the Column (3) variation excludes comments with fewer than 30 characters. For the third model $\beta1$ is statistically significant and positive, and the difference of 0.32 stars between hedonic and utilitarian reviews is smaller than in the second but bigger than in the basic model. Column (4) contains the results of the combination of both a minimum probability distance of 30% and a minimum number of characters of 30 characters. To even further strengthen the selection criteria for the classification into hedonic and utilitarian comments we changed the distance between the topic probability from 30% to 70% in our fifth model (Column (5)). In line with our hypothesis, when a review refers to hedonic feature of the rated service, the rating is significantly more positive than when a review refers to a utilitarian feature. In the combined model which only considers comments containing at least 30 characters and a 70% difference between the 1st and 2nd topic, a review that refers to a hedonistic service feature is on average 0.48 stars higher than a review that refers to a utilitarian service feature. The biggest difference can be observed in model (5), which has the strictest selection criteria for the assignment into the hedonic and utilitarian review categories. In order to compare the rating difference of 0.26 to 0.48-stars with the results of studies that investigate drivers of online ratings, the coefficient range of the standard deviation is calculated. In our study, the coefficient range lies between 32% for (1) and 57% for (5). This is a comparatively strong result, as the coefficients found in other studies range below these figures (9-23% for Proserpio and Zervas 2017; 12-35% for Chen et al. 2018).

Table 2. Preliminary Results

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Dependent Variable	COMMENT_RATING	COMMENT_RATING	COMMENT_RATING	COMMENT_RATING	COMMENT_RATING
COMMENT_HEDONIC	0.26258***	0.39269***	0.31826***	0.41232***	0.47920***
	(0.01711)	(0.02768)	(0.02044)	(0.02986)	(0.05257)
REVIEW_HELPFUL	-0.03299	-0.01246	-0.04282*	-0.01342	0.01988
	(0.02065)	(0.03367)	(0.02298)	(0.03518)	(0.05370)
REVIEW_PHOTO	0.11070***	0.14556***	0.11991***	0.14706***	0.23743***
	(0.02867)	(0.04327)	(0.03199)	(0.04636)	(0.08418)
REVIEW_MANAGEMENT	-0.51896***	-0.49726***	-0.53295***	-0.49064***	-0.31414
	(0.13013)	(0.18076)	(0.14133)	(0.18850)	(0.35658)
REVIEW_LENGTH	-0.00056***	-0.00078***	-0.00047***	-0.00072***	-0.00090***
	(0.00008)	(0.00012)	(0.00008)	(0.00012)	(0.00028)
AUTHOR_REVIEWS	-0.00001	-0.00007	0.00003	-0.00009	-0.00001
	(0.00006)	(0.00012)	(0.00007)	(0.00013)	(0.00029)
AUTHOR_PHOTOS	0.00001	0.00001*	0.00001	0.00002***	0.00001
	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00005)
Constant	4.45194***	4.40191***	4.38906***	4.37344***	4.29984***
	(0.01738)	(0.02818)	(0.02193)	(0.03108)	(0.05530)
Theater-Year-level fixed effects	✓	✓	✓	✓	✓
Observations	12,380	4,960	9,577	4,516	1,681
R-squared	0.10418	0.18175	0.11529	0.18474	0.28726

Our results suggest that, depending on the rated feature, if the reviewer comments more on a hedonic rather than a utilitarian feature, the associated rating of their review seems to be higher. Thus, this supports our notion that self-enhancement is amplified for reviews commenting on hedonic characteristics of products or services.

Conclusion and Next Steps

Online reviews have become an essential feature for consumer decision-making. Therefore, understanding the factors that drive the valence of online reviews such as the social influence bias (Muchnik et al. 2013) or the underreporting bias (Hu et al. 2017) is crucial to improve their effectiveness in supporting consumers. In this study, we analyze how reviews focusing on either utilitarian or hedonic characteristics of a product or service differ in their ratings (e.g., for a restaurant visit, waiting time versus atmosphere). Conducting an ordinary least squares regression with fixed effects on a large data set of Google Maps reviews for theaters, our preliminary results suggest that reviews focusing more on hedonic aspects (e.g., the performance of the play) are associated with a 0.48-star higher rating than those focusing more on utilitarian aspects (e.g., the cloakroom or parking near the theater). Drawing on theory of self-enhancement, we hypothesize that our results are driven by the consumer's need to signal competence and superiority (Angelis et al. 2012; Chen and Lurie 2013). We postulate that consumers are hesitant in giving low ratings for hedonic characteristics as this could signal their poor ability in choosing and judging hedonic experiences such as a theater play. In contrast, reviews on utilitarian aspects such as poor car parking facilities should be more likely to result in low ratings, since potential readers should be more likely to agree with such opinions. To the best of our knowledge, we are the first to investigate how a reviewer's choice of review content might affect their reviewing behavior and thus contribute to the literature on the drivers of online ratings. The key feature of this study, that sets our work apart from previous studies examining self-enhancement behavior, is that the role of self-enhancement for online reviews is refined by distinguishing reviews by their content as being more focused on hedonic or utilitarian features. We contribute to the literature on the evaluation of hedonic and utilitarian aspects by providing empirical support for differences in evaluation behavior depending on the nature of the evaluated aspect. Consequently, even for the same good, the impact of self-enhancement and justification might differ depending on the evaluation criteria. Our preliminary results also provide practical implications for both consumers and designers of online review systems, suggesting that the hedonic or utilitarian focus of a review accounts for potential differences in ratings. The online marketplace Amazon already offers its customers machine learning based star ratings for certain product features derived from review texts, e.g. it provides separate star ratings for features like the fingerprint reader or battery life in addition to the overall rating. Our results suggest that differentiating between hedonic and utilitarian features of a product in this setting would also be helpful for customer decision making.

We plan to extend this research-in-progress in five major ways. First, we aim to validate our identification of hedonic- and utilitarian-focused reviews by letting independent human coders analyze a sample of our data set e.g., as in (Poniatowski et al. 2019). Second, we plan to perform a comprehensive text-mining and sentiment analysis on the review texts to investigate whether reviews on hedonic characteristics are written in a more self-centered manner. In this way, we will be able to provide further insights on whether self-enhancement behavior is driving our results. Third, reviewers with a positive and optimistic attitude might give systematically more positive reviews than reviewers with a negative and pessimistic outlook. To account for such reviewer heterogeneity in our empirical analysis, we plan to enhance our data set so that we are able to include reviewer-level fixed effects. Fourth, since theater visits are inherently hedonic we would also like to extend our findings to other domains, such as restaurants and laptops. Finally, as the reviewer's decision on the review's content is endogenous, we aim to conduct a supplementary experiment in which we nudge reviewers towards writing about either hedonic or utilitarian aspects and subsequently analyze their reviewing behavior. Nudging reviewers towards a certain review content (e.g., by providing review templates) should help to mitigate concerns regarding the endogeneity of our results.

References

Angelis, M. D., Bonezzi, A., Peluso, A. M., Rucker, D. D., and Costabile, M. 2012. "On braggarts and gossips: A self-enhancement account of word-of-mouth generation and transmission," *Journal of Marketing Research* (49:4), pp. 551–563.

Batra, R., and Ahtola, O. T. 1991. "Measuring the hedonic and utilitarian sources of consumer attitudes," *Marketing letters* (2:2), pp. 159–170.

Berger, J., and Iyengar, R. 2013. "Communication channels and word of mouth: How the medium shapes the message," *Journal of consumer research* (40:3), pp. 567–579.

Charters, S. 2006. "Aesthetic products and aesthetic consumption: A review," *Consumption, Markets and Culture* (9:3), pp. 235–255.

Chen, P.-Y., Hong, Y., and Liu, Y. 2018. "The Value of Multidimensional Rating Systems: Evidence from a Natural Experiment and Randomized Experiments," *Management Science* (64:10), pp. 4629–4647.

Chen, Z., and Lurie, N. H. 2013. "Temporal contiguity and negativity bias in the impact of online word of mouth," *Journal of Marketing Research* (50:4), pp. 463–476.

Chevalier, J. A., and Mayzlin, D. 2006. "The effect of word of mouth on sales: Online book reviews," *Journal of Marketing Research* (43:3), pp. 345–354.

Debortoli, S., Müller, O., Junglas, I. A., and Vom Brocke, J. 2016. "Text mining for information systems researchers: an annotated topic modeling tutorial," *CAIS* (39), p. 7.

Dhar, R., and Wertenbroch, K. 2000. "Consumer choice between hedonic and utilitarian goods," *Journal of Marketing Research* (37:1), pp. 60–71.

Gutt, D., Neumann, J., Zimmermann, S., Kundisch, D., and Chen, J. 2019. "Design of review systems–A strategic instrument to shape online reviewing behavior and economic outcomes," *The Journal of Strategic Information Systems*.

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., and Gremler, D. D. 2004. "Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet?" *Journal of interactive marketing* (18:1), pp. 38–52.

Hirschman, E. C., and Holbrook, M. B. 1982. "Hedonic consumption: emerging concepts, methods and propositions," *Journal of marketing* (46:3), pp. 92–101.

Ho, Y.-C., Wu, J., and Tan, Y. 2017. "Disconfirmation effect on online rating behavior: A structural model," *Information Systems Research* (28:3), pp. 626–642.

Holbrook, M. B., and Schindler, R. M. 1994. "Age, sex, and attitude toward the past as predictors of consumers' aesthetic tastes for cultural products," *Journal of Marketing Research*, pp. 412–422.

Hu, N., Pavlou, P. A., and Zhang, J. J. 2017. "On Self-Selection Biases in Online Product Reviews," *MIS quarterly* (41:2), pp. 449–471.

Karimi, S., and Wang, F. 2017. "Online review helpfulness: Impact of reviewer profile image," *Decision support systems* (96), pp. 39–48.

Kuan, K. K. Y., Hui, K.-L., Prasarnphanich, P., and Lai, H.-Y. 2015. "What makes a review voted? An empirical investigation of review voting in online review systems," *Journal of the Association for Information Systems* (16:1), p. 48.

Li, X., and Hitt, L. M. 2008. "Self-selection and information role of online product reviews," *Information Systems Research* (19:4), pp. 456–474.

Mizerski, R. W. 1982. "An attribution explanation of the disproportionate influence of unfavorable information," *Journal of consumer research* (9:3), pp. 301–310.

Muchnik, L., Aral, S., and Taylor, S. J. 2013. "Social influence bias: A randomized experiment," *Science* (341:6146), pp. 647–651.

Okada, E. M. 2005. "Justification effects on consumer choice of hedonic and utilitarian goods," *Journal of Marketing Research* (42:1), pp. 43–53.

Poniatowski, M., Neumann, J., Görzen, T., and Kundisch, D. (eds.) 2019. "Organizing Their Thoughts: How Online Review Templates Affect the Review Text," in *Proceedings of the Twenty-Seventh European Conference on Information Systems (ECIS), Research-in-Progress,* Stockholm, Sweden (Forthcoming).

Proserpio, D., and Zervas, G. 2017. "Online Reputation Management: Estimating the Impact of Management Responses on Consumer Reviews," *Marketing Science* (36:5), pp. 645–665.

Strahilevitz, M., and Myers, J. G. 1998. "Donations to charity as purchase incentives: How well they work may depend on what you are trying to sell," *Journal of consumer research* (24:4), pp. 434–446.

Tesser, A. 1988. "Toward a self-evaluation maintenance model of social behavior," in *Advances in experimental social psychology*: Elsevier, pp. 181–227.

Tirunillai, S., and Tellis, G. J. 2014. "Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation," *Journal of Marketing Research*, pp. 463–479.

Wu, C., Che, H., Chan, T. Y., and Lu, X. 2015. "The economic value of online reviews," *Marketing Science* (34:5), pp. 739–754.